Essays in Industrial Organization and Mechanism Design

Dissertation
der Wirtschaftswissenschaftlichen Fakultät
der Universität Zürich

zur Erlangung der Würde
eines Doktors der Ökonomie

vorgelegt von

Dennis Gärtner

von Deutschland

genehmigt auf Antrag von

Prof. Dr. A. Schmutzler
Prof. Dr. M. Wolf

Die Wirtschaftswissenschaftliche Fakultät der Universtität Zürich gestattet hierdurch die Drucklegung der vorliegenden Dissertation, ohne damit zu den darin ausgesprochenen Anschauungen Stellung zu nehmen.

Zürich, den 8. Februar 2006

Der Dekan: Prof. Dr. H. P. Wehrli

Contents

Chapter 3:
Are There Waves in Merger Activity After All? 83
Dennis Gärtner and Daniel Halbheer

Preface

In writing this dissertation, I have enjoyed the help and support of a number of people whom I would like to thank here.

First of all, I am indebted to Armin Schmutzler, my thesis advisor, for his guidance and support. Amongst many other things, I am particularly grateful for his introducing me to the field of 'robust comparative statics', which has not only been a valuable tool in thinking about the questions treated in Chapters 1 and 2 of this dissertation, but which I believe more generally has had a profound impact on my way of structuring and analyzing economic issues. Moreover, I am grateful to Michael Wolf for co-supervising this dissertation.

A substantial part of the research leading to Chapters 1 and 2 of this dissertation was conducted during my stay at the Institut d'Economie Industrielle (IDEI) at the University of Toulouse from March through August 2005. I am greatly indebted to Jean Tirole and the Swiss National Science Foundation for making this visit possible. Aude Schloesing of the University of Toulouse provided truly extraordinary assistance in organizing my visit.

The research presented in Chapter 3 is joint work with Daniel Halbheer. Collaborating with him was not only productive and inspiring—but also lots of fun. I would further like to express my thanks to Stefan Buehler, from who's advice I have profited on countless occasions while writing this dissertation. My sincere thanks also go to other former and present colleagues of mine: Zava Aydemir, Katrin Bernath, Johannes Binswanger, Andreas Polk, Dario Sacco, and Christian Stoff.

Finally, special thanks of course go to family and friends for allowing me to always count on their full support and encouragement during this endeavor.

Zürich, February 2006 Dennis Gärtner

Introduction

This dissertation deals with two topical issues in the field of industrial organization and regulation: corporate mergers and dynamic regulation. In three separate essays, we provide a positive analysis of three particular questions in these areas of research.

First, Chapter 1 takes on the issue of optimal dynamic regulation and procurement when the supplier's technology is (i) subject to learning effects, and (ii) privately known to the supplier. That private knowledge on technology is a key factor in regulatory practice has been a central theme in the literature for years (see the influential work of Laffont and Tirole, 1993). However, relatively little is still known about *dynamic* aspects of regulation and the interplay between regulation and innovation. This is particularly unfortunate because innovation plays a prominent role in many traditional fields of government regulation such as telecommunications or electric utilities. In such industries, adequate regulatory practice must invariably take account of its dynamic impact on innovation.

Chapter 1 contributes to closing this gap by considering a well-documented kind of innovation: innovation through 'learning by doing'. Under learning by doing, the regulator's task is to induce a level of production that takes careful account of future cost-savings induced through learning effects. However, this task is again complicated by firms enjoying superior knowledge on their technology. Hence, a key question is whether this private knowledge should lead the regulator to under- or overexploit learning effects relative to the full-information benchmark—that is, whether he induces inefficiently little or inefficiently much innovation. We find that the answer crucially depends on how firms' learning potential and their absolute level of efficiency are related: If learning leads inherently more efficient agents to expand their lead over less efficient agents, then learning effects will be underexploited. If, however, learning leads inherently less efficient agents to catch up, then learning effects will be overexploited.

To our knowledge, this insight—particularly the possibility of learning effects being overexploited—is new to the regulation and procurement literature. Although recent studies by Lewis and Yildirim (2002a,b) also investigate regulation and procurement under learning by doing, they conclude that private information should always lead to an underexploitation of learning effects. We argue in Chapter 1 that the reason for our difference in findings lies in Lewis and Yildirim's considering information only on the cost side of learning-by-doing and neglecting the fact that agents may hold private information also on their inherent learning rates.

Chapter 2 takes up another area of industrial organization in which private information allegedly plays an important role: corporate mergers. There is considerable evidence that mergers are often unprofitable. Moreover, there is anecdotal evidence abound describing merger cases in which at least one party expressed ex-post regret about the merger, in some cases even explicitly attributing this to having misjudged the merger partner.[1] In Chapter 2, we explore the extent to which these two observations may be explained by merger candidates holding superior information on their own firm. We consider a setting in which two potential merger partners each possess private information pertaining both to the profitability of the prospective merged entity and to each firm's stand-alone profits. Using a mechanism-design approach, we investigate the scope for merger negotiation processes ('merger mechanisms') which gather agents' dispersed information and reach a merger decision—subject to the constraint that the merger decision avoids outcomes involving ex-post regret on behalf of any of the merging parties. Importantly, what sets this problem apart from the classical problem of bilateral trade under asymmetric information (see Myerson and Satterthwaite, 1983) is that, in addition to using transfers and the trade-decision (the merger decision, in our case) to elicit agents' private information, the mechanism can condition agents' payoffs if trade occurs on the *ex-post value* of trade, by specifying how profits in the new entity are to be shared.

In this setting, we show that if agents' information is relevant only to profits of the prospective merged entity, then the fact that this information is privately dispersed poses no problem: the negotiation process will lead agents to reveal this information and implement a merger if and only if the merger is profitable, and it will never result in any ex-post regret. As soon as private information is relevant also to firms' profits if the merger is *not* implemented, however, we obtain a markedly different result: In this case, *any* merger negotiation process must necessarily result in ex-post regret for *some* constellation of private information. The important lesson to be drawn from this second result is that generally, asymmetric information indeed poses a real problem to merger negotiations: Even if *jointly*, parties possess all the information to assess the profitability of a merger deal, under any ever so cleverly designed negotiation process, parties necessarily still run the risk of entering a merger deal which is *idiosyncratically* unprofitable if this information is dispersed across parties and relevant to their outside option. More generally, our results indicate that the problem which asymmetric information poses to efficient merger decisions may lie

[1] A telling case in point is the 1998 merger between Hypobank and Vereinsbank, in which "it took more than two years for Vereinsbank to discover the full horror of its partner's balance sheet" (The Economist, July 20th 2000, 'How mergers go wrong').

more in its *inhibiting* mergers that would be efficient than in its provoking mergers that are inefficient. This result, while in line with classical results on bilateral trade, suggests that the widespread 'merger-mania'-discussion may be focussing too much attention on the frequent occurrence of unprofitable mergers than on profitable merger opportunities that are foregone.

Finally, Chapter 3 investigates a further and perhaps even more widespread conventional wisdom concerning mergers which contends that mergers come in waves. Motivated by the recent surge in potential explanations of such a phenomenon,[2] our focus in Chapter 3 is to empirically (re-)evaluate the underlying merger-wave hypothesis using more recent data. Starting from Nelson's (1959) seminal observation that aggregate merger activity is characterized by "large bursts of activity separated by lengthy intervals of very low activity," we motivate a time-series model of merger activity in which waves are caused by switches in an unobserved state.

We find that such a model provides valuable insight into both US and UK merger activity over the last 30 years by clearly identifying wave periods (i.e., states). Somewhat surprisingly, however, while we obtain a sharp indication of a wave in US merger activity beginning in 1995, we find no evidence whatsoever of a wave in the 80s. This finding is at odds not only with conventional wisdom, but also with previous findings by Town (1992) and Linn and Zhu (1997) who identify an 80s merger wave using a similar model. We argue that this apparent contradiction is caused by a further key feature of our analysis: the use of more recent Bayesian estimation methods which take adequate account of uncertainty about the model's parameters rather than just uncertainty about the unobserved states.

Eventually, we hope that the update on the empirical facts concerning merger waves which we present in Chapter 3 will serve as a useful input for the ongoing discussion on their cause.

We conclude this overview with a note to the reader concerning the hierarchical structure of this dissertation, which reproduces the format familiar from handbook series in economics: Each chapter forms an independent unit, complete with its own appendix and list of references. Moreover, sections, propositions, equations, figures etc. in each chapter are numbered independently. Any references *across* chapters are made explicit. This format appeared most adequate and readable given the fact that the individual chapters are essentially self contained.

[2]We might add that, in keeping with the theme of the other two chapters, some of these theoretical explanations are again based on informational asymmetries (see Rhodes-Kropf and Viswanathan, 2004, for instance).

References

Laffont, Jean-Jacques, and Jean Tirole (1993) *A Theory of Incentives in Procurement and Regulation* (Cambridge, Mass.: MIT Press)

Lewis, Tracy R., and Huseyin Yildirim (2002a) 'Learning by doing and dynamic regulation.' *RAND Journal of Economics* 33(1), 22–36

— (2002b) 'Managing dynamic competition.' *American Economic Review* 92(4), 779–797

Linn, Scott C., and Zhen Zhu (1997) 'Aggregate merger activity: New evidence on the wave hypothesis.' *Southern Economic Journal* 64(1), 130–146

Myerson, Roger B., and Mark A. Satterthwaite (1983) 'Efficient mechanisms for bilateral trading.' *Journal of Economic Theory* 29(2), 265–281

Nelson, Ralph L. (1959) *Merger Movements in the American Industry, 1895–1956* (Princeton: Princeton University Press)

Rhodes-Kropf, Matthew, and S. Viswanathan (2004) 'Market valuation and merger waves.' *Journal of Finance* 59(6), 2685–2718

Town, Robert J. (1992) 'Merger waves and the structure of merger and acquisition activities.' *Journal of Applied Econometrics* 7, Supplement: Special Issue on Nonlinear Dynamics and Econometrics, S83–S100

Chapter 1

Incentive Contracts under Learning by Doing

Dennis Gärtner

1 Introduction

Supply relations in which a firm procures goods or services from another firm are typically ridden with problems of asymmetric information. The canonical case considered in the literature is that of a supplier enjoying private information concerning his marginal cost of production (see for instance Laffont and Martimort, 2002, for a textbook treatment). A characteristic feature of these models is that the supplier's marginal costs are assumed exogenously given. However, the production of many of the goods typically traded in such supplier relations involves learning effects in the sense that the supplying firm can lower its marginal costs as it gains experience with rising production volumes.[1] Such relations will therefore involve endogenous cost structures as the supplier's marginal costs come to depend not only on exogenous factors, but also on the endogenous volume of trade. This paper explores the impact of endogenizing marginal costs on optimal procurement: Does private information concerning suppliers' extent of learning effects lead to contracts that exploit learning

[1]Learning effects have been documented in numerous industries, such as in the production of airplanes (Wright, 1936; Alchian, 1963), ships (Rapping, 1965; Thompson, 2001; Thornton and Thompson, 2001), chemicals (Stobaugh and Townsend, 1975; Lieberman, 1984), machine tools (Hirsch, 1952, 1956), computers and semiconductors (Nye, 1996; Gruber, 1998), electrical equipment (Preston and Keachie, 1964; Sultan, 1975), nuclear power (Joskow and Rozanski, 1979; Roberts and Burwell, 1981; Lester and McCabe, 1993) and in the weapons industry (Fox, 1988; Gansler, 1989). While these studies all pertain to cost reducing learning effects, a recent strand of literature also seeks do document *qualitative* learning effects in production (see Moul, 2001).

effects only to an inefficiently low extent? More generally, how does the endogenous formation of agents' private information affect contractual arrangements?

Such issues arise not only in traditional vertical procurement relations. They concern also the question of how to regulate a monopolistic service supplier with private information on costs (Baron and Myerson, 1982; Laffont and Tirole, 1986). While we already have a fairly good understanding of optimal regulation in stationary settings (see the influential work of Laffont and Tirole, 1993), relatively little is still known about the impact of regulation in dynamic settings where suppliers' technology is endogenous.[2] Specifically, technological improvements through learning effects should play a significant role in electricity or telecommunication markets, for instance, making it important to understand the long-run interplay between regulation and innovation.

To tackle these issues, we set up a simple model of procurement over two periods, where in each period, the supplier produces and sells some amount of a good to the procuring firm in exchange for a monetary transfer. First-period marginal costs are exogenously given (and publicly known). However, to capture learning effects, second-period marginal costs are assumed to depend on the level of first-period production. Furthermore, the strength of this learning effect is assumed to be known only to the supplier.

Our main finding is that whether private information causes learning-effects to be over- or underexploited relative to the efficient benchmark crucially depends on how agents' learning rates differ. If learning causes inherently more efficient agents to extend their technological lead over less efficient agents, then quantities procured will underexploit learning effects. In this case, the principal's incentive to underexploit learning effects may even be so strong as to let first-period output fall short of its static optimum, implying that learning effects are not exploited at all. However, if learning causes inherently less efficient agents to catch up on more efficient agents, then the principal has an incentive to *overexploit* learning effects. This incentive not only leads to inefficiently high levels of output in the first period—it may also lead to an inefficiently high level of overall trade.

We first derive these results in full generality under the assumption of *full commitment*, implying that parties can sign a contract which settles both periods' exchanges at the beginning of the relationship, and—more importantly—that they can commit not to renege on this contract. In a further step we show, however, that our main findings do not crucially depend on this assumption by considering the opposite case of *spot commitment*, where the principal can commit only to exchanges in the current period. Even though limited commitment is generally understood to be detrimental to long-term investments, we show that if learning effects lead the

[2] A notable exception is Baron and Besanko (1984), who consider dynamic procurement relations in a full-commitment setting. While they allow for innovation in the supplier's technology by endogenous unobservable effort, they do not consider learning effects in production.

less efficient agent to catch up, overexploitation can result also in the case of spot commitment.

The basic intuition for why distortionary incentives depend on how agents' learning rates differ is simple: By the familiar rent-efficiency tradeoff, distortions in output are driven by the principal's incentive to limit the rent payable to more efficient agents, which in turn corresponds to their cost advantage over less efficient agents. If learning magnifies efficiency differences, the usual downward distortion results. However, if learning leads less efficient agents to catch up, the cost advantage enjoyed by more efficient agents decreases with first-period output, in which case upward distortions, and thereby inefficiently high investments into learning, result.

The paper closest in spirit to ours is Lewis and Yildirim (2002a). In the context of optimal regulation under spot commitment, the authors analyze how dynamic regulation deals with suppliers' learning effects under private information (in a companion paper, Lewis and Yildirim, 2002b, also apply these results to vertical procurement relations). One of the key conclusions is that optimal policy indeed encourages learning effects, but at an inefficiently low level. However, their model differs from ours in an important respect. Private information in their model pertains exclusively to the *cost* side of learning: while agent and principal have symmetric knowledge concerning the impact of higher output today on marginal costs tomorrow (i.e. on learning effects themselves), each period involves a transitory 'cost shock' which offsets production costs in that period only and which is known only to the agent. Hence, if we think of learning effects as inducing a cost-benefit tradeoff—where costs are inefficiently high output today from a static viewpoint and benefits are lower marginal costs tomorrow—then private information in their model concerns only the cost side of this tradeoff. In contrast, this paper focusses on the benefit side by letting agents enjoy private information on their learning technology.

An important contribution of this paper to the existing literature thus lies in showing that focussing merely on cost-side private information misses key issues concerning optimal regulation and procurement when learning effects are present. Particularly, it fails to identify the possibility of an *over*exploitation of learning effects and of an inefficiently high level of overall trade.

On a more basic level, this paper's contribution may be seen as sharpening basic economic intuition concerning the connection between asymmetric information and the volume of trade: It has become a virtual commonplace to associate private information with inefficiently low trade. Our results bring back to mind that this intuition depends crucially on the presumption that increased trade exacerbates the value of private information and thereby informational rents. While such a structure arises naturally in many models, we argue that learning by doing provides a case in point where it is just as natural for the reverse to be true.

The rest of this paper is organized as follows. Section 2 sets up a basic two-period model of learning by doing and describes the full-information benchmark.

Section 3 presents the optimal contract under asymmetric information and full commitment. We discuss both its efficiency properties and whether it even makes use of learning economies in the first place. Section 4 investigates contracts under spot commitment. We analyze the additional restrictions on output schedules relative to full commitment and show that the main insights of Section 3—particularly the possibility of inefficiently *high* trade—generalize to the case of spot commitment. Finally, Section 5 concludes with a brief discussion of the results and further possible applications.

2 A Simple Model of Learning By Doing

This section presents the basic model and, as a benchmark for our later analysis, characterizes the efficient allocation.

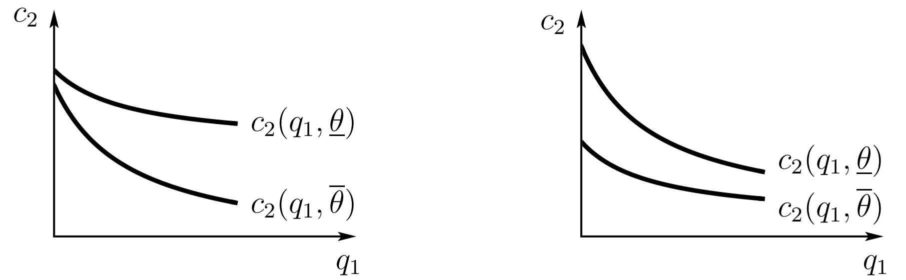
2.1 Setting up the Model

We consider a simple model in which a principal procures a good from an agent over two periods $t \in \{1, 2\}$. Let $q_t \in Q_t$ denote the amount of the good procured in period t and let $z_t \in \mathbb{R}$ denote the monetary transfer from the principal to the agent in that period. Unless stated differently, we let $Q_t \equiv \mathbb{R}_{\geqslant 0}$. In each period t, let the principal's utility be given by $v_t = S(q_t) - z_t$, where $S' > 0$ and $S'' < 0$. The principal discounts with $\delta \in (0, 1)$, overall utility from transactions over the two periods being $V = v_1 + \delta v_2$. The agent produces the good at a constant marginal cost $c_t > 0$ within each period, yielding a utility of $u_t = z_t - c_t q_t$ in each period t. The agent discounts with the same factor δ, leading to an overall utility of $U = u_1 + \delta u_2$. Both principal and agent are assumed risk neutral.[3]

The remaining assumptions detail the agent's production technology and its dependence on private information (we briefly discuss their relaxation in the context of contracts under full commitment in Section 3.5). To focus our analysis, we assume the agent's first-period marginal costs c_1 to be observable.[4] However, we let the agent possess private information concerning the structure of second-period costs. Private information is represented by the scalar θ (the agent's "type"), which is drawn from a commonly known distribution over Θ, and which the agent privately observes

[3]Letting principal and agent share the same risk attitude and time preference focusses our analysis by avoiding motives to trade risk or intertemporal utility.

[4]As we discuss in Section 3.5 in more detail, introducing private information on first-period costs c_1 introduces a second, competing motive for the principal to distort quantities in order to reduce *first-period* informational rent. This second motive in isolation already being well understood in the standard framework *without* learning, letting c_1 be observable serves to make this papers' contribution more transparent.



(a) More Efficient Agent Learns Faster.

(b) Less Efficient Agent Learns Faster.

Figure 1: Types of Learning Effects.

prior to first-period production (and prior to contracting). For simplicity, most of our analysis will assume two types $\Theta = \{\overline{\theta}, \underline{\theta}\}$ with $\overline{\theta} > \underline{\theta}$ and $\text{Prob}(\theta = \overline{\theta}) = \nu$.

Second-period marginal costs c_2 are a function of θ and first-period output q_1. We model the presence of learning effects in production by letting $\partial c_2/\partial q_1 < 0$: The higher first-period production, the lower the marginal costs of production in period 2 (for any given type θ).[5]

Next, we assume that c_2 is strictly decreasing in θ for all q_1. Thus, an agent with a higher θ is more efficient in that he produces any output schedule $\mathbf{q} = (q_1, q_2)$ at a lower cost. Note that this assumption represents more than a mere normalization of the type space Θ: Since it is assumed to hold for all q_1, it will provide a key sorting condition in our derivation of the optimal contract under asymmetric information.

We call $|\partial c_2/\partial q_1|$ the agent's *learning rate*, and say that an agent *learns faster* if he has a higher learning rate. Note that an agent may learn faster even though he is less efficient (i.e., has a lower θ). Indeed, key aspects of our analysis will crucially depend on whether learning rates increase or decrease in θ. To facilitate this, we assume that learning rates either increase or decrease in θ for all q_1. Figure 1 illustrates the relevant constellations.

The following example not only provides an illustration of the setup, but—due to its analytical tractability—will prove useful for numerical examples given further below:

Example 1. Let second-period costs be given by $c_2(q_1, \theta) = c(\theta) - \gamma(\theta) q_1$ with c strictly increasing in θ, and with $\gamma > 0$, and let the principal's objective function be given by $S(q_t) = aq_t - bq_t^2$, where $a, b > 0$. Then more efficient agents learn faster if γ is increasing in θ, whereas less efficient agents learn faster if γ is decreasing in θ.[6]

[5] Note that we assume marginal costs to be constant *within* each period but change discontinuously from one period to the next. This assumption serves to isolate learning effects from simple scale economies. Indeed, what distinguishes the two is that learning by doing depends on both previous production volumes *and* on time.

[6] To make this example entirely compatible with our assumptions, the range of permissible q_1

In what follows, we will consider various settings for negotiating the exchanged quantities and transfers. Common to these settings, however, is the usual assumption of full bargaining power resting with the principal: The principal offers a contract (or a choice of contracts) to the agent, which the agent can decide to accept or reject. If the agent rejects, he obtains a (type-independent) reservation payoff of zero and negotiations end.

2.2 The Efficient Full-Information Benchmark

As a point of comparison for our later analysis, we first consider the efficient full-information benchmark. Assume for a moment that the agent's type θ is known to the principal. For any output schedule $\mathbf{q} = (q_1, q_2)$, the joint surplus of trade is then given by

$$W(\mathbf{q};\theta) \equiv S(q_1) - c_1 q_1 + \delta[S(q_2) - c_2(q_1,\theta)q_2]. \tag{1}$$

Given his bargaining power and any known type θ, the informed principal will offer a contract $(\mathbf{q}^*(\theta), \mathbf{z}^*(\theta))$ which is efficient (i.e., first best), specifying production levels $\mathbf{q}^*(\theta) = (q_1^*(\theta), q_2^*(\theta))$ which maximize joint surplus $W(\mathbf{q};\theta)$, and payments $\mathbf{z}^*(\theta) = (z_1^*(\theta), z_2^*(\theta))$ which leave the agent his reservation utility.[7]

For later comparisons, we define *conditional* first-best output levels as follows. For any $(q_2;\theta)$, let $\hat{q}_1^*(q_2;\theta) \equiv \arg\max_{q_2} W(q_1, q_2;\theta)$ and, similarly, for any (q_1,θ) let $\hat{q}_2^*(q_1;\theta) \equiv \arg\max_{q_1} W(q_1, q_2;\theta)$.

Lemma 2.1. *The first-best output schedule $\mathbf{q}^*(\theta)$ and the contingent first-best output levels $\hat{q}_1^*(q_2;\theta)$ and $\hat{q}_2^*(q_1;\theta)$ have the following properties:*

(a) $\hat{q}_2^*$ is increasing in both q_1 and θ;

(b) $\hat{q}_1^*$ is increasing in q_2, and increasing (decreasing) in θ if more (less) efficient agents learn faster;

(c) $\mathbf{q}^*$ is increasing in θ if more efficient agents learn faster.

The proof of this and all later results is presented in the Appendix. As a general matter, we establish comparative static results such as Lemma 2.1 using supermodular analysis (cf. Milgrom and Roberts, 1990; Topkis, 1998). This approach exploits basic complementarity relations among arguments of the objective function and avoids imposing any unnecessary concavitity assumptions on objectives.[8] The

and q_2, (i.e., Q_1 and Q_2) must be bounded from above so as to ensure $S'(q_t) > 0$ and $c_2(q_1, \underline{\theta}) > c_2(q_1, \overline{\theta}) > 0$.

[7]Note that the first-best transfer schedule $\mathbf{z}^*$ will never be unique. Indeed, if $(\mathbf{q}^*, \mathbf{z}^*)$ is a first-best contract, then any contract $(\mathbf{q}^*, \tilde{\mathbf{z}})$ with the same discounted value of transfers (i.e., with $\tilde{z}_1 + \delta\tilde{z}_2 = z_1^* + \delta z_2^*$) will also be first-best.

[8]Nonetheless, readers unfamiliar with this technique will quickly verify the results under additional concavity assumptions by means of the first-order approach.

latter is particularly valuable in our setting due to the concavity which learning effects naturally introduce into the cost function.[9]

These technical issues aside, the intuition for the above results is conceivably simple: Higher first-period output lowers the cost of additional second-period output, thereby raising incentives to expand the latter. Conversely, higher second-period output raises incentives to lower that output's costs through learning effects by expanding first-period output. Thus, each period's conditionally efficient output rises in the other period's output level. Moreover, a higher θ makes additional second-period output less costly, which is why $\hat{q}_2^*$ is increasing in θ. Incentives to raise first-period output in turn rise in the agent's learning rate, which is why $\hat{q}_1^*$'s response to a change in θ depends on how the learning rate changes in θ. Finally, if more efficient agents also learn faster, these effects complement each other, making $\mathbf{q}^*$ rise in θ. Note that no such robust comparative result is available if the *less* efficient agent learns faster: a rise in θ then provides direct incentives to raise q_2 and lower q_1, which are counteracted however by the complementarity between q_1 and q_2, leaving the overall result ambiguous.

3 Contracts under Full Commitment

In contractual problems with investment characteristics, the outcome is generally sensitive to the level of intertemporal commitment available to the principal (see for instance Fudenberg et al., 1990). In this section, we investigate our problem of learning by doing under the most extreme form of commitment: We assume that at the start of period one (but after the agent has learned his type), the principal can offer a contract settling all future exchange which cannot be reneged on.

3.1 Characterizing the Optimal Contract

The full-commitment setting has the convenient property that, by the revelation principle and the stationarity of private information, we may equivalently restrict our attention to truth-revealing mechanisms of the type $\{\mathbf{q}(\tilde{\theta}), \mathbf{z}(\tilde{\theta})\}_{\tilde{\theta}\in\Theta}$ which specify

[9] Supermodular analysis extends also to situations where optimizers are not unique. In this case, comparative static results are interpretable in terms of ordering relations among *sets*. Although all our results permit such an interpretation, to avoid tedious notation, we will be somewhat loose in distinguishing between the *set* of optimizers and its individual elements.

Moreover, while comparative statics derived by this technique are very general in terms of covering also the possibility of corner solutions, this comes at the cost of all results applying only in a *weak* sense (i.e., "increasing" in Lemma 2.1 is to be read as "nondecreasing"). Since essentially all complementarity relations underlying our results are in fact *strict*, strict versions of our comparative static predictions are easily established for interior maximizers under mild additional conditions (see Edlin and Shannon, 1998, for technical details).

contracts (i.e., exchanged quantities and transfers) for each type, and where these contracts are designed so as to make it optimal for the agent to truthfully reveal his type. For any such contract, we let

$$U(\theta) \equiv z_1(\theta) - c_1 q_1(\theta) + \delta\big\{z_2(\theta) - c_2[q_1(\theta), \theta] q_2(\theta)\big\} \tag{2}$$

denote the θ-type's *equilibrium rent*.

To relax notation in this section, for any function of $\theta \in \{\overline{\theta}, \underline{\theta}\}$, we let an upper (lower) bar indicate that the function is evaluated at $\overline{\theta}$ ($\underline{\theta}$) and drop the argument θ. Thus, for instance, $\overline{U} \equiv U(\overline{\theta})$ and $\underline{U} \equiv U(\underline{\theta})$. Finally, we let

$$\Phi(\mathbf{q}) \equiv \delta q_2[\underline{c_2}(q_1) - \overline{c_2}(q_1)] \tag{3}$$

denote the *cost advantage* enjoyed by the $\overline{\theta}$-agent over the $\underline{\theta}$-agent for any output schedule $\mathbf{q} = (q_1, q_2)$. Intuitively, this cost advantage measures the *value* of private information enjoyed by the more efficient $\overline{\theta}$-type.

With this notation in place, the optimal contract under full commitment can be characterized as follows:

Proposition 3.1. *The menu of contracts offered by the uninformed principal under full commitment is such that production schedules $\overline{\mathbf{q}}^{\mathrm{SB}}$ and $\underline{\mathbf{q}}^{\mathrm{SB}}$ solve*

$$\overline{\mathbf{q}}^{\mathrm{SB}} \in \arg\max_{\mathbf{q}} \overline{W}(\mathbf{q}) \quad \textit{and} \quad \underline{\mathbf{q}}^{\mathrm{SB}} \in \arg\max_{\mathbf{q}} \big\{\underline{W}(\mathbf{q}) - \tfrac{\nu}{1-\nu}\Phi(\mathbf{q})\big\}. \tag{4}$$

Transfers $\overline{\mathbf{z}}^{\mathrm{SB}}$ and $\underline{\mathbf{z}}^{\mathrm{SB}}$ = are chosen such that types' equilibrium rents are $\overline{U} = \Phi(\underline{\mathbf{q}}^{\mathrm{SB}})$ and $\underline{U} = 0$.

This result is easily understood by recognizing that, despite the presence of learning effects, sorting (i.e., the relevance of incentive constraints) is entirely driven by the assumption that c_2 is decreasing in θ. Since this unambiguously makes the $\overline{\theta}$-agent more efficient, the relevant incentive problem is keeping *him* from falsely reporting $\underline{\theta}$ by granting him a rent equal to his cost advantage $\Phi(\underline{\mathbf{q}})$ for the corresponding production schedule—and leaving the $\underline{\theta}$-agent a rent of zero. Deducting these rents from joint surplus, the principal is left with *reduced form profits* (i.e., incorporating the optimal choice of transfers $\overline{\mathbf{z}}$ and $\underline{\mathbf{z}}$) of

$$\Pi(\overline{\mathbf{q}}, \underline{\mathbf{q}}) \equiv \nu\big[\overline{W}(\overline{\mathbf{q}}) - \Phi(\underline{\mathbf{q}})\big] + (1-\nu)\underline{W}(\underline{\mathbf{q}}), \tag{5}$$

maximization of which corresponds to condition (4).

The objective function (5) embodies the usual rent-efficiency tradeoff faced by an uninformed principal: His menu of contracts trades off expected joint surplus $\nu\overline{W}(\overline{\mathbf{q}}) + (1-\nu)\underline{W}(\underline{\mathbf{q}})$ against the expected rent payments $\nu\Phi(\underline{\mathbf{q}})$ required to induce truthful reporting by the $\overline{\theta}$-type. This tradeoff leads to inefficiencies whose precise nature we analyze next.

3.2 Partial Distortionary Incentives

A trivial implication of Proposition 3.1 is the usual 'no distortion at the top'-result: In spite of information being private, the efficient $\overline{\theta}$-type still produces first-best quantities in both periods, so $\overline{\mathbf{q}}^{\mathrm{SB}} = \overline{\mathbf{q}}^{*}$. This leaves us with an investigation of the nature of distortions 'at the bottom', that is, of inefficiencies inherent in the contract offered to the inefficient $\underline{\theta}$-type.

The rent-efficiency tradeoff responsible for this distortion involves the simultaneous use of *two* instruments, $\underline{q_1}^{\mathrm{SB}}$ and $\underline{q_2}^{\mathrm{SB}}$. To clarify their individual roles and make the principal's motives more transparent, we first analyze what we shall call 'partial' distortionary incentives. In analogy to the conditional first-best production schedules $\underline{\hat{q}_1}^{*}(q_2)$ and $\underline{\hat{q}_2}^{*}(q_1)$, we let $\underline{\hat{q}_1}^{\mathrm{SB}}(q_2)$ and $\underline{\hat{q}_2}^{\mathrm{SB}}(q_1)$ denote the levels of q_1 and q_2, respectively, which maximize $\underline{W}(\mathbf{q}) - \frac{\nu}{1-\nu}\Phi(\mathbf{q})$ *conditional* on the other period's output level. With these definitions in place, the following partial distortionary motives can be identified:

Proposition 3.2. *Under full commitment, the uninformed principal faces the following (partial) distortionary incentives in designing the $\underline{\theta}$-type's production schedule:*

(a) *Conditional on any first-period output q_1, the uninformed principal will distort second-period output* downward, *so $\underline{\hat{q}_2}^{\mathrm{SB}}(q_1) \leqslant \underline{\hat{q}_2}^{*}(q_1)$ for all $q_1 \in Q_1$.*

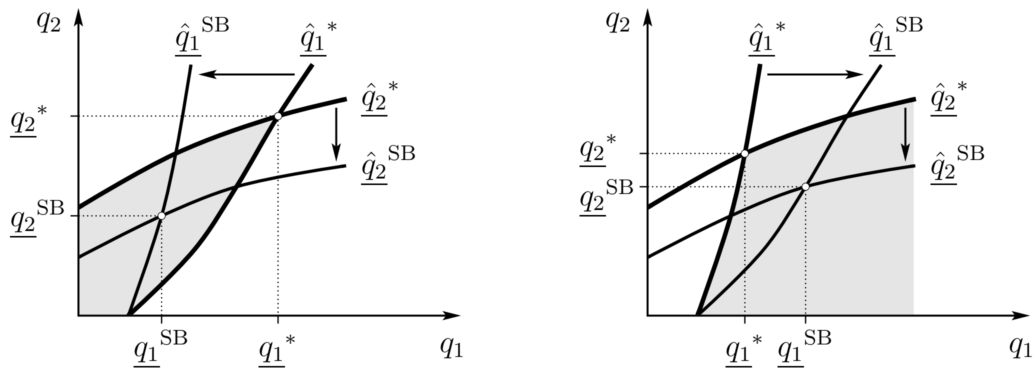
(b) *Conditional on any second-period output q_2, the uninformed principal will distort first-period output*

  (i) downward *if the more efficient agent learns faster, so $\underline{\hat{q}_1}^{\mathrm{SB}}(q_2) \leqslant \underline{\hat{q}_1}^{*}(q_2)$ for all $q_2 \in Q_2$, and*

  (ii) upward *if the less efficient agent learns faster, so $\underline{\hat{q}_1}^{\mathrm{SB}}(q_2) \geqslant \underline{\hat{q}_1}^{*}(q_2)$ for all $q_2 \in Q_2$.*

Part (a) concerning second-period distortionary incentives is not surprising: Given a first-period output level, second-period marginal costs c_2 are a datum, and hence the principal's optimization problem is identical to the standard one-period model of procurement with privately known and constant marginal costs, for which downward distortion (i.e., inefficiently low trade) is a well-known result.

More interestingly, Part (b) identifies the distortionary incentives involved in choosing first-period output—and thereby the extent to which learning effects are under- or overexploited for any given second-period output. That the direction of these distortions crucially depends on which agent learns faster is quickly understood by referring back to our illustration of the two cases in Figure 1: If the more efficient agent learns faster, *decreasing* q_1 leads types' second-period costs c_2 to converge, thereby reducing the $\overline{\theta}$-type's cost advantage (for fixed $\underline{q_2}$) and hence the rent payable to him. Conversely, if the *less* efficient agent learns faster, the same effect is achieved by an *increase* in q_1.



(a) More Efficient Agent Learns Faster.

(b) Less Efficient Agent Learns Faster.

Figure 2: The Direction of Overall Distortions.

3.3 Overall Distortions in Trade

The partial distortions analyzed above provide a direct measure of the over- or underexploitation of learning effects by asking whether the uninformed principal's contracts can be pareto-improved upon by expanding or reducing first-period output. Since the principal ultimately distorts both periods' output schedules simultaneously, however, attaining the full welfare optimum will also require simultaneous adjustments in both quantities.

As the next result shows, if the more efficient type learns faster, attaining the full welfare optimum requires *expanding* both periods' output: Partial distortions in this case are representative of overall distortions, so overall trade is inefficiently low.

Proposition 3.3. *Under full commitment, if the more efficient type learns faster, then private information causes an overall downward distortion in both first- and second-period output for the inefficient type, so* $\underline{\mathbf{q}}^{\mathrm{SB}} \leqslant \underline{\mathbf{q}}^{*}$.

Figure 2(a) illustrates the results under the additional assumption that both $\underline{W}$ and $\underline{W} - \nu\Phi$ are strictly concave. By the implied uniqueness of the maximizers and by Lemma 2.1, the conditional first-best outputs $\hat{\underline{q_1}}^*(q_2)$ and $\hat{\underline{q_2}}^*(q_1)$ are increasing functions. Moreover, concavity of $\underline{W}$ implies that the $\hat{\underline{q_1}}^*$-curve crosses the $\hat{\underline{q_2}}^*$-curve from below at $\mathbf{q}^* = (\underline{q_1}^*, \underline{q_2}^*)$ in $(\underline{q_1}, \underline{q_2})$-space. Now by Proposition 3.2(a), the $\hat{\underline{q_2}}^{\mathrm{SB}}$-curve will lie south of the $\hat{\underline{q_2}}^*$-curve, and by Proposition 3.2(bi), the $\hat{\underline{q_1}}^{\mathrm{SB}}$-curve will lie west of the $\hat{\underline{q_1}}^*$-curve. Hence, the equilibrium under private information—determined by the intersection of the $\hat{\underline{q_2}}^{\mathrm{SB}}$- and the $\hat{\underline{q_1}}^{\mathrm{SB}}$-curve—must lie in the shaded area in Figure 2(a). As illustrated, this area must lie in the southwest quadrant of $\underline{\mathbf{q}}^*$.

Why an analogous argument fails when the *less* efficient agent learns faster is illustrated in Figure 2(b). Again, the $\hat{\underline{q_2}}^{\mathrm{SB}}$-curve must lie south of the $\hat{\underline{q_2}}^*$-curve.

However, Proposition 3.2(bii) in this case tells us that there will be an upward distortion in first-period output given any second-period output, so that the new equilibrium must lie east of the $\underline{\hat{q}_1}^*$-curve. Hence, only equilibria with $\underline{q_1}^{\mathrm{SB}} < \underline{q_1}^*$ and $\underline{q_2}^{\mathrm{SB}} > \underline{q_2}^*$ can be excluded.[10]

Particularly, if the less efficient agent learns faster, it is possible for overall distortions to be upward in both periods, so that $\underline{q_1}^{\mathrm{SB}} > \underline{q_1}^*$ and $\underline{q_2}^{\mathrm{SB}} > \underline{q_2}^*$. This is illustrated by the following extension to Example 1:

Example 2. Assume the value of output to the principal is given by $S(q_t) = 100q_t - 80q_t^2$, cost structures are $c_1 = 75$, $\overline{c_2}(q_1) = 30-60q_1$, and $\underline{c_2}(q_1) = 50-95q_1$, types are equally likely, and the common discount factor is $\delta = 0.7$. First-best production then entails $\overline{\mathbf{q}}^* = (0.30, 0.55)$ and $\underline{\mathbf{q}}^* = (0.38, 0.54)$, whereas contracts under asymmetric information and full commitment will entail $\underline{\mathbf{q}}^{\mathrm{SB}} = (0.49, 0.58)$ for the $\underline{\theta}$-type.

Alternative parameterizations of Example 1 will produce the other two possible directions in overall distortions.

3.4 Are Learning Effects Exploited At All?

Having gauged the outcome under incomplete information against efficient benchmarks, this section investigates whether downward distortionary incentives can be so severe as to eliminate the exploitation of learning effects altogether. For comparison, we consider the outcome which results if either first-period output has no impact on second-period marginal costs, so $\frac{\partial}{\partial q_1} c_2 \equiv 0$, or both principal and agent behave myopically, so $\delta = 0$. The resulting choice of q_1 in either case will maximize first-period surplus $S(q_1) - c_1 q_1$ alone. Motivated by this, we introduce the following terminology:

Definition 3.4. Let $q_1^\circ \equiv \arg\max_{q_1}\{S(q_1) - c_1 q_1\}$. A first-period output level q_1 *exploits* learning effects if $q_1 \geqslant q_1^\circ$; it *neglects* learning effects if $q_1 \leqslant q^\circ$.

Equivalently (recall that $S'' < 0$), learning effects are exploited if the marginal benefit of first-period output S' (weakly) falls short of marginal costs c_1, and neglected if the reverse holds.

Obviously, first-best quantities always exploit learning effects. This need not be true for $\underline{q_1}^{\mathrm{SB}}$, the first-period quantity procured from the $\underline{\theta}$-type. The following result gives sufficient conditions for either case:

[10] A simple generalization of this last argument (generalized beyond the graphical analysis' additional assumptions) runs as follows. Let $\underline{\mathbf{q}}^*$ denote the first-best output schedule (if the first-best output schedule is not unique, let $\underline{\mathbf{q}}^*$ denote any first-best schedule such that there exists no other first-best output schedule involving lower first- and higher second-period output). Then for any $\mathbf{q} = (q_1, q_2)$ with $q_1 < \underline{q_1}^*$ and $q_2 > \underline{q_2}^*$, we have $\underline{W}(\mathbf{q}) < \underline{W}(\mathbf{q}^*)$. Moreover, $\Phi(\mathbf{q}) \leqslant \Phi(\underline{\mathbf{q}}^*)$ for any such $\mathbf{q}$ because Φ is decreasing in q_1 and increasing in q_2. But then $(1-\nu)\underline{W}(\mathbf{q}) - \nu\Phi(\mathbf{q}) < (1-\nu)\underline{W}(\mathbf{q}^*) - \nu\Phi(\mathbf{q}^*)$, so that no such $\mathbf{q}$ can maximize the uninformed principal's objective in (4).

Proposition 3.5. *The contract offered to the $\underline{\theta}$-agent* exploits *learning effects if*

$$\left|\tfrac{\partial}{\partial q_1}\underline{c_2}(q_1)\right| \geqslant \nu \cdot \left|\tfrac{\partial}{\partial q_1}\overline{c_2}(q_1)\right| \tag{6}$$

for all $q_1 \in Q_1$; it neglects *learning effects if (6) is reversed for all $q_1 \in Q_1$.*

Thus, learning effects are exploited if either the efficient agent does not learn too much faster than the inefficient agent, or if efficient types are scarce enough. Intuitively, both ensure that the principal's rent-efficiency tradeoff is sufficiently in favor of efficiency—the former by reducing the efficient type's cost advantage and thereby his rent, the latter by making it less likely that such a rent will have to be paid in the first place.

In relation to our previous results in Section 3.3, Proposition 3.5 shows that even though downward distortions in q_1 may ensue if the *inefficient* agent learns faster, they will never be so strong as to eliminate the exploitation of learning effects altogether. In relation to the previous literature, Proposition 3.5 points out that this may however be the case if the *efficient* agent learns faster, depending on the distribution of types and how strongly learning rates differ. Particularly, the possibility of learning effects being neglected is absent in Lewis and Yildirim's (2002a) model.

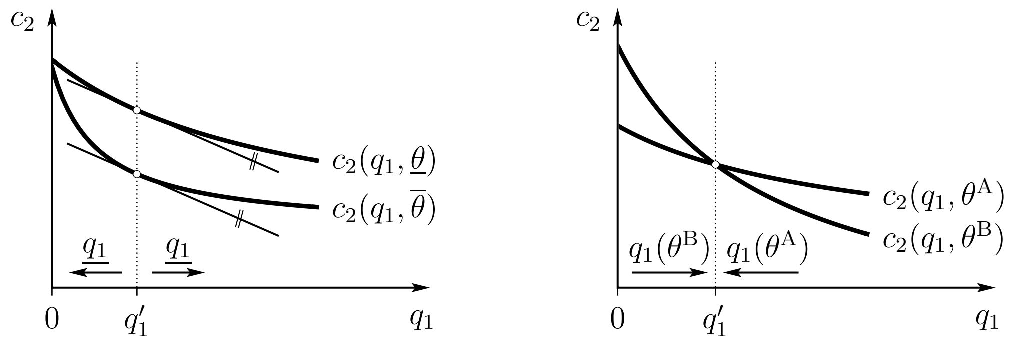
3.5 Extensions and Limitations

Our analysis thus far has relied on several simplifying assumptions. Before the next section proceeds to relax what might seem the most serious and restrictive one—the assumption of full commitment—, we briefly discuss other possible extensions and the challenge they pose to our findings.

'Varying Learning Advantages': Assuming that one of the agents unambiguously learns faster has simplified our identification of first-period distortions, but has been immaterial to the derivation of the optimal contract itself in Proposition 3.1. Without this assumption, it is still true that only the $\underline{\theta}$-type's output is distorted, $\underline{q_2}$ is distorted downward given $\underline{q_1}$, but the direction of the distortion in $\underline{q_1}$ given $\underline{q_2}$ will be ambiguous.

More specifically, consider the cost functions shown in Figure 3(a), where the efficient $\overline{\theta}$-agent learns faster for $q_1 < q_1'$ and slower for $q_1 > q_1'$. Our characterization of partial distortions in $\underline{q_1}$ is then still valid to the extent that they will be downward *if* $\underline{q_1} < q_1'$, and upward *if* $\underline{q_1} > q_1'$. However, which regime is relevant is ambiguous and depends, inter alia, on the value of output relative to its costs.

Intersecting Cost Curves: In contrast, the assumption that the $\underline{\theta}$-type has lower second-period costs c_2 for any q_1 has indeed been vital to our derivation of the optimal contract by determining which of the incentive constraints must bind. This assumption precludes, however, the possibility of one agent 'overtaking' the other due to learning effects.



(a) Agent with Higher Learning Rate Ambiguous.

(b) Agent with Higher Second-Period Costs Ambiguous.

Figure 3: Examples of more General Second-Period Cost Functions.

Such a case is illustrated in Figure 3(b). Here, a straightforward extension to our previous analysis shows that (partial) distortions in $q_1(\theta^{\mathrm{B}})$ can only occur for $q_1(\theta^{\mathrm{B}}) < q_1'$ and will be upward, whereas distortions in $q_1(\theta^{\mathrm{A}})$ can only occur for $q_1(\theta^{\mathrm{A}}) > q_1'$ and will be downward. However, which of these distortions occurs in the optimum is again ambiguous, and both may in fact occur simultaneously.

The previous two examples generalize the main theme of our above analysis in the following straightforward way: Distortions, if they occur, aim to reduce the (locally) less efficient agent's cost-disadvantage. Whether this requires an increase or a decrease in first-period output depends on relative learning rates.

Private Information on First-Period Costs: We have assumed type-independent first-period costs c_1 to focus on the role of asymmetric information on the *returns* to learning by doing. If c_1 also depends on θ, we may generalize the $\overline{\theta}$-type's comparative cost advantage to $\Phi(\underline{q_1}, q_2) \equiv [c_1(\underline{\theta}) - c_1(\overline{\theta})]q_1 + \delta[c_2(q_1, \underline{\theta}) - c_2(q_1, \overline{\theta})]q_2$. Under the assumption that $c_1(\overline{\theta}) \leqslant c_1(\underline{\theta})$, the $\overline{\theta}$-type is still unambiguously more efficient in *both* periods, and the derivation of the optimal contract goes through unchanged.[11] Particularly, the $\overline{\theta}$-type's output schedule remains undistorted. There is now, however, an additional motive to distort $\underline{q_1}$ downward in order to reduce the $\overline{\theta}$-type's rent for first-period cost advantages, which—depending on which agent learns faster—will either reinforce or counteract the distortionary incentives identified above.

Finally, we note that if $c_2(q_1, \overline{\theta}) = c_2(q_1, \underline{\theta})$ for all q_1, in addition, so that private information concerns the cost-side *exclusively*, then we are essentially in the setting considered by Lewis and Yildirim (2002a,b), and learning-effects will always be *under*exploited for the $\underline{\theta}$-type.

[11] If $c_1(\overline{\theta}) > c_1(\underline{\theta})$, we again face a situation in which it is unclear which incentive constraint binds at the optimum.

Alternative Valuations of Output: We have assumed that the principal values output at $S(q_1) + \delta S(q_2)$. However, our key results concerning (partial) distortionary incentives in Proposition 3.2 are robust to more general valuation functions $\tilde{S}(\mathbf{q})$.[12] A particularly interesting extension involves letting $\tilde{S}$ depend on q_2 alone: In this case, our model represents a pure investment problem with privately known returns, but where the *level* of investment is contractible. Reinterpreting our above results, underinvestment then ensues (in both a partial and an overall sense) if investment returns (in terms of cost savings) become *more* sensitive to private information with higher investment levels, whereas overinvestment (at least in a partial sense) occurs in the reverse case.

4 Contracts under Spot Commitment

The above analysis has been based on the premise that in the first period, the principal can fully commit to a contract spanning both procurement periods. Particularly, this requires the principal being able to preclude any interim renegotiations on the original contract, even if they are mutually advantageous. This may be somewhat troubling, as advantageous renegotiations indeed exist at the beginning of the second period whenever the agent turns out to be bad: By the nature of the truth-revealing contract, the principal learns this in the first period already. But then there is no longer a point in having the agent produce an inefficiently low quantity in the second period. Indeed, the principal can always offer a new contract which induces *efficient* production by the bad agent and constitutes a strict Pareto-improvement over the old contract. To complete this well-known argument, agents' anticipation of this incentive to renege will in turn change their own first-period incentives, making it impossible to implement the above contract under limited commitment.

Limited commitment is generally understood to be a deterrent to long-term investments (see for instance Fudenberg et al., 1990). As the exploitation of learning effects over several periods is inherently long termed (the costs of higher output today leads to lower production costs only tomorrow), we would expect investment into learning to be affected by the inability to commit to long-term contracts. Specifically, we have seen above that under full commitment, learning effects can be exploited even when agents are privately informed about their learning abilities. We

[12]Except for our analysis in Section 3.4, all above results in fact easily generalize to cases in which the valuation function $\tilde{S}(\mathbf{q})$ displays complementarities in q_1 and q_2 in the sense that $\partial^2\tilde{S}/\partial q_1 \partial q_2 \geqslant 0$. To understand this, note that the analysis has employed the particular form of valuation function only to the extent that it implies additive separability of valuations in q_1 and q_2, which in turn implies that, due to cost-side effects alone, the surplus function W is complementary in q_1 and q_2. Comparative static results reliant on this complementarity (i.e., Lemma 2.1 and Proposition 3.3) are thus robust to valuation functions which preserve this complementarity.

have further seen that in some circumstances, private information may even lead to learning effects being over-exploited relative to the social optimum. It should thus be interesting to see how these results are altered by limited commitment.

To analyze this, we assume in this section that rather than being able to offer contingent contracts describing all quantities and transfers q_1, q_2, z_1, and z_2 at the beginning of period one, the principal is limited to offering *spot contracts* in each period: At the beginning of period one, he can offer a menu which settles first-period exchanges q_1 and z_1, whereas he can offer a menu which settles second-period exchanges q_2 and z_2 only at the beginning of period two. Furthermore, in each period, the agent can decline, which yields him a reservation utility of 0 and terminates the game.[13]

It is important to stress that this section assumes that the principal is not *able* to commit to a second-period contract in period one. Indeed, by a standard argument (see for instance Salanié, 1997, p. 146), the principal can reproduce any sequence of spot contracts (and, more generally, any sequence of complete contracts under limited commitment) by fully committing in period one. Thus, if the principal *can* fully commit, it will be (weakly) optimal for him to do so, and the analysis of Section 3 applies.

However, there are cases where it is indeed realistic to assume that the principal cannot commit. First, under virtually all legal systems, courts of law cannot prevent parties from renegotiating on a contract if all parties agree to do so. Hence, pure contractual long-term commitments with scope for pareto-improving renegotiations are essentially meaningless.[14] Second, regulatory authorities often cannot commit to decisions beyond the life-span of the current administration. Finally, it appears to be a simple matter of fact that real-world procurement and regulatory relations are often governed by short-term contracts, even if the relationship itself is long termed (see Lewis and Yildirim, 2002a, on the latter).

We investigate spot contracts in five steps. First, Section 4.1 analyzes the extent to which spot commitment is compatible with full revelation of information in the first period. We show that full first-period revelation—and thereby the full-commitment allocation—is typically not achievable due to the well-known ratchet effect. Next, Section 4.2 sets the stage for more general mechanisms with only *partial* first-period revelation of information by appealing to an extended revelation principle due to Bester and Strausz (2001). Based thereon, Sections 4.3 and 4.4 character-

[13]An intermediate level of commitment which we do not analyze here is that of *long-term commitment* (see for instance Laffont and Tirole, 1990), which grants the principal the opportunity to offer contracts spanning all periods of gameplay, but does not let him commit not to renege, i.e. to offer a new contract which supersedes the original one, in later periods.

[14]Parties may however find a *non*-contractual way to commit. For instance, if the principal procures some input into his own production from the agent, he may be able to commit to a certain second-period quantity by means of irreversible first-period investments that are specific to the level of second-period quantity, such as by choosing the size of his assembly plant.

ize the sequence of equilibrium spot contracts. Finally, using this characterization, Section 4.5 presents numeric examples to show that even under spot-commitment, it is possible for learning effects to be over-exploited.

4.1 First-Period Information Revelation and the Ratchet Effect

In the full-commitment setting of Section 3, the principal offers a contract such that the agent fully reveals himself at the start of the first period and thereby determines the terms of trade for the entire relationship. Under spot commitment, inducing such reporting behavior is much more difficult and in most cases even impossible. The next result, which is essentially an application of the well-known 'ratchet' result (cf. Freixas et al., 1985; Laffont and Tirole, 1987, 1988) to our model, shows that full first-period revelation can be an equilibrium phenomenon only under heavy restrictions on quantities implemented thereby:

Lemma 4.1. *Under spot contracting, both agents truthfully revealing their type in the first period can only be an equilibrium if the quantities procured from the inefficient type,* $(\underline{q_1}, \underline{q_2})$*, are such that the efficient agent has no cost advantage, that is, if* $\Phi(\underline{q_1}, \underline{q_2}) = 0$.

See the Appendix for the proof.

The intuition for this observation is straightforward: First, full information revelation in period one implies that the principal is fully informed about the agent's type when designing the second-period contract, which lets him extract the full surplus in the second period and leaves each type a second-period rent of zero. Now suppose the principal offers a contract to the $\underline{\theta}$-agent for which the $\overline{\theta}$-agent has a *strict* second-period cost-advantage, that is, for which $\Phi(\underline{q_1}, \underline{q_2}) = \delta \underline{q_2}[c_2(\underline{q_1}, \underline{\theta}) - c_2(\underline{q_1}, \underline{\theta})] > 0$. Since limited commitment prevents the principal from paying the $\overline{\theta}$-agent a positive second-period rent, the principal can only keep the $\overline{\theta}$-agent from misreporting and grabbing the resulting positive second-period rent if the false report leads to a strict *decrease* in his first-period rent. However, since first-period rents are independent of θ due to identical first-period production costs c_1, a $\underline{\theta}$-agent could then strictly *increase* his overall payoff by a "take-the-money-and-run"-strategy, that is, by misreporting in period one (yielding a strictly positive first-period rent), and by subsequently rejecting the second-period contract to obtain his reservation payoff of zero.

We have thus shown that truthful revelation in the first-period can be an equilibrium only if the quantities procured from the $\underline{\theta}$-agent are such that the $\overline{\theta}$-agent has no cost-advantage. Additionally, however, due to the spot nature of commitment, the principal cannot simply commit to offering a contract with this property in the first period: such a contract must be *sequentially optimal*. This poses no problem if there exists a first-period q_1 which equates agents' second-period costs. Whenever this is not the case, however, then the only way to keep the $\overline{\theta}$-agent from having a cost advantage over the $\underline{\theta}$-agent under the latter's contract is by shutting down the $\underline{\theta}$-agent in the second period, so $\underline{q_2} = 0$—and this shutdown must be sequentially optimal, that is, it must constitute a "credible threat":

Corollary 4.2. *Truthful revelation in the first period can only be an equilibrium if* $c_2(q_1, \overline{\theta}) = c_2(q_1, \underline{\theta})$ *for some* $q_1 \in Q_1$*, or if* $\arg\max_{q_2}\{S(q_2) - c_2(q_1, \underline{\theta})q_2\} = 0$ *for some* $q_1 \in Q_1$.

Corollary 4.2 implies in particular that truth-telling cannot be an equilibrium in environments for which $c_2(q_1, \overline{\theta}) < c_2(q_1, \underline{\theta})$ for all q_1 (so that the $\underline{\theta}$-type never fully catches up) and $S'(0) > c_2(0, \underline{\theta})$ (so that positive second-period production by the $\underline{\theta}$-type is efficient even if he produced nothing in the first period, i.e., if he enjoyed no learning). As an immediate consequence, full revelation of information cannot be an equilibrium in a large class of environments.

As an alternative to *full* first-period revelation, the principal may of course make the first-period allocation altogether independent of any first-period message. This will trivially circumvent any first-period incentive problems—but will typically entail high efficiency losses. As we will see, the optimal sequence of spot contracts can—and typically will—involve revelation behavior *in between* these two polar cases: Agents will reveal their type only *partially* in period one, in a way to be made precise in the next section.

We will also see that the described constraints to *full* first-period revelation are paralleled by similar constraints to *partial* revelation: First, more information revelation by the efficient type in period one reduces his rent in period two, which represents an instance of the ratchet effect. Therefore, the efficient type can only be induced to reveal more information if he is compensated with a higher first-period rent. Second, however, the higher first-period rent to the efficient agent will make it more profitable for the inefficient agent to report falsely in period one, since he can always 'take the money and run'. Consequently, as under full first-period revelation, *both* types' first-period incentive constraints will bind in a partially revealing equilibrium.[15] On the other hand, we will see that the efficient agent will generally obtain a strictly positive second-period for revealing his residual information in period two.

[15] This clear result concerning the relevance of both incentive constraint stands in stark contrast to the usual ambiguity under spot contracts. Indeed, in their two-type model of spot-contracting with *constant* (but privately) known costs, Laffont and Tirole (1987) emphasize the point that, depending on the specific parameterization, four distinct types of equilibria can result depending on which incentive constraint binds. As we will see, the problem of ambiguity is avoided in our model by the simplifying assumption of identical *first*-period costs.

4.2 Spot Contracts and the Revelation Principle

In full generality, interaction in our two-period learning-by-doing model with spot commitment proceeds as follows: First, in each period $t = 1, 2$, the principal announces a message space M_t and a corresponding menu of outputs and transfers, $q_t : M_t \to Q_t$ and $z_t : M_t \to \mathbb{R}$, specifying the quantity purchased and the transfer made in period t for any message $m_t \in M_t$. Second, in each period, the agent decides whether to accept the contract or not, and in the former case announces a message $m_t \in M_t$ according to a reporting strategy which is given by a probability distribution over messages M_t for each possible type $\theta \in \{\overline{\theta}, \underline{\theta}\}$. Moreover, players' strategies in each period of course depend on the history of the game to date. As usual, we require equilibrium strategies to be optimal in the sense of a *perfect Bayesian equilibrium*.

As in any mechanism design problem, a main obstacle in the search for equilibria of this game comes from the principal's infinite degrees of freedom in designing the message space M_t. In settings in which the principal can fully commit to all future actions (such as in Section 3), this problem is easily resolved through the revelation principle, according to which there is no loss in generality in (i) restricting the message space to type space and (ii) restricting agents' reporting strategies to strategies in which agents report their true type in equilibrium. With spot commitment, however, the classical revelation principle is only applicable to the last (i.e., second) period, whereas reducing the dimensionality of the set of *first*-period strategies requires a recent extension thereof due to Bester and Strausz (2001).

More specifically, consider first the second period: For any conceivable history of the game (and correspondingly updated beliefs of the principal), the problem of designing the optimal second-period contract is identical to the standard adverse selection model. Thus, we may restrict the set of second-period mechanisms $\{M_2, q_2(\cdot), z_2(\cdot)\}$ to ones in which $M_2 = \{\overline{\theta}, \underline{\theta}\}$ and $q_2(\cdot), z_2(\cdot)$ is such that the agent finds it optimal to report his true type with probability one.

This is not so for the first period, since by assumption, the principal cannot commit himself to contracts offered in the second period. As illustrated in Section 4.1, the agent's first-period message reveals information to the principal which the principal may use in designing the second-period contract to the agent's disadvantage. However, due to an extension of the revelation principle in Bester and Strausz (2001), in solving for the optimal contract, we may restrict our attention to contracts where (i) the agent's message space corresponds to the type space (so $M_1 = \{\overline{\theta}, \underline{\theta}\}$), and (ii) truth-telling is an optimal strategy for each type of agent in the sense that he reveals his true type with some strictly positive probability (but not necessarily equal to 1).[16]

[16] There are three noteworthy differences between the classical revelation principle and that presented in Bester and Strausz (2001): *First*, as noted, the extended revelation principle requires including also 'weakly truth-revealing' mechanisms in which agents reveal their true type only with

Hence, without further loss of generality, we may reduce our mechanism-design problem to one in which (i) $M_t = \{\overline{\theta}, \underline{\theta}\}$ in each period $t = 1, 2$, (ii) agents find it optimal to always report their true type in the second period, and (iii) agents of any type $\theta \in \{\overline{\theta}, \underline{\theta}\}$ find it optimal to report their true type with some probability $p_1(\theta) > 0$ in the first period (and report the false type with probability $1 - p_1(\theta)$).[17]

Equipped with this major simplification, we now proceed to characterize the optimal series of spot contracts. In the spirit of backward induction, we first characterize the set of sequentially optimal second-period mechanisms for any history of the game in Section 4.3, and then proceed to characterize the optimal first-period spot contract in Section 4.4.

4.3 The Second-Period Spot Contract

4.3.1 The Formation of Second-Period Beliefs

At the beginning of the second period, the principal updates his beliefs on the agent's type based on the history of the game and agents' reporting strategies $p_1(\theta)$. Letting $\nu_2(\tilde{\theta}_1)$ denote the assessed probability that the agent is of type $\overline{\theta}$ if he sent signal $\tilde{\theta}_1$ in the first period, Bayes' Law delivers

$$\nu_2(\tilde{\theta}_1) = \begin{cases} \dfrac{p_1(\overline{\theta}) \cdot \nu}{p_1(\overline{\theta}) \cdot \nu + [1 - p_1(\underline{\theta})] \cdot (1 - \nu)}, & \tilde{\theta}_1 = \overline{\theta}, \\ \dfrac{[1 - p_1(\overline{\theta})] \cdot \nu}{[1 - p_1(\overline{\theta})] \cdot \nu + p_1(\underline{\theta}) \cdot (1 - \nu)}, & \tilde{\theta}_1 = \underline{\theta}. \end{cases} \tag{7}$$

Note that this belief is always well-defined since each type reports truthfully with some strictly positive probability, so that each signal is observed with nonzero probability in equilibrium.

Condition (7) formalizes the way in which agents' first-period reporting strategies $p_1(\cdot)$ affect the informativeness of the first-period report $\tilde{\theta}_1$. Instructive polar cases arise for $p_1(\overline{\theta}) = 1 - p_1(\underline{\theta})$, in which case $\nu_2(\overline{\theta}) = \nu_2(\underline{\theta}) = \nu$, so that the first-period report $\tilde{\theta}_1$ is entirely uninformative, and for $p_1(\overline{\theta}) = p_1(\underline{\theta}) = 1$, in which case $\nu_2(\overline{\theta}) = 1$

---

a certain nonzero probability. *Second*, while the classical revelation principle contends that *any* incentive feasible mechanism can be replicated by a truth-revealing one, the extended revelation principle covers only *incentive efficient* ones (i.e., mechanisms such there exists no other incentive feasible mechanism which yields a strictly higher payoff to the principal and the same payoff to the agent). The latter presents no restriction, however, if we are only interested in finding contracts which maximize the principal's profits. *Finally*, in terms of allocational outcomes, the procedure delivers *a*, but not necessarily *all*, mechanisms which maximize the principal's payoff. Particularly, while the principal cannot *strictly* increase his payoff with a richer message space, it does not preclude the existence of mechanisms with richer message spaces which permit the principal to achieve the *same* payoff.

[17] We note that a similar simplifying assumption was already made in Laffont and Tirole (1987, 1993) in the context of dynamic regulation with spot commitment, but without the theoretical underpinning provided by Bester and Strausz (2001).

and $\nu_2(\underline{\theta}) = 0$, so that first-period reports fully reveal the agents' type (as analyzed in Section 4.1 above).

4.3.2 The Optimal Second-Period Continuation Contract

Next, we consider the principal's optimal design of the second-period contract given his belief $\nu_2(\cdot)$ and given any history of the game. Observe first that at the start of the second period, for any given first-period contract $(q_1(\tilde{\theta}_1), z_1(\tilde{\theta}_1))$ and any reporting strategy $p_1(\cdot)$, both the history of the game (i.e., first-period outputs and transfers) and the principal's beliefs are completely summarized by the first-period report $\tilde{\theta}_1$. Hence, for any $\{p_1(\cdot), q_1(\cdot), z_1(\cdot)\}$, we may let $(q_2(\tilde{\theta}_1, \tilde{\theta}_2), z_2(\tilde{\theta}_1, \tilde{\theta}_2))$ denote the second-period contract offered to an agent who reported $\tilde{\theta}_1$ in the first period and who reports $\tilde{\theta}_2$ in the second period.

As noted above, the revelation principle lets us restrict attention to second-period contracts which induce truthful revelation, so $\tilde{\theta}_2 = \theta$ for an agent of type θ and any first-period report $\tilde{\theta}_1$. Letting $u_2(\tilde{\theta}_1, \theta) \equiv z_2(\theta) - c[q_1(\tilde{\theta}_1), \theta]q_2(\theta)$ denote the second-period rent obtained by an agent of type θ who reported $\tilde{\theta}_1$ in period one, we obtain the following characterization of optimal second-period contracts:

Proposition 4.3. *For any first-period strategies $q_1(\cdot), z_1(\cdot), p_1(\cdot)$ and any first-period report θ_1, the second-period continuation contract $q_2(\tilde{\theta}_1, \theta), z_2(\tilde{\theta}_1, \theta)$ must be such that*

(a) outputs satisfy

$$q_2(\tilde{\theta}_1, \overline{\theta}) \in \arg\max_{q_2} \Big\{ \nu_2(\tilde{\theta}_1) \cdot \big[S(q_2) - c_2[q_1(\tilde{\theta}_1), \overline{\theta}] \cdot q_2 \big] \Big\} \tag{8}$$

$$q_2(\tilde{\theta}_1, \underline{\theta}) \in \arg\max_{q_2} \Big\{ [1 - \nu_2(\tilde{\theta}_1)] \cdot \big[S(q_2) - c_2[q_1(\tilde{\theta}_1), \underline{\theta}] \cdot q_2 \big] - \nu_2(\tilde{\theta}_1) \cdot \Phi[q_1(\tilde{\theta}_1), q_2]/\delta \Big\}, \tag{9}$$

$$\Phi[q_1(\tilde{\theta}_1), q_2(\tilde{\theta}_1, \overline{\theta})] \geqslant \Phi[q_1(\tilde{\theta}_1), q_2(\tilde{\theta}_1, \underline{\theta})], \tag{10}$$

where $\nu_2(\tilde{\theta}_1)$ is determined by (7),

(b) transfers $z_2(\tilde{\theta}_1, \theta)$ are such that $u_2(\tilde{\theta}_1, \underline{\theta}) = 0$ and $u_2(\tilde{\theta}_1, \overline{\theta}) = \Phi[q_1(\tilde{\theta}_1), q_2(\underline{\theta})]/\delta$. Moreover, for $\nu_2(\tilde{\theta}_1) > 0$, any $q_2(\tilde{\theta}_1, \theta)$ which satisfies (8) and (9) will satisfy (10).

See the Appendix for the proof.

Although somewhat obscured by notation including all possible histories of the game, Proposition 4.3 represents standard results: Conditional on first-period output, for $\nu_2(\tilde{\theta}_1) \in (0, 1)$, the $\overline{\theta}$-type's second-period output is efficient, and the $\underline{\theta}$-type's output is distorted downward from its conditionally efficient level in order to reduce the rent payable to the $\overline{\theta}$-type.

However, since the principal's belief $\nu_2(\tilde{\theta}_1)$ is endogenous, we must also anticipate cases in which $\nu_2(\tilde{\theta}_1) = 1$ and $\nu_2(\tilde{\theta}_1) = 0$: In contrast to Proposition 3.1, these cases are not precluded by the assumption that $\nu \in (0,1)$. For $\nu_2(\tilde{\theta}_1) = 1$, the principal attaches no probability to the agent being of type $\underline{\theta}$ and will set $q_2(\tilde{\theta}_1, \underline{\theta})$ with the sole objective of eliminating the $\overline{\theta}$-agent's rent, which he can always achieve by setting $q_2(\tilde{\theta}_1, \underline{\theta}) = 0$, so that $\Phi[q_1(\tilde{\theta}_1), q_2(\tilde{\theta}_1, \underline{\theta})] = 0$. If $\nu_2(\tilde{\theta}_1) = 0$, the $\overline{\theta}$-type's output will be efficient, and *any* output of the $\overline{\theta}$-type will satisfy (8). Only in this latter case will the implementability condition (10) be relevant, in the form of additional restrictions on $q_2(\tilde{\theta}_1, \overline{\theta})$ to ensure that incentive compatibility conditions can be met. Without further loss of generality, however, we may simply replace condition (8) by

$$q_2(\tilde{\theta}_1, \overline{\theta}) \in \arg\max_{q_2} \big[S(q_2) - c_2[q_1(\tilde{\theta}_1), \overline{\theta}] \cdot q_2\big], \tag{8'}$$

which allows us to neglect (10) altogether.[18]

4.4 The First-Period Spot Contract

In Proposition 4.3, we have fully characterized feasible outcomes of the second period for any conceivable first-period outcomes and strategies. Using this characterization, we now proceed to characterize the optimal first-period contract $q_1(\cdot), z_1(\cdot)$.

We have already made use of the fact that, by Bester and Strausz (2001), there is no loss of generality in restricting first-period mechanisms to *direct* mechanisms (i.e., with message space $\{\overline{\theta}, \underline{\theta}\}$). Moreover, by Bester and Strausz's result, the principal's optimization problem may be formulated as choosing (i) the first-period contract $(q_1(\cdot), z_1(\cdot))$, (ii) the agent's reporting strategy $p_1(\cdot)$, and (iii) the second-period continuation contract $(q_2(\cdot), z_2(\cdot))$ so as to maximize his expected payoff,[19] subject to

(i) $p_1(\cdot) > 0$;

(ii) the continuation contract $(q_2(\tilde{\theta}_1, \theta), z_2(\tilde{\theta}_1, \theta))$ satisfying Proposition 4.3;

(iii) the first-period reporting strategy p_1 being optimal for the agent *given* the impact this has on his continuation payoff;

(iv) first-period participation constraints.

To operationalize the latter two requirements, in the by now familiar fashion, we let $u_1(\tilde{\theta}_1) \equiv z_1(\tilde{\theta}_1) - c_1 \cdot q_1(\tilde{\theta}_1)$ denote the first-period rent of an agent who reports $\tilde{\theta}_1$ in the first-period. Note that this first-period rent is independent of the agent's true type θ due to the fact that we have assumed away first-period cost asymmetries. Moreover, we let $U(\theta) \equiv u_1(\theta) + \delta u_2(\theta, \theta)$ denote the overall discounted rent of an

[18]That any $q_1(\tilde{\theta}_1, \overline{\theta})$ and $q_1(\tilde{\theta}_1, \underline{\theta})$ satisfying (8') and (9) satisfy (10) also for $\nu_2(\tilde{\theta}_1) = 0$ is seen by trivial extension of the proof of Proposition 4.3 in the Appendix. To understand why this is without loss of generality, note that if $\nu_2(\tilde{\theta}_1) = 0$, the principal's expected profits are independent of $q_2(\tilde{\theta}_1, \overline{\theta})$ since he attaches zero probability to the corresponding sequence of reports.

[19]To understand the fact that the *principal* maximizes over the *agent's* reporting strategy $p_1(\cdot)$, note that this maximization is of course subject to the strategy being optimal for the agent. A similar comment applies to the principal maximizing ex-ante over $q_2(\cdot), z_2(\cdot)$.

agent of type θ who reports truthfully in period one (it will become clear immediately that $U(\theta)$ is in fact the utility of an agent of type θ for *any* first-period report made with strictly positive probability in equilibrium).

Since $p_1(\theta) > 0$ for all θ, the reporting strategy being optimal requires that the overall discounted rent obtained from truthful reporting must be at least as large as that obtained from any other report, that is, $U(\theta) = \max_{\tilde{\theta}_1}\{u_1(\tilde{\theta}_1) + \delta u_2(\tilde{\theta}_1, \theta)\}$ for all θ. Using the second-period rents $u_2(\tilde{\theta}_1, \theta)$ derived in Proposition 4.3, these conditions can be expressed as

$$U(\overline{\theta}) \geqslant U(\underline{\theta}) + \Phi[q_1(\underline{\theta}), q_2(\underline{\theta}, \underline{\theta})] \tag{11}$$

$$U(\underline{\theta}) \geqslant U(\overline{\theta}) - \Phi[q_1(\overline{\theta}), q_2(\overline{\theta}, \underline{\theta})]. \tag{12}$$

Moreover, whenever $p_1(\theta) < 1$ for some type θ, truth-telling must be a *weakly* optimal strategy, in which case the corresponding incentive constraint must bind:

$$[1 - p_1(\overline{\theta})]\big\{U(\overline{\theta}) - U(\underline{\theta}) - \Phi[q_1(\underline{\theta}), q_2(\underline{\theta}, \underline{\theta})]\big\} = 0 \tag{13}$$

$$[1 - p_1(\underline{\theta})]\big\{U(\underline{\theta}) - U(\overline{\theta}) + \Phi[q_1(\overline{\theta}), q_2(\overline{\theta}, \underline{\theta})]\big\} = 0. \tag{14}$$

Finally, the discounted rent which accrues to the agent over the two periods must at least match his outside opportunity of zero in order for the agent to participate:

$$U(\overline{\theta}) \geqslant 0 \tag{15}$$

$$U(\underline{\theta}) \geqslant 0. \tag{16}$$

Note that the agent will be willing to accept a first-period contract yielding a negative rent u_1, provided that he is compensated by a sufficiently large rent in the second-period contract.[20]

The incentive constraints (11) and (12) immediately deliver the implementability condition

$$\Phi[q_1(\overline{\theta}), q_2(\overline{\theta}, \underline{\theta})] \geqslant \Phi[q_1(\underline{\theta}), q_2(\underline{\theta}, \underline{\theta})], \tag{17}$$

which says that the second-period rent obtained by an efficient type, $u_2(\tilde{\theta}_1, \overline{\theta})$, must be maximal under truth-telling in period one ($\tilde{\theta}_1 = \overline{\theta}$).

Following the usual argument, (11) and the fact that $\Phi \geqslant 0$ implies that $U(\overline{\theta}) \geqslant U(\underline{\theta})$, so that (11) and (16) imply (15). Thus, we may neglect constraint (15). Moreover, (16) must bind (i.e., $U(\underline{\theta}) = 0$) in equilibrium: Otherwise, the principal

[20] Moreover, we are implicitly assuming that it is optimal for the principal to induce participation by either type of agent in every period. This is without loss of generality, however, because non-participation by any type θ in any period $t = 1, 2$ is equivalent in terms of payoffs and incentive feasibility constraints to a mechanism which leads to $q_t = 0$ and $z_t = 0$ for that type.

It is important to note that this does *not*, however, imply that the 'take-the-money-and-run'-actions described in Section 4.1 pose no restriction on incentive design. They are simply not an equilibrium phenomenon (although $q_2 = 0$ and $z_2 = 0$ in the second period is essentially equivalent to 'running away').

could reduce $U(\overline{\theta})$ and $U(\underline{\theta})$ by the same small amount without violating any of the remaining constraints, and thereby strictly raise his expected payoff.[21] Given this, constraints (11) and (12) simplify to

$$\Phi[q_1(\overline{\theta}), q_2(\overline{\theta}, \underline{\theta})] \geqslant U(\overline{\theta}) \geqslant \Phi[q_1(\underline{\theta}), q_2(\underline{\theta}, \underline{\theta})]. \tag{18}$$

The line of arguments up to this point should appear familiar from the full-commitment case (see the proof of Proposition 3.1). The difference here comes from the additional requirements of sequential rationality (see Proposition 4.3) and constraints (13) and (14), which—in the next step of the argument—cause the range of permissible $U(\overline{\theta})$ implied by (18) to collapse to a single point:

Lemma 4.4. *In equilibrium, the implementability condition (17) must bind. Equivalently, both types' first-period incentive constraints (11) and (12) must bind in equilibrium.*

See the Appendix for the proof.[22]

The important implication of Lemma 4.4 is that in any spot equilibrium, the first-period report $\tilde{\theta}_1$ cannot be used to *separate* agents in that agents' first- and second-period rents (and thereby also the cost advantage enjoyed by the $\overline{\theta}$-agent) must be *independent* of first-period reports. Observe that this parallels the result derived in Lemma 4.1 for *full* first-period revelation—the difference being that under *partial* first-period revelation, the $\overline{\theta}$-agent may obtain a strictly positive second-period rent on residual information not revealed in the first period.

To appreciate and understand Lemma 4.4, note that its clear-cut conclusion concerning the relevance of both incentive constraint stands in stark contrast to the usual ambiguity under spot contracts. Indeed, a key result in Laffont and Tirole's (1987) two-type model of spot-contracting with *constant* (but privately known) marginal costs is that, depending on the specific parameterization, four distinct types of equilibria will result depending on which incentive constraint binds—an ambiguity which is reminiscent of the standard one-period analysis when the Spence-Mirrlees

[21] Recall that the principal's payoff corresponds to expected surplus minus expected rents. By (13) and (14), the rent payable to any agent θ must be $U(\theta)$ irrespectively of his first-period report whenever false reporting is part of the equilibrium (i.e., if $p_1(\theta) < 1$), so that expected rents may simply be written as $\nu U(\overline{\theta}) + (1-\nu)U(\underline{\theta})$. Hence, the principal's expected payoff is strictly decreasing in both $U(\overline{\theta})$ and $U(\underline{\theta})$.

[22] We note that the proof also shows that any equilibrium with $p_1(\underline{\theta}) = 1$ must entail $\Phi[q_1(\overline{\theta}), q_2(\overline{\theta}, \underline{\theta})] = \Phi[q_1(\underline{\theta}), q_2(\underline{\theta}, \underline{\theta})] = 0$, which generalizes Lemma 4.1 by dropping the requirement that $p_1(\overline{\theta}) = 1$. To understand this, recall that the crucial point in the argument leading to Lemma 4.1 was simply that a $\overline{\theta}$-type gets no rent in the second-period. But even if the $\overline{\theta}$-type does not reveal himself with *certainty* in period one, so long as the $\underline{\theta}$-type does, the $\overline{\theta}$-type announcing his true type (which must be weakly optimal) will fully reveal his type and therefore leave him a second-period rent of zero. Hence, the $\overline{\theta}$-type receives no second-period rent, and invoking the "take-the-money-and-run"-argument on behalf of the $\underline{\theta}$-type leads to the conclusion that neither type can receive a first-period rent.

sorting condition is violated (cf. Guesnerie and Laffont, 1984). This ambiguity is circumvented in our setting by focussing on the *returns* to learning by doing: Assuming that first-period costs c_1 are type-independent implies that first-period rents u_1 are type-independent as well, so that first-period sorting (or 'separation') can be based only on *second* period rents u_2. Thus, in contrast to Laffont and Tirole (1987), there is no conflict between first- and second-period sorting conditions. Moreover, since (i) first-period separation can only be based on second-period rents, but (ii) the ratchet effect keeps the principal from being able to commit to raising second-period rents due to first-period information revelation, first-period separation is altogether infeasible in the sense stated in Lemma 4.4.

Thanks to Lemma 4.4, we may now describe the equilibrium sequence of spot contracts and the accompanying reporting strategies by means of a straightforward optimization problem:

Proposition 4.5. *(a) In the spot equilibrium, $p_1(\cdot)$, $q_1(\cdot)$ and $q_2(\cdot,\cdot)$ must maximize the principal's expected profits, given by*

$$\begin{aligned}
&\nu p_1(\overline{\theta})\Big\{S[q_1(\overline{\theta})] - c_1 q_1(\overline{\theta}) + \delta\big[S[q_2(\overline{\theta},\overline{\theta})] - c_2[q_1(\overline{\theta}),\overline{\theta}]q_2(\overline{\theta},\overline{\theta})\big]\Big\}\\
&+\nu(1-p_1(\overline{\theta}))\Big\{S[q_1(\underline{\theta})] - c_1 q_1(\underline{\theta}) + \delta\big[S[q_2(\underline{\theta},\overline{\theta})] - c_2[q_1(\underline{\theta}),\overline{\theta}]q_2(\underline{\theta},\overline{\theta})\big]\Big\}\\
&+(1-\nu)p_1(\underline{\theta})\Big\{S[q_1(\underline{\theta})] - c_1 q_1(\underline{\theta}) + \delta\big[S[q_2(\underline{\theta},\underline{\theta})] - c_2[q_1(\underline{\theta}),\underline{\theta}]q_2(\underline{\theta},\underline{\theta})\big]\Big\}\\
&+(1-\nu)(1-p_1(\underline{\theta}))\Big\{S[q_1(\overline{\theta})] - c_1 q_1(\overline{\theta}) + \delta\big[S[q_2(\overline{\theta},\underline{\theta})] - c_2[q_1(\overline{\theta}),\underline{\theta}]q_2(\overline{\theta},\underline{\theta})\big]\Big\}\\
&-\nu\Phi[q_1(\underline{\theta}), q_2(\underline{\theta},\underline{\theta})],
\end{aligned} \tag{19}$$

subject to
(a1) $p_1(\overline{\theta}), p_1(\underline{\theta}) > 0$*;*
(a2) $p_1(\cdot)$, $q_1(\cdot)$, *and* $q_2(\cdot,\cdot)$ *satisfying Proposition 4.3, and*
(a3) $q_1(\cdot)$ *and* $q_2(\cdot,\underline{\theta})$ *satisfying*

$$\Phi[q_1(\overline{\theta}), q_2(\overline{\theta},\underline{\theta})] = \Phi[q_1(\underline{\theta}), q_2(\underline{\theta},\underline{\theta})]. \tag{20}$$

(b) Moreover, the accompanying transfers, $z_1(\cdot)$ and $z_2(\cdot)$ are chosen such that

(b1) $u_2(\overline{\theta},\overline{\theta}) = u_2(\underline{\theta},\overline{\theta}) = \Phi[q_1(\overline{\theta}), q_2(\overline{\theta},\underline{\theta})]$ *and* $u_2(\underline{\theta},\underline{\theta}) = u_2(\overline{\theta},\underline{\theta}) = 0$*, and*

(b2) $u_1(\overline{\theta}) = u_1(\underline{\theta}) = 0$.

Proposition 4.5 reformulates the problem of finding the equilibrium under spot commitment as an optimization problem which is considerably more straightforward than its first appearance might suggest. First, note that in the usual fashion, part (b) of Proposition 4.5 concerning equilibrium transfers is irrelevant to finding the equilibrium values of the allocationally relevant strategies $q_1(\cdot)$, $q_2(\cdot)$, and $p_1(\cdot)$.

Second, owing to Lemma 4.4, the optimization problem is subject to no inequality constraints other than the immediate constraints on the optimizers themselves.[23] Finally, it is worth noting that for $p_1(\overline{\theta}), p_1(\underline{\theta}) < 1$, objective and constraints are invariant to relabeling the first-period message space, so that there is no loss in generality in restricting the maximizers to reporting strategies such that $p_1(\overline{\theta}) \geqslant 1 - p_1(\underline{\theta})$, which formalizes the idea that the first-period *message* $\overline{\theta}$ is 'primarily designed for' the corresponding *type* $\overline{\theta}$, and similarly for the first-period report $\underline{\theta}$.

To establish a link between Proposition 4.5 and the full-commitment case, observe that the restrictions stemming from spot commitment lie entirely in constraints (a2) and (a3): It is only due to these constraints that we cannot simply set $p_1(\overline{\theta}) = p_1(\underline{\theta}) = 1$ and choose $q_1(\cdot)$ and $q_2(\cdot)$ accordingly to replicate the full-commitment outcome. Moreover, from our above discussions, the source of these constraints is intuitive: (a2) embodies the requirement that second-period contracts be sequentially rational under spot commitment, and (a3) represents the fact that first-period separation of types is not possible under spot commitment due to the ratchet effect.

4.5 Numerical Results

Rather than derive results concerning the direction of distortions at the same level of generality as in Section 3 for the full-commitment case, we confine ourselves in this section to presenting results for two specific examples. One reason for this is that, as we shall see, in contrast to the full-commitment case, the partial nature of first-period information revelation under spot commitment suggests multiple meaningful ways to measure these distortions. However, we will supply an example for which distortions in output are *upward* for all these measures, which generalizes this paper's main theme by showing that upward distortions in output are possible under learning by doing—even under spot contracts.

The examples in this section are based on specific parameterizations of Example 1 in Section 2. We determine the equilibrium sequence of contracts and reporting behavior by means of the optimization problem in Proposition 4.5. This optimization may be implemented in terms of a basic three-dimensional grid search over $(p_1(\overline{\theta}), p_1(\underline{\theta}), q_1(\overline{\theta}))$: Each such tuple uniquely determines $q_2(\overline{\theta}, \overline{\theta})$ and $q_2(\overline{\theta}, \underline{\theta})$ by requirement (a2). We may then construct the set of tuples $(q_1(\underline{\theta}), q_2(\underline{\theta}, \underline{\theta}))$ which jointly satisfy requirements (a2) and (a3). In the setup of Example 1, where c_2 and S' are linear in q_2, this results in a quadratic condition, producing at most two such tuples. Finally, for every such tuple, the resulting $q_1(\underline{\theta})$ lets us uniquely determine $q_1(\underline{\theta}, \overline{\theta})$ by requirement (a2). In this fashion, for every $(p_1(\overline{\theta}), p_1(\underline{\theta}), q_1(\overline{\theta}))$, we may determine the maximal expected profits as given by (19), which completes the

[23]Recall from Section 4.3 that the inequality constraint (10) in Proposition 4.3 can be ignored at no loss of generality by replacing condition (8) by condition (8').

	Period 1		*Period 2*			
	$q_1(\overline{\theta})$	$q_1(\underline{\theta})$	$q_2(\overline{\theta},\overline{\theta})$	$q_2(\underline{\theta},\overline{\theta})$	$q_2(\underline{\theta},\underline{\theta})$	$q_2(\overline{\theta},\underline{\theta})$
Probability	63%	37%	50%	0%	37%	13%
Spot Contract (SC)	0.341	0.444	0.565	0.604	0.576	0.319
(1) q_1^*, q_2^*	0.301	0.380	0.550	–	0.538	–
(2) $q_1^*, q_2^* \| p_1^{\text{SC}}$	0.314	0.380	0.555	0.580	0.538	0.500
(3) $q_1^* \| p_1^{\text{SC}}, q_2^{\text{SC}}$	0.301	0.396				
(4) $q_2^* \| p_1^{\text{SC}}, q_1^{\text{SC}}$			0.565	0.604	0.576	0.514

Note: Row 'Probability' shows probability of observing quantity in spot equilibrium; row 'Spot Contract (SC)' shows quantities in equilibrium spot contract; remaining rows are explained in the text.

Table 1: Evaluating Output Distortions in Example 3.

algorithm.

4.5.1 Example 3: Inefficient Agent Learns Faster

We first consider an example for which the inefficient agent learns faster, which we know under full commitment causes the inefficient agent's first-period output to be inefficiently high—at least given second-period output. Specifically, we use the same setting used in Example 2 in Section 3 to illustrate the possibility of overall upward distortions under full commitment:

Example 3. For the setting given in Example 2, equilibrium reporting strategies under spot commitment are given by $p_1(\overline{\theta}) = 1$ and $p_1(\underline{\theta}) = 0.74$ for equilibrium output menus given by $q_1(\overline{\theta}) = 0.34$, $q_1(\underline{\theta}) = 0.44$, $q_2(\overline{\theta},\overline{\theta}) = 0.57$, $q_2(\overline{\theta},\underline{\theta}) = 0.60$[24], $q_2(\underline{\theta},\underline{\theta}) = 0.58$, and $q_2(\underline{\theta},\overline{\theta}) = 0.32$.

As noted, there are multiple ways to gauge distortions in the spot contract of Example 3 (comparisons are summarized in Table 1):

(1) Comparison with Unconstrained First Best (q_1^*, q_2^*)*:* First-best quantities as described in Section 2 are shown in the third row of Table 1. Distortions are easy to evaluate for the $\overline{\theta}$-type since he produces deterministic quantities in the spot equilibrium (recall $p(\overline{\theta}) = 1$): His output is higher under the spot contract than under the dynamic first-best. Comparisons are complicated for the $\underline{\theta}$-agent by his equilibrium quantities under the spot contract being stochastic. However, his *average* first-period quantity $p_1(\underline{\theta})q_1(\underline{\theta}) + (1 - p_1(\underline{\theta}))q_1(\overline{\theta}) = 0.42$ exceeds $q_1^*(\underline{\theta})$, whereas his average second-period output of 0.51 falls short of $q_2^*(\underline{\theta})$. Average *overall* second-

[24]Since $\nu_2(\overline{\theta}) = 1$, the equilibrium value of $q_2(\overline{\theta},\underline{\theta})$ is not unique (see Proposition 4.3).

period output (i.e., averaged over both agents and reports) under the spot contract (0.54) is higher, however, than average first-best second-period output (0.52).

(2) Comparison with First Best Conditional on First-Period Reporting $(q_1^*, q_2^*|p_1^{\mathrm{SC}})$*:* Next, efficient levels of $q_1(\cdot)$ and $q_2(\cdot)$ *given agents reporting strategies under the spot contract* are given in the fourth row of Table 1. Compared to this benchmark, there is a clear upward distortion in both agents' period output and the $\overline{\theta}$-type's second-period output. The direction of the distortion in the $\underline{\theta}$-type's second-period output depends on his first-period report, but the average distortion under the spot contract is again upward.

(3) Comparison with Efficient First-Period Output Conditional on First-Period Reporting and Second-Period Output $(q_1^*|p_1^{\mathrm{SC}}, q_2^{\mathrm{SC}})$*:* Row five of Table 1 shows that given (i) agents' reporting strategies and (ii) second-period output, both type's first-period output is inefficiently high under spot contracting.

(4) Comparison with Efficient Second-Period Output Conditional on First-Period Reporting and First-Period Output $(q_2^*|p_1^{\mathrm{SC}}, q_1^{\mathrm{SC}})$*:* Not surprisingly given Proposition 4.3, the last line of Table 1 shows that there is no distortion in the $\overline{\theta}$-agent's second-period output given his first-period output. Moreover, since $p_1(\overline{\theta}) = 1$ and therefore $\nu_2(\underline{\theta}) = 0$, the $\underline{\theta}$-agent's second-period is also efficient under truthful reporting by this measure, whereas it is inefficiently low for a $\overline{\theta}$-report.

There is a final interesting comparative exercise not shown in Table 1: Imagine a principal who faces the constraints of Proposition 4.5 regarding his strategies, but who maximizes expected *surplus* rather than expected *profits*. Technically, this simply means ignoring the last term (the expected rent payment) in the objective function (19) of Proposition 4.5. Intuitively, we might think of this as first-and second-period contracts being designed by two *different* principals ('principal t', $t = 1, 2$), where the 'benevolent' principal 1 maximizes expected surplus while taking into account that principal 2 observes first-period messages and behaves opportunistically in designing the second-period contract.[25]

It turns out in the setting of Example 3 that such a benevolent first-period principal will not collect *any* first-period information: He will simply set $q_1 = 0.341$, which will entail second-period quantities of $q_2(\overline{\theta}) = 0.565$ and $q_2(\underline{\theta}) = 0.465$ conditional on the second-period report.[26] Compared to this benchmark, the above

[25] The term 'benevolent' may be somewhat misleading in the context of Laffont and Tirole's (1986) regulation model. In that model, the regulatory authority behaves like a 'rational' rather than a 'benevolent' principal in trading off joint surplus against informational rents not because it is selfish, but because it is assumed either (i) that the authority is at least partly guided by distributional concerns in favor of consumers, or (ii) that levying taxes to pay rents to the firm involves some form of deadweight loss.

[26] We note in passing that—while clearly inefficient compared to the dynamic first best—such a benevolent first-period principal still exploits learning effects in terms of Definition 3.4, since the optimal *static* first-period output in Example 3 would involve $q_1 = 0.16$ for both types.

results show that the opportunistic (i.e., rational) first-period principal will implement only an upward distortion in the $\underline{\theta}$-type's first-period quantities by offering him a first-period quantity of $q_1 = 0.444$ for a truthful report—which a $\underline{\theta}$-type will choose with probability 0.74, while it is never chosen by a $\overline{\theta}$-type.

This last comparative exercise is particularly instructive in that it provides a comprehensive account of first-period distortionary motives—distortions in first-period quantities and the distortions in first-period reporting behavior induced thereby. Moreover, it is intuitive that a 'benevolent first-period principal' will generally not induce any information revelation in the first period: Since he does not care about rents paid to the agent in the second period, he will be interested in implementing type-dependent first-period quantities only to the extent that they enhance second-period efficiency. However, second-period efficiency can only be enhanced by making it more likely that the efficient $\overline{\theta}$-type produces at a real cost advantage over the $\underline{\theta}$-type. This in turn is precluded by Lemma 4.4, by which second-period cost advantages must be identical for any first-period report due to the ratchet effect precluding any first-period separation. Hence, the benevolent first-period principal will implement report-independent first-period quantities—and it is quickly seen that this is compatible only with non-informative reporting behavior on behalf of both types.[27,28] This also illustrates that an *opportunistic* principal will not implement report-dependent first-period quantities to raise efficiency—but only to reduce rents payable in the second period, in a sense reminiscent of the usual rent-efficiency tradeoff.

Mainly, however, the discussion surrounding Example 3 has shown that, along an entire range of possible benchmarks, it is possible for learning effects to be exploited to an inefficiently high extent even under spot commitment.

4.5.2 Example 4: Efficient Agent Learns Faster

To round off this section of numerical examples, Example 4 below presents a setting in which the more efficient agent also learns faster. As one may expect, results in this example are less surprising in that the spot equilibrium entails unambiguous downward distortion in quantities traded:

Example 4. Assume the value of output to the principal is given by $S(q_t) = 100q_t - 80q_t^2$, cost structures are $c_1 = 75$, $c_2(q_1, \overline{\theta}) = 60 - 70q_1$, and $c_2(q_1, \underline{\theta}) = 65 - 60q_1$, $\nu = 0.5$, and $\delta = 0.7$. The spot equilibrium then involves no revelation of information

[27] More precisely, given a first-period message space $M_1 = \{\overline{\theta}, \underline{\theta}\}$, any report-independent quantities $q_1(\overline{\theta}) = q_1(\underline{\theta})$ and any reporting strategy $p_1(\underline{\theta})$ of the $\underline{\theta}$-type, any informative reporting strategy $p_1(\overline{\theta}) \neq 1 - p_1(\underline{\theta})$ on behalf of the $\overline{\theta}$-type will be suboptimal since revealing his type in the first period only serves to diminish his expected second-period utility.

[28] This insight is in line with the general notion that hold-up problems prevent private information on long-term investment opportunities from being put to efficient use under short-term contracts (see for instance Fudenberg et al., 1990).

	Period 1		*Period 2*			
	$q_1(\overline{\theta})$	$q_1(\underline{\theta})$	$q_2(\overline{\theta},\overline{\theta})$	$q_2(\underline{\theta},\overline{\theta})$	$q_2(\underline{\theta},\underline{\theta})$	$q_2(\overline{\theta},\underline{\theta})$
Probability	50%	50%	25%	25%	25%	25%
Spot Contract (SC)	0.237	0.237	0.354	0.354	0.262	0.252
(1) q_1^*, q_2^*	0.267	0.237	0.368	–	0.308	–
(2) $q_1^*, q_2^* \| p_1^{SC}$	0.253	0.253	0.361	0.361	0.313	0.313
(3) $q_1^* \| p_1^{SC}, q_2^{SC}$	0.245	0.245				
(4) $q_2^* \| p_1^{SC}, q_1^{SC}$			0.354	0.354	0.308	0.308

Note: See the explanation in Table 1 and in the text.

Table 2: Evaluating Output Distortions in Example 4.

in the first period with $q_1 = 0.237$, and second-period outputs (contingent only on second-period reports) of $q_2(\overline{\theta}) = 0.354$ and $q_2(\underline{\theta}) = 0.262$.

Comparisons of output distortions are shown in Table 2.[29] Relative to all benchmarks, outputs are distorted downward in the equilibrium spot contract. Moreover, in this setting, the 'benevolent first-period principal' will implement the report-independent first-period quantity of $q_1 = 0.251$, resulting in second-period quantities of $q_2(\overline{\theta}) = 0.360$ and $q_2(\underline{\theta}) = 0.267$. Hence, also relative to this benchmark, quantities are unambiguously distorted downward.

5 Conclusion

Our model has analyzed how the introduction of privately known learning capabilities into the standard dynamic model of adverse selection influences incentive design. Contrary to previous work by Lewis and Yildirim (2002a), we have considered a setting in which agents are privately informed about the *rate* at which they learn rather than just the cost side.

The focus of our investigation has been on whether this information being private leads to an under- or an overexploitation of learning effects relative to the efficient level. Under full commitment, we have shown that this crucially depends on whether learning effects let inherently more efficient agents expand their lead, or whether they enable inherently less efficient agents to catch up. In the first case, we obtain results similar to Lewis and Yildirim's in that learning effects will be under-exploited—

[29]To present the results in the same format as Table 1, Table 2 equivalently presents the equilibrium in Example 4 as a mechanism *with* first-period message but where contracts are independent of this message and reporting strategies are completely uninformative (i.e., $p_1(\overline{\theta}) = 1 - p_1(\underline{\theta})$).

the only difference being that in our case, this distortion can be so strong as to eliminate the exploitation of learning effects altogether. In the second case, we obtain entirely new results: If learning effects let inherently less efficient agents catch up, the principal has an incentive to *over*exploit learning effects. Moreover, we have shown that this effect is not driven by the full-commitment assumption: An overexploitation of learning effects may result also under spot commitment, despite the general notion that limited commitment tends to deter rather than encourage long-term investments.

More generally, our analysis has shown that in order to predict an under- or overexploitation of learning effects in dynamic adverse selection settings, it is important to identify whether these learning effects serve to magnify or to diminish existing differences in efficiency between types. Concerning vertical procurement relationships, for instance, we may seek to categorize supplying industries along these lines according to their technology. For example, consider rather simple low-tech inputs produced in more traditional 'bread-and-butter' industries where there is little scope for large technological improvements. Even if there is originally some scope for improvements through learning by doing, we would eventually expect all agents to 'catch on to the trick' (some types sooner, some later), after which there is little scope for further improvement. Thus, we would expect learning effects to quickly subside and to equalize agents' productivity. In such industries, our model would predict learning effects to be over- rather than under-exploited. In contrast, consider suppliers of more high-tech products such as the computer chip industry. Here, we would expect significant scope for long-run improvements in production technologies. Further, we would expect inherently more innovative and creative suppliers to ever increase their lead over less efficient suppliers through accumulated learning effects. For such suppliers, our model predicts learning effects to be under- rather than overexploited. Similar technological arguments may be applied to the regulation of monopolistic suppliers.

One may also imagine applications of our model outside of the realm of pure procurement and regulation settings. Take, for instance, labor contracts of the type considered in Miyazaki's (1977) 'internal labor market rat race', where employees' productivity on the job is privately known, and labor contracts specify how hard an agent is expected to work on the job. Assume, in addition, that how hard an agent works today influences his future productivity on the job. If we expect hard work to make a less efficient worker catch up with the more efficient worker's productivity, we should expect the employer to ask agents to work inefficiently hard on the job—in a sense interpretable as an aggravation of the 'rat race'. If, on the other hand, we expect harder work today to magnify productivity differences between workers (as might be the case on more creative jobs), we should expect the employer to relax workers' workload below the efficient level.

Other potential fields of application may include credit market problems in the style of Freixas and Laffont (1990), where the borrower is privately informed about the returns to his project. If the future productivity of capital in the project systematically depends on the size of the loan today, then our analysis would predict inefficiently high first-period loans if raising loans has a larger impact on productivity in inherently less productive projects, and vice versa if productivity in inherently more productive projects is more strongly affected. Finally, the insights may be applied to models of discrimination in quantity or quality by a monopolistic supplier of a consumption good (see Maskin and Riley, 1984; Mussa and Rosen, 1978) if we assume that consumers get used to or even addicted to the good, so that consuming more (or a higher quality) of the good today increases consumers' willingness to pay tomorrow.[30] The learning-speed criterion of our model then pertains to whether customers with a higher willingness to pay for the good also get used to the good faster, or whether it is the customers who value the good less who get used to it at a faster rate. Our analysis predicts an underexploitation of the 'addiction factor' in the former case, and an overexploitation in the latter.

Acknowledgements

This work has benefited from comments by Bernard Salanié and seminar participants at the University of Zürich. Moreover, I am indebted to Armin Schmutzler and Jean Tirole for invaluable discussions based on earlier versions. Financial support for this project from the Swiss National Science Foundation under Grant No. PBZH1-108484 is gratefully acknowledged.

Appendix: Proofs

A.1 Proofs of Results in Section 2

Proof of Lemma 2.1. Using elementary robust comparative statics and supermodular analysis (cf. Topkis, 1998), the results require identifying suitable complementarity relations among q_1, q_2 and θ in the objective function W. Observe first that $\frac{\partial^2}{\partial q_1 \partial q_2} W = -\delta \frac{\partial}{\partial q_1} c_2 > 0$, so that W has increasing differences in $\mathbf{q}$, implying that $\hat{q}_1^*$ and $\hat{q}_2^*$ are both increasing in

[30] The idea of such 'rational addiction' has been introduced by Becker and Murphy (1988), albeit in the context of a competitive market. Boone and Shapiro (2005) have more recently investigated a setting quite similar to the application we have in mind here, the main difference being that the discriminating monopolist sells on 'anonymous' spot markets, that is, any information revealed to him in the first-period is useless in the second-period). However, in many settings, even if the seller has a sizeable number of clients, he can nonetheless design a 'personalized' sequence of contracts, as is done for instance by video rental chains by means of membership cards.

the quantity produced in the other period. Next, $\frac{\partial}{\partial q_2}W = -\delta c_2$, which is increasing in θ since we have assumed c_2 to decrease in θ. Hence, W has increasing differences in (q_2, θ), implying that $\hat{q}_1^*$ is increasing in θ. Finally, $\frac{\partial}{\partial q_1}W = -\delta q_2 \frac{\partial}{\partial q_1}c_2$. Thus, W has increasing differences in (q_1, θ) if more efficient agents learn faster ($|\partial c_2/\partial q_1|$ increasing in θ), and decreasing differences if less efficient agents learn faster, yielding the comparative static result for $\hat{q}_1^*$ in θ. Finally, if more efficient agents learn faster, W will thereby have increasing differences in all pairs of arguments, implying that $\mathbf{q}^*$ is increasing in θ. □

A.2 Proofs of Results in Section 3

Proof of Proposition 3.1. Using the Φ-function defined in (3), the incentive constraints may be written compactly as

$$\overline{U} \geqslant \underline{U} + \Phi(\underline{\mathbf{q}}) \tag{A.1}$$

$$\underline{U} \geqslant \overline{U} - \Phi(\overline{\mathbf{q}}). \tag{A.2}$$

Using the surplus functions $\overline{W}$ and $\underline{W}$ and agents' rents $\overline{U}$ and $\underline{U}$, the principal's payoff from any contract may be written as

$$\nu[\overline{W}(\overline{\mathbf{q}}) - \overline{U}] + (1-\nu)[\underline{W}(\underline{\mathbf{q}}) - \underline{U}]. \tag{A.3}$$

The principal maximizes this payoff by choice of $\{\overline{\mathbf{q}}, \overline{U}\}$ and $\{\underline{\mathbf{q}}, \underline{U}\}$, subject to incentive constraints (A.1) and (A.2), and subject to the participation constraints, which we restate here for easy reference:

$$\overline{U} \geqslant 0 \tag{A.4}$$

$$\underline{U} \geqslant 0. \tag{A.5}$$

Observe first that only allocations satifsfying the *implementability condition*

$$\Phi(\overline{\mathbf{q}}) \geqslant \Phi(\underline{\mathbf{q}}) \tag{A.6}$$

can be realized. This condition follows from combining (A.1) and (A.2).

Next, we argue that for any menu of allocations $\{\overline{\mathbf{q}}, \underline{\mathbf{q}}\}$ satisfying (A.6), the principal will optimally set $\underline{U} = 0$ and $\overline{U} = \Phi(\underline{\mathbf{q}})$. To see this, note first that constraint (A.4) may be neglected: Since $\Phi \geqslant 0$, it is implied by (A.1) and (A.5). Given this insight, the principal must optimally set $\underline{U} = 0$: Otherwise, he could decrease $\underline{U}$ and $\overline{U}$ by the same small amount without violating any of the remaining constraints, and thereby strictly increase his payoff (A.3). But then the remaining constraints, (A.1) and (A.2), simplify to $\Phi(\underline{\mathbf{q}}) \leqslant \overline{U} \leqslant \Phi(\overline{\mathbf{q}})$, so that the principal must optimally set $\overline{U} = \Phi(\underline{\mathbf{q}})$: If not, we could decrease $\overline{U}$ by a small amount without violating any of the remaining constraints. Finally, this implies that the only remaining constraint, (A.2), simply becomes (A.6).

Hence, the principal's optimization problem may be restated as choosing $\overline{\mathbf{q}}$ and $\underline{\mathbf{q}}$ so as to maximize

$$\nu[\overline{W}(\overline{\mathbf{q}}) - \Phi(\underline{\mathbf{q}})] + (1-\nu)\underline{W}(\underline{\mathbf{q}}) \tag{A.7}$$

subject to (A.6). Due to the additive separable structure and since $\nu \in (0,1)$, this is equivalent to condition (4). To complete the proof, it therefore remains to be shown that any $\{\overline{\mathbf{q}}, \underline{\mathbf{q}}\}$ satisfying (4) also satisfy (A.6). The former implies in particular that

$$\overline{W}(\overline{\mathbf{q}}) \geqslant \overline{W}(\underline{\mathbf{q}}) \quad \text{and} \quad \underline{W}(\underline{\mathbf{q}}) - \frac{\nu}{1-\nu}\Phi(\underline{\mathbf{q}}) \geqslant \underline{W}(\overline{\mathbf{q}}) - \frac{\nu}{1-\nu}\Phi(\overline{\mathbf{q}}). \tag{A.8}$$

Since $\overline{W}(\mathbf{q}) - \underline{W}(\mathbf{q}) = \Phi(\mathbf{q})$ for any $\mathbf{q}$ by the definition of $\overline{W}$ and $\underline{W}$, the inequalities in (A.8) may be added to yield $\Phi(\overline{\mathbf{q}})/(1-\nu) \geqslant \Phi(\underline{\mathbf{q}})/(1-\nu)$, which implies (A.6). □

Proof of Proposition 3.2. We again employ supermodular analysis to derive the results. To this end, define the real-valued function $g(\mathbf{q};\tau)$ such that

$$g(\mathbf{q};\tau) = \begin{cases} \underline{W}(\mathbf{q}), & \text{for } \tau = 0, \\ \underline{W}(\mathbf{q}) - \frac{\nu}{1-\nu}\Phi(\mathbf{q}), & \text{for } \tau = 1. \end{cases} \tag{A.9}$$

(a) For any $\mathbf{q}$,

$$\tfrac{\partial}{\partial q_2} g(\mathbf{q};1) - \tfrac{\partial}{\partial q_2} g(\mathbf{q};0) = -\tfrac{\nu}{1-\nu}\,\tfrac{\partial}{\partial q_2}\Phi(\mathbf{q}), \tag{A.10}$$

which is strictly negative (see the definition of Φ). Thus, for any q_1, g has strictly decreasing differences in q_2 and τ, implying that $\arg\max_{q_2} g(q_1, q_2; \tau)$ is decreasing in τ for any q_1. By definition of g, this set corresponds to the conditional second-period first-best for $\tau = 0$ and to the conditional second-period second-best, which proves the claim.

(b) For any $\mathbf{q}$,

$$\tfrac{\partial}{\partial q_1} g(\mathbf{q};1) - \tfrac{\partial}{\partial q_1} g(\mathbf{q};0) = -\tfrac{\nu}{1-\nu}\,\tfrac{\partial}{\partial q_1}\Phi(\mathbf{q}), \tag{A.11}$$

the sign of which depends on whether the more efficient $\overline{\theta}$-agent also learns faster ($\frac{\partial}{\partial q_1}\Phi(\mathbf{q}) \geqslant 0$) or whether the $\underline{\theta}$-agent learns faster ($\frac{\partial}{\partial q_1}\Phi(\mathbf{q}) \leqslant 0$), with each of these inequalities being strict for $q_2 > 0$. Thus, whenever the more efficient agent learns faster, g has decreasing differences in q_1 and τ for any q_2 (and strictly so for any $q_2 > 0$), which proves claim (bi). On the other hand, whenever the less efficient agent learns faster, g has *increasing* differences in q_1 and τ for any q_2, thereby proving part (bii). □

Proof of Proposition 3.3. We again employ the auxiliary function $g(\mathbf{q};\tau)$ defined in (A.9) in the proof of Proposition 3.2, so that $\underline{\mathbf{q}}^* \in \arg\max_{\mathbf{q}} g(\mathbf{q};0)$ and $\underline{\mathbf{q}}^{\mathrm{SB}} \in \arg\max_{\mathbf{q}} g(\mathbf{q};1)$. As established there, for the case at hand, g has strictly decreasing differences in q_2 and τ for any q_1, and strictly decreasing differences in q_1 and τ for any $q_2 > 0$. Hence, g has strictly decreasing differences in $(\mathbf{q}, \tau)$. Moreover, g is supermodular in $\mathbf{q}$ for $\tau = 0$ since W is supermodular in $\mathbf{q}$ (see the proof of Proposition 2.1). Hence, the set of maximizers of g is decreasing in τ, which proves the claim.[31] □

[31] Note that supermodularity of g in $\mathbf{q}$ for $\tau = 1$ is *not* required for the proof (and not generally satisfied). The interested reader is invited to verify that, given a *binary* parameter space, Theorem 2.8.1 in Topkis (1998) in fact requires supermodularity of the objective for only *one* of the two parameter values: Together with increasing differences, this is easily seen to imply Topkis' condition (2.8.1), which produces the result by Lemma 2.8.1.

Proof of Proposition 3.5. Observe first that $\underline{\hat{q}_1}^{SB}(0) = q_1^\circ$ since the principal's optimal choice of $\underline{q_1}$ conditional on $q_2 = 0$ simply maximizes first-period surplus.[32] Moreover, $\underline{\hat{q}_1}^{SB}(\overline{q_2^{SB}}) = q_1^{SB}$ by definition of the full and conditional optima. Since $q_2^{SB} \geqslant 0$, learning effects will thus be exploited if $\underline{\hat{q}_1}^{SB}(q_2)$ is increasing in q_2, and unexploited if it is decreasing. This in turn depends on whether the principal's objective $\underline{W} - \nu\Phi$ has increasing or decreasing differences in (q_1, q_2). Using the definitions of $\underline{W}$ and Φ, we have

$$\tfrac{\partial^2}{\partial q_1 \partial q_2}\left\{\underline{W}(\mathbf{q}) - \tfrac{\nu}{1-\nu}\Phi(\mathbf{q})\right\} = \tfrac{\delta}{1-\nu}\left[\left|\tfrac{\partial}{\partial q_1}\underline{c_2}(q_1)\right| - \nu \cdot \left|\tfrac{\partial}{\partial q_1}\overline{c_2}(q_1)\right|\right],$$

so that the former will be the case whenever condition (6) holds (and the latter whenever (6) is reversed), which completes the proof. □

A.3 Proofs of Results in Section 4

Proof of Lemma 4.1. Consider an arbitrary first-period mechanism. Let $\overline{z_1}$ and $\overline{q_1}$ denote the resulting first-period transfer and quantities for the $\overline{\theta}$-type (if the equilibrium outcome is not deterministic, let them denote any particular realization), and $\underline{z_1}$ and $\underline{q_1}$ those for the $\underline{\theta}$-type. If agents truthfully reveal their type through this mechanism, then the principal knows the true type when offering the new contract at the beginning of the second period. Hence, the second-period contract will enable the principal to extract the full second-period surplus and leave either type its reservation payoff of 0.

Now consider agents' incentives to mimic the other type. First, consider the $\underline{\theta}$-agent. By mimicking the $\overline{\theta}$-agent in the first period, he obtains a first-period rent of $\overline{z_1} - c_1\overline{q_1}$. In the second period, he can always quit (given that the more efficient agent obtains no second-period rent, this will typically be optimal), giving him a second-period payoff of 0. Hence, the $\underline{\theta}$-type not mimicking the $\overline{\theta}$-type in the first period can only be optimal if

$$\underline{z_1} - c_1\underline{q_1} \geqslant \overline{z_1} - c_1\overline{q_1}.$$

Next, consider a $\overline{\theta}$-type mimicking the less efficient $\underline{\theta}$-type. This gives him a rent of $\underline{z_1} - c_1\underline{q_1}$ in the first period, and second-period marginal costs below those of the $\underline{\theta}$-type. Since the $\underline{\theta}$-type receives a zero second-period rent from his contract, the more efficient $\overline{\theta}$-type can always obtain a positive second-period rent of $\Phi(\underline{q_1}, \underline{q_2})/\delta$ from mimicking the inefficient type also in the second period.[33] Thus, the efficient $\overline{\theta}$-type not mimicking the inefficient $\underline{\theta}$-type can only be optimal if

$$\overline{z_1} - c_1\overline{q_1} \geqslant \underline{z_1} - c_1\underline{q_1} + \Phi(\underline{q_1}, \underline{q_2}).$$

Adding up the two incentive constraints shows that incentive compatibility requires that $\Phi(\underline{q_1}, \underline{q_2}) \leqslant 0$. Since Φ is nonnegative by construction, this immediately implies that $\Phi(\underline{q_1}, \underline{q_2}) = 0$, which in turn implies that $\overline{z_1} - c_1\overline{q_1} = \underline{z_1} - c_1\underline{q_1}$, so that agents must obtain the same first-period rent. □

[32] The same is true of the conditional first-best output, i.e. $\underline{\hat{q}_1}^*(0) = q_1^\circ$, which is why the $\underline{\hat{q}_1}^{SB}$- and $\underline{\hat{q}_1}^*$-curves meet at $q_2 = 0$ in Figure 2.

[33] To see this, note that the $\overline{\theta}$-agent's obtains a second-period rent of $\underline{z_2} - c_2(\underline{q_1}, \overline{\theta})\underline{q_2}$ from mimicking the $\underline{\theta}$-agent. Given that the $\underline{\theta}$-agent's second-period rent of $\underline{z_2} - c_2(\underline{q_1}, \underline{\theta})\underline{q_2}$ is zero, the claim follows from the definition of Φ.

Proof of Proposition 4.3. In terms of rent/output-contracts and suppressing the conditioning of q_2, u_2 and ν_2 on $\tilde{\theta}_1$, the principal's problem may be written as choosing $\{q_2(\tilde{\theta}_2), u_2(\tilde{\theta}_2)\}$ so as to maximize

$$\nu_2 \cdot \{S[q_2(\overline{\theta})] - c_2(q_1, \overline{\theta}) \cdot q_2(\overline{\theta})\} + (1 - \nu_2) \cdot \{S[q_2(\underline{\theta})] - c_2(q_1, \underline{\theta}) \cdot q_2(\underline{\theta})\} \\ - \nu_2 \cdot u_2(\overline{\theta}) - (1 - \nu_2) u_2(\underline{\theta})$$

subject to the constraints

$$\begin{aligned} u_2(\overline{\theta}) &\geqslant u_2(\underline{\theta}) + \Phi[q_1, q_2(\underline{\theta})]/\delta \\ u_2(\underline{\theta}) &\geqslant u_2(\overline{\theta}) - \Phi[q_1, q_2(\overline{\theta})]/\delta \\ u_2(\overline{\theta}) &\geqslant 0 \\ u_2(\underline{\theta}) &\geqslant 0. \end{aligned}$$

Adding the first two constraints immediately yields the (second-period) implementability condition (10). Replicating the argument for the full-commitment case, it is immediately seen that the principal will optimally set $u_2(\underline{\theta}) = 0$ and $u_2(\overline{\theta}) = \Phi[q_1, q_2(\underline{\theta})]/\delta$ given *any* $q_2(\tilde{\theta}_2)$, so that the quantities specified in the contract must maximize

$$\nu_2 \cdot \{S[q_2(\overline{\theta})] - c_2(q_1, \overline{\theta}) \cdot q_2(\overline{\theta})\} + (1 - \nu_2) \cdot \{S[q_2(\underline{\theta})] - c_2(q_1, \underline{\theta}) \cdot q_2(\underline{\theta})\} \\ - \nu_2 \cdot \Phi[q_1, q_2(\underline{\theta})]/\delta \quad \text{(A.12)}$$

subject to (10). Noting that (A.12) is additively separable in terms depending on $q_2(\overline{\theta})$ and on $q_2(\underline{\theta})$ immediately yields the characterization of the optimal second-period quantities given in (8) through (10).

Finally, to show that (10) holds whenever $\nu_2 > 0$, observe that (8) and (9) imply in particular that

$$\nu_2\{S[q_2(\overline{\theta})] - c_2(q_1, \overline{\theta}) q_2(\overline{\theta})\} \geqslant \nu_2\{S[q_2(\underline{\theta})] - c_2(q_1, \overline{\theta}) q_2(\underline{\theta})\} \quad \text{(A.13)}$$

and

$$(1 - \nu_2)\{S[q_2(\underline{\theta})] - c_2(q_1, \underline{\theta}) q_2(\underline{\theta})\} - \nu_2 \Phi[q_1, q_2(\underline{\theta})] \geqslant \\ (1 - \nu_2)\{S[q_2(\overline{\theta})] - c_2(q_1, \underline{\theta}) q_2(\overline{\theta})\} - \nu_2 \Phi[q_1, q_2(\overline{\theta})] \quad \text{(A.14)}$$

Multiplying condition (A.13) by $(1 - \nu_2)/\nu_2$ and combining it with condition (A.14) yields (10). □

To prove Lemma 4.4, we use the following properties of the second-period continuation outcome as described by Proposition 4.3:

Lemma A.1. *For any first-period strategies $q_1(\cdot), z_1(\cdot)$ and any first-period report $\tilde{\theta}_1$,*

(a) if $\nu_2(\tilde{\theta}_1) \in (0,1)$, the set of $q_2(\tilde{\theta}_1, \cdot)$ which satisfy Proposition 4.3 is unique. Denoting this solution by $\hat{q}_2(\tilde{\theta}_1, \cdot)$, both $\hat{q}_2(\tilde{\theta}_1, \underline{\theta})$ and $\Phi[q_1(\cdot), \hat{q}_2(\tilde{\theta}_1, \underline{\theta})]$ are continuous and strictly decreasing in $\nu_2(\tilde{\theta}_1)$ whenever $\Phi[q_1(\cdot), \hat{q}_2(\tilde{\theta}_1, \underline{\theta})] > 0$;

(b) if $\nu_2(\tilde{\theta}_1) = 1$, then $\Phi[q_1(\tilde{\theta}_1), q_2(\tilde{\theta}_1, \underline{\theta})] = 0$ for any $q_2(\tilde{\theta}_1, \underline{\theta})$ satisfying Proposition 4.3.

Proof of Lemma A.1. (a) Note first that the objective function is strictly concave for $\nu_2(\tilde{\theta}_1) \in (0,1)$ due to the assumption that $S'' < 0$ and the linearity of the second-period cost function (and thereby Φ) in q_2. This implies uniqueness, continuity, and validity of the first-order approach for interior solutions to condition (8) in Proposition 4.3. Dropping the conditioning on $\tilde{\theta}_1$, the first-order condition for $q_2(\underline{\theta})$ is

$$S'[q_2(\underline{\theta})] = c_2(q_1, \underline{\theta}) + \frac{\nu_2}{1-\nu_2}[c_2(q_1, \underline{\theta}) - c_2(q_1, \overline{\theta})]. \tag{A.15}$$

(Note that the first-order condition must bind since $\Phi > 0$ implies that $q_2(\underline{\theta}) > 0$.) Since $\Phi > 0$, $c_2(q_1, \underline{\theta}) - c_2(q_1, \overline{\theta}) > 0$, so that the right-hand side of (A.15) is strictly increasing in ν_2. With the assumption that $S'' < 0$, the claim that $q_2(\underline{\theta})$ is strictly increasing immediately follows. Moreover, for $\Phi > 0$, Φ is strictly increasing in $q_2(\underline{\theta})$, which proves the second claim.

(b) For $\nu_2(\tilde{\theta}_1) = 1$, the principal attaches no probability to the agent being of type $\underline{\theta}$ and therefore—by condition(9) in Proposition 4.3—will set $q_2(\tilde{\theta}_1, \underline{\theta})$ with the sole objective of eliminating the $\overline{\theta}$-agent's rent, which he can always achieve by setting $q_2(\tilde{\theta}_1, \underline{\theta}) = 0$, so that $\Phi[q_1(\tilde{\theta}_1), q_2(\tilde{\theta}_1, \underline{\theta})] = 0$. □

Proof of Lemma 4.4. For $p_1(\overline{\theta}), p_1(\underline{\theta}) < 1$, (17) must obviously bind by (13) and (14): If both agents are indifferent between first-period messages, both incentive constraints (and thereby (17)) must bind. For $p_1(\underline{\theta}) = 1$, this is implied by sequential rationality: $p_1(\underline{\theta}) = 1$ implies $\nu_2(\overline{\theta}) = 1$, so $\Phi[q_1(\overline{\theta}), q_2(\overline{\theta}, \underline{\theta})] = 0$ by Lemma A.1(b), which in turn implies $\Phi[q_1(\underline{\theta}), q_2(\underline{\theta}, \underline{\theta})] = 0$ by (17) and the fact that $\Phi \geqslant 0$ by assumption.

It thus remains to be argued that (17) must bind for $p(\overline{\theta}) = 1$ and $p_1(\underline{\theta}) < 1$. This is the only constellation of $p_1(\overline{\theta})$ and $p_1(\underline{\theta})$ in which the principal *could* leave (17) slack without violating either incentive compatibility or sequential rationality. However, in this case, doing so will always yield him a suboptimal level of profits.

To this end, consider any set of strategies $\mathcal{S}' \equiv \{p'_1, q'_1, q'_2\}$[34] such that (i) $p'_1(\overline{\theta}) = 1$ and $p'_1(\underline{\theta}) < 1$, such that (ii) $q'_2(\cdot)$ satisfies Proposition 4.3 (i.e., sequential rationality), and such that (iii) the implementability condition (17) is slack (i.e., such that $\Phi[q'_1(\overline{\theta}), q'_2(\overline{\theta}, \underline{\theta})] > \Phi[q'_1(\underline{\theta}), q'_2(\underline{\theta}, \underline{\theta})]$). We will show that $\mathcal{S}'$ cannot maximize the principal's expected payoff by constructing two alternative sets of strategies, $\mathcal{S}''$ and $\mathcal{S}'''$, both of which satisfy condition

[34]To ease notation, we suppress the strategies z_1 and z_2. Note that for any p_1, q_1, q_2, second-period rents are determined by Proposition 4.3, and optimal first-period rents satisfy $U(\underline{\theta}) = 0$ and $U(\overline{\theta}) = \Phi[q_1(\overline{\theta}), q_2(\overline{\theta}, \underline{\theta})]$ (see the discussion in the text immediately prior to Lemma 4.4), which in turn implicitly determines the corresponding strategies z_1, z_2.

(17) and Proposition 4.3 (i.e., sequential optimality of $q_2(\cdot)$), and by arguing that either $\mathcal{S}''$ or $\mathcal{S}'''$ must yield a strictly higher expected payoff to the principal.

Consider first the strategies $\mathcal{S}'' \equiv \{p_1'', q_1'', q_2''\}$, which differ from $\mathcal{S}'$ only in $p_1(\underline{\theta})$ and $q_2(\overline{\theta}, \underline{\theta})$: We choose $p_1''(\overline{\theta}) > p_1'(\overline{\theta})$ and $q_2''(\overline{\theta}, \underline{\theta})$ such that $q_2''(\overline{\theta}, \underline{\theta})$ satisfies Proposition 4.3 and that $\mathcal{S}''$ satisfies the implementability condition (17), that is, that

$$\Phi[q_1'(\overline{\theta}), q_2''(\overline{\theta}, \underline{\theta})] \geqslant \Phi[q_1'(\underline{\theta}), q_2'(\underline{\theta}, \underline{\theta})]. \tag{A.16}$$

This is possible by choosing $p_1''(\overline{\theta}) > p_1'(\overline{\theta})$ small enough, due to Lemma A.1.[35]

Before comparing the payoffs resulting under $\mathcal{S}'$ and $\mathcal{S}''$, we need to ensure that setting $q_2''(\cdot) = q_2'(\cdot)$ for the remaining elements of $\mathcal{S}''$ does not lead $\mathcal{S}''$ to violate Proposition 4.3. But given that $q_1'(\underline{\theta}, \cdot)$ satisfy Proposition 4.3 for $\mathcal{S}'$, they do also for $\mathcal{S}''$ since $\nu_2(\underline{\theta}) = 0$ for both strategy sets. The same is true of $q_2(\overline{\theta}, \overline{\theta})$ since $\nu_2(\overline{\theta}) > 0$ for both $\mathcal{S}'$ and $\mathcal{S}''$.

Hence, we may proceed to compare the principal's payoffs under $\mathcal{S}'$ and $\mathcal{S}''$. To avoid tedious notation and make the proof more transparent, for any strategy set $\mathcal{S}$ and any report $\tilde{\theta}_1, \theta$, we let $\Gamma_1(\tilde{\theta}) \equiv S[q_1(\tilde{\theta}_1)] - c_1 q_1(\tilde{\theta}_1)$ and $\Gamma_2(\tilde{\theta}_1, \theta) \equiv S[q_2(\tilde{\theta}_2, \theta)] - c_2[q_1(\tilde{\theta}_1), \theta] q_2(\tilde{\theta}_2, \theta)$ denote the first- and second-period surplus generated by an agent of type θ who reports $\tilde{\theta}_1$ in the first period, and we let $\hat{\Phi}(\tilde{\theta}_1, \underline{\theta}) \equiv \Phi[q_1(\tilde{\theta}_1), q_2(\tilde{\theta}_1, \underline{\theta})]$ denote the second-period rent obtained by a $\overline{\theta}$-agent who reports $\tilde{\theta}_1$ in the first period. Moreover, we let $\Pi(\mathcal{S})$ denote the principal's expected profits under $\mathcal{S}$.

Using this notation, the principal's profits under any $\mathcal{S}$ with $p_1(\overline{\theta}) = 1$ may be compactly written as

$$\begin{aligned}\Pi(\mathcal{S}) = \quad & \nu\{\Gamma_1(\overline{\theta}) + \delta\Gamma_2(\overline{\theta}, \overline{\theta}) - \hat{\Phi}(\overline{\theta}, \underline{\theta})\} \\ & +(1-\nu)p_1(\underline{\theta})\{\Gamma_1(\underline{\theta}) + \delta\Gamma_2(\underline{\theta}, \underline{\theta})\} \\ & +(1-\nu)(1-p_1(\underline{\theta}))\{\Gamma_1(\overline{\theta}) + \delta\Gamma_2(\overline{\theta}, \underline{\theta})\}.\end{aligned} \tag{A.17}$$

where, to determine rents payable to the agent, we have used the insight that $U(\underline{\theta}) = 0$ (see the discussion in the main text), which for $p_1(\underline{\theta}) < 1$ in turn implies $U(\overline{\theta}) = \Phi[q_1(\overline{\theta}), q_2(\overline{\theta}, \underline{\theta})]$ by (14), that is, by the fact that the $\underline{\theta}$-agent's incentive constraint must bind.

Using (A.17) along with the fact that the strategy sets $\mathcal{S}'$ and $\mathcal{S}''$ differ only in $p_1(\underline{\theta})$ and $q_2(\overline{\theta}, \underline{\theta})$, and thereby only in $p_1(\underline{\theta})$, $\Gamma_2(\overline{\theta}, \underline{\theta})$, and $\hat{\Phi}(\overline{\theta}, \underline{\theta})$ in terms of the 'reduced form',

---

[35] Recall that by Lemma A.1, for any given $q_1(\cdot)$, $\Phi[q_1(\overline{\theta}), q_2(\overline{\theta}, \underline{\theta})]$ is strictly, but *continuously* decreasing in $\nu_2(\overline{\theta})$ for $\nu_2(\overline{\theta}) \in (0,1)$ and $\Phi[q_1(\overline{\theta}), q_2(\overline{\theta}, \underline{\theta})] > 0$. Here, under $\mathcal{S}'$, $\nu_2'(\overline{\theta}) \in (0,1)$ since $p_1'(\overline{\theta}) = 1$ and $p_1'(\underline{\theta}) < 1$, and $\Phi[q_1'(\overline{\theta}), q_2'(\overline{\theta}, \underline{\theta})] > 0$ since otherwise, (17) could not have been slack under $\mathcal{S}'$ in the first place. Thus, starting from $\mathcal{S}'$, we may marginally increase $p_1'(\underline{\theta})$ and adjust $q_2'(\overline{\theta}, \underline{\theta})]$ in accordance with Proposition 4.3 *without* violating (17).

Moreover, while not essential for the proof, we note that it is in fact always possible to choose $p_1''(\underline{\theta}) > p_1'(\underline{\theta})$ such that (A.16) binds by choosing $p_1''(\underline{\theta})$ large enough: for $p_1''(\underline{\theta}) = 1$, $\nu_2''(\overline{\theta}) = 0$, so that $\Phi[q_1'(\overline{\theta}), q_2''(\overline{\theta}, \underline{\theta})] = 0$ by Lemma A.1(b). By the continuity of $\Phi[q_1'(\overline{\theta}), q_2''(\overline{\theta}, \underline{\theta})]$ in $\nu_2(\overline{\theta})$, there must therefore exist a $p_1''(\overline{\theta}) \in (p_1'(\overline{\theta}), 1]$ for which (A.16) binds.

we obtain

$$\begin{aligned}\Pi(\mathcal{S}'') - \Pi(\mathcal{S}') \\ &= (1-\nu)(p_1''(\underline{\theta}) - p_1'(\underline{\theta}))\Big\{\big[\Gamma_1'(\underline{\theta}) + \delta\Gamma_2'(\underline{\theta},\underline{\theta})\big] - \big[\Gamma_1'(\overline{\theta}) + \delta\Gamma_2'(\overline{\theta},\underline{\theta})\big]\Big\} \\ &\quad + (1-\nu)(1-p_1''(\underline{\theta}))\delta\big[\Gamma_2''(\overline{\theta},\underline{\theta}) - \Gamma_2'(\overline{\theta},\underline{\theta})\big] - \nu\big[\hat{\Phi}''(\overline{\theta},\underline{\theta}) - \hat{\Phi}'(\overline{\theta},\underline{\theta})\big]. \quad (A.18)\end{aligned}$$

By (A.18), the difference in payoffs between $\mathcal{S}''$ and $\mathcal{S}'$ may be decomposed into (i) the effect of the increase in $p_1(\underline{\theta})$ (first term on right-hand side), and (ii) the effect of the ensuing decrease in $q_2(\underline{\theta},\underline{\theta})$ (second term).

To obtain a lower bound on the second term, we use the fact that $q_2''(\overline{\theta},\underline{\theta})$ must be sequentially optimal in $\mathcal{S}''$ as described Proposition 4.3, with $\nu_2''(\overline{\theta}) = \nu/[\nu+(1-p_1''(\underline{\theta}))(1-\nu)]$ and $q_1''(\overline{\theta}) = q_1'(\overline{\theta})$, so

$$q_2''(\overline{\theta},\underline{\theta}) \in \arg\max_{q_2}(1-\nu)(1-p_1''(\underline{\theta}))\delta\big[S(q_2) - c_2[q_1'(\overline{\theta}),\underline{\theta}]q_2\big] - \nu\Phi[q_1'(\overline{\theta}),q_2], \quad (A.19)$$

which implies in particular that

$$(1-\nu)(1-p_1''(\underline{\theta}))\delta\big[\Gamma_2''(\overline{\theta},\underline{\theta}) - \Gamma_2'(\overline{\theta},\underline{\theta})\big] - \nu\big[\hat{\Phi}''(\overline{\theta},\underline{\theta}) - \hat{\Phi}'(\overline{\theta},\underline{\theta})\big] > 0. \quad (A.20)$$

Inequality (A.20) is strict due to the result that maximizers of (A.19) are unique and strictly increasing in $p_1(\underline{\theta})$ by Lemma A.1. Hence, letting

$$\kappa' \equiv (1-\nu)\big\{\big[\Gamma_1'(\underline{\theta}) + \delta\Gamma_2'(\underline{\theta},\underline{\theta})\big] - \big[\Gamma_1'(\overline{\theta}) + \delta\Gamma_2'(\overline{\theta},\underline{\theta})\big]\big\} \quad (A.21)$$

denote the (weighted) increase in discounted surplus from an agent of type $\underline{\theta}$ reporting truthfully rather than lying under $\mathcal{S}'$, we have

$$\Pi(\mathcal{S}'') - \Pi(\mathcal{S}') > (p_1''(\underline{\theta}) - p_1'(\underline{\theta})) \cdot \kappa, \quad (A.22)$$

implying that $\mathcal{S}''$ will strictly dominate $\mathcal{S}'$ in terms of the principal's profits whenever under $\mathcal{S}'$, the $\overline{\theta}$-agent generates a (weakly) larger surplus by truth-telling rather than lying.[36]

If this is *not* the case, so the $\overline{\theta}$-agent's surplus is strictly higher under false reporting, then the separation of types in the first period is not profitable in the first place and the principal will find it profitable to *decrease* $p_1(\underline{\theta})$ relative to $\mathcal{S}'$. We will show particularly that a strategy set involving *no first-period separation at all* will be more profitable than $\mathcal{S}'$ whenever $\kappa' < 0$.

To this end, define the strategy set $\mathcal{S}'''$ such that, for any report $\tilde{\theta}_1 \in \{\overline{\theta},\underline{\theta}\}$, $q_1'''(\tilde{\theta}_1) = q_1'(\overline{\theta})$, $q_2'''(\tilde{\theta}_1,\overline{\theta}) = q_2'(\overline{\theta},\overline{\theta})$, reporting strategies are uninformative (so $p_1'''(\overline{\theta}) = 1 - p_1'''(\underline{\theta})$, implying that $\nu_2(\overline{\theta}) = \nu_2(\underline{\theta}) = \nu$), and $q_2'''(\overline{\theta},\underline{\theta}) = q_2'''(\underline{\theta},\underline{\theta})$ is such as to satisfy Proposition

[36]Note that the argument presented can essentially be understood as an application of the envelope theorem—after noting that for any given $q_1(\cdot)$, first- and second-period objectives are congruent and sequential optimization will therefore lead to a maximization of overall discounted profits, and that (as argued in the proof of Lemma A.1) the second-period objective is concave in second-period quantities.

4.3. Note that the implementability condition (17) trivially binds since output schedules are independent of first-period messages. Moreover, sequential optimality of $q_2'(\tilde{\theta}_1, \overline{\theta})$ in $\mathcal{S}'''$ is again implied by the sequential optimality in $\mathcal{S}'''$ since $\nu_2 > 1$ in both cases. To simplify notation, we suppress conditioning on the (meaningless) report $\tilde{\theta}_1$, letting $q_1''' \equiv q_1'''(\overline{\theta}) = q_1'''(\underline{\theta})$, letting $q_2'''(\overline{\theta}) \equiv q_2'''(\overline{\theta}, \overline{\theta}) = q_2'''(\underline{\theta}, \overline{\theta})$, and $q_2'''(\underline{\theta}) \equiv q_2'''(\overline{\theta}, \underline{\theta}) = q_2'''(\underline{\theta}, \underline{\theta})$ (and similarly suppressing $\tilde{\theta}_1$ in Γ_t''' and $\hat{\Phi}'''$), so that the principal's profits under $\mathcal{S}'''$ may be written as

$$\Pi(\mathcal{S}''') = \nu[\Gamma_1'(\overline{\theta}) + \delta\Gamma_2'(\overline{\theta}, \overline{\theta}) - \hat{\Phi}'''(\underline{\theta})] + (1-\nu)[\Gamma_1'(\overline{\theta}) + \delta\Gamma_2'''(\underline{\theta})], \tag{A.23}$$

so that—again using the generic characterization of payoffs in (A.17) for $\mathcal{S}'$ —the difference in payoffs between $\mathcal{S}'''$ and $\mathcal{S}'$ becomes

$$\begin{aligned}\Pi(\mathcal{S}''') - \Pi(\mathcal{S}') = (1-\nu)p_1'(\underline{\theta})\Big\{\big[\Gamma_1'(\overline{\theta}) + \delta\Gamma_2'(\overline{\theta}, \underline{\theta})\big] - \big[\Gamma_1'(\underline{\theta}) + \delta\Gamma_2'(\underline{\theta}, \underline{\theta})\big]\Big\} \\ + (1-\nu)\delta\big[\Gamma_2'''(\underline{\theta}) - \Gamma_2'(\overline{\theta}, \underline{\theta})\big] - \nu\big[\tilde{\Phi}'''(\underline{\theta}) - \tilde{\Phi}'(\overline{\theta}, \underline{\theta})\big].\end{aligned}$$

Again, sequential optimality of $q_2'''(\cdot)$ yields a lower bound of zero on the second term: Given that $\nu_2 = \nu$ (note that ν_2 is also independent of the report $\tilde{\theta}_1$ under $\mathcal{S}'''$), $q_2'''(\underline{\theta})$ must satisfy

$$q_2'''(\underline{\theta}) \in \arg\max_{q_2}(1-\nu)\delta\big[S(q_2) - c_2[q_1'(\overline{\theta}), \underline{\theta}]q_2 - \nu\Phi[q_1'(\overline{\theta}), q_2],$$

implying in particular that

$$(1-\nu)\delta\big[\Gamma_2'''(\underline{\theta}) - \Gamma_2'(\overline{\theta}, \underline{\theta})\big] - \nu\big[\hat{\Phi}'''(\underline{\theta}) - \hat{\Phi}'(\overline{\theta}, \underline{\theta})\big] > 0,$$

and hence, using the definition of κ' in (A.21),

$$\Pi(\mathcal{S}''') - \Pi(\mathcal{S}') > -p_1'(\underline{\theta}) \cdot \kappa. \tag{A.24}$$

We have thus shown that, depending on the sign of κ' under $\mathcal{S}'$, either $\mathcal{S}''$ or $\mathcal{S}'''$ (or both) must yield a strictly higher payoff to the principal than $\mathcal{S}'$, which completes the proof.[37]

□

References

Alchian, Armen (1963) 'Reliability of progress curve in airframe production.' *Econometrica* 31(4), 679–93

[37]We note for completeness that, by suitable choice of $p_1''(\underline{\theta})$ for $\mathcal{S}''$(see Footnote 35), the proof in fact shows that any $\mathcal{S}'$ with $p_1'(\overline{\theta}) = 1$, $p_1'(\underline{\theta}) < 1$, and for which the implementability condition (17) is slack is dominated not just by *some* other $\mathcal{S}$, but by some $\mathcal{S}$ for which (17) binds.

Baron, David P., and David Besanko (1984) 'Regulation and information in a continuing relationship.' *Information Economics and Policy* 1(3), 267–302

Baron, David P., and Roger B. Myerson (1982) 'Regulating a monopolist with unknown costs.' *Econometrica* 50(4), 911–930

Becker, Gary S., and Kevin M. Murphy (1988) 'A theory of rational addiction.' *Journal of Political Economy* 96(4), 675–700

Bester, Helmut, and R. Strausz (2001) 'Contracting with imperfect commitment and the revelation principle: The single agent case.' *Econometrica* 69(4), 1077–1098

Boone, Jan, and Joel Shapiro (2005) 'Selling to consumers with endogenous types.' Working paper, Universitat Pompeu Fabra, Barcelona

Edlin, Aaron S., and Chris Shannon (1998) 'Strict monotonicity in comparative statics.' *Journal of Economic Theory* 81(1), 201–219

Fox, J. Ronald (1988) *The Defense Management Challenge: Weapons Acquisition* (Boston, Mass.: Harvard Business School Press)

Freixas, Xavier, and Jean-Jacques Laffont (1990) 'Optimal banking contracts.' In *Essays in Honor of Edmond Malinvaud, Vol. 2, Macroeconomics,* ed. Paul Champsaur (Cambridge, Mass.: MIT Press)

Freixas, Xavier, Roger Guesnerie, and Jean Tirole (1985) 'Planning under incomplete information and the ratchet effect.' *Review of Economic Studies* 52(2), 173–191

Fudenberg, Drew, Bengt Holmstrom, and Paul Milgrom (1990) 'Short-term contracts and long-term agency relationships.' *Journal of Economic Theory* 51(1), 1–31

Gansler, Jacques S. (1989) *Affording Defense* (Cambridge, Mass.: MIT Press)

Gruber, Harald (1998) 'Learning by doing and spillovers: Further evidence for the semiconductor industry.' *Review of Industrial Organization* 13(6), 697–711

Guesnerie, Roger, and Jean-Jacques Laffont (1984) 'A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm.' *Journal of Public Economics* 25(3), 329–369

Hirsch, Werner Z. (1952) 'Manufacturing progress functions.' *Review of Economic Studies* 34(2), 143–55

— (1956) 'Firm progress ratios.' *Econometrica* 24(2), 136–43

Joskow, Paul L., and George A. Rozanski (1979) 'The effects of learning by doing on nuclear power plant operating reliability.' *Review of Economic Studies* 61(2), 161–68

Laffont, Jean-Jacques, and David Martimort (2002) *The Theory of Incentives: The Principal Agent Model* (Princeton, New Jersey: Princeton University Press)

Laffont, Jean-Jacques, and Jean Tirole (1986) 'Using cost observation to regulate firms.' *Journal of Political Economy* 94(4), 614–641

— (1987) 'Comparative statics of the optimal dynamic incentive compatible contract.' *European Economic Review* 31(4), 901–926

— (1988) 'The dynamics of incentive contracts.' *Econometrica* 56(5), 1153–1175

— (1990) 'Adverse selection and renegotiation in procurement.' *Review of Economic Studies* 57(4), 597–562

— (1993) *A Theory of Incentives in Procurement and Regulation* (Cambridge, Mass.: MIT Press)

Lester, Richard K., and Mark J. McCabe (1993) 'The effect of industrial structure on learning by doing in nuclear power plant operation.' *RAND Journal of Economics* 24(3), 418–38

Lewis, Tracy R., and Huseyin Yildirim (2002a) 'Learning by doing and dynamic regulation.' *RAND Journal of Economics* 33(1), 22–36

— (2002b) 'Managing dynamic competition.' *American Economic Review* 92(4), 779–797

Lieberman, Marvin B. (1984) 'The learning curve and pricing in the chemical processing industries.' *RAND Journal of Economics* 15(2), 213–28

Maskin, Eric, and John Riley (1984) 'Monopoly with incomplete information.' *RAND Journal of Economics* 15(2), 171–196

Milgrom, Paul, and John Roberts (1990) 'The economics of modern manufacturing: Technology, strategy and organization.' *American Economic Review* 80(3), 511–528

Mirrlees, James A. (1971) 'An exploration in the theory of optimum income taxation.' *Review of Economic Studies* 38(114), 175–208

Miyazaki, Hajime (1977) 'The rat race and internal labor markets.' *Bell Journal of Economics* 8(2), 394–418

Moul, Charles (2001) 'Evidence of qualitative learning-by-doing from the advent of the 'talkie'.' *Journal of Industrial Economics* 49(1), 97–109

Mussa, Michael, and Sherwin Rosen (1978) 'Monopoly and product quality.' *Journal of Economic Theory* 18(2), 301–317

Nye, William W. (1996) 'Firm-specific learning by doing in semiconductor production: Some evidence from the 1986 trade agreement.' *Review of Industrial Organization* 11(3), 383–94

Preston, Lee E., and E. C. Keachie (1964) 'Cost functions and progress functions: An integration.' *American Economic Review* 54(2), 100–07

Rapping, Leonard (1965) 'Learning and World War II production functions.' *Review of Economic Studies* 47(1), 81–86

Roberts, Peter C., and Calvin C. Burwell (1981) *The Learning Function in Nuclear Reactor Opteration and its Implications for Siting Policy* (Oak Ridge, TN: Institute for Energy Analysis)

Salanié, Bernard (1997) *The Economics of Contracts* (Cambridge, Mass.: MIT Press)

Stobaugh, Robert B., and Patrick L. Townsend (1975) 'Price forecasting and strategic planning: The case of petrochemicals.' *Journal of Marketing Research* 12, 19–29

Sultan, Ralph G. M. (1975) 'Pricing in electrical oligopoly.' Mimeo, Harvard University

Thompson, Peter (2001) 'How much did the liberty shipbuilders learn? New evidence for an old case study.' *Journal of Political Economy* 109(1), 103–37

Thornton, Rebecca A., and Peter Thompson (2001) 'Learning from experience and learning from others: An exploration of learning and spillovers in wartime shipbuilding.' *American Economic Review* 91(5), 1350–68

Topkis, Donald M. (1998) *Supermodularity and Complementarity* (Princeton, New Jersey: Princeton University Press)

Wright, T. P. (1936) 'Factors affecting the cost of airplanes.' *Journal of Aeronautical Sciences* 2, 122–28

Chapter 2

Ex-Post Implementable Merger Mechanisms

Dennis Gärtner and Armin Schmutzler

1 Introduction

There is a widespread perception that merger decisions are not necessarily efficient. The focus of the literature has been on the frequent occurrence of unhappy marriages, where at least one of the involved firms regrets the decision with the benefit of hindsight.[1] To explain these events, the literature has invoked agency arguments, emphasizing the incentives of managers to expand their empire (cf. Jensen, 1986; Shleifer and Vishny, 1988). The present paper offers an alternative route towards understanding the intrinsic problems of merger decisions by examining the conflict of interests between firms rather than agency conflicts within firms. We argue that, absent any misalignment of incentives *within* either firm, the misalignment of incentives *between* firms can still present a fundamental impediment to the implementation of efficient merger decisions. We show that it is usually impossible to simultaneously prevent unhappy marriages and guarantee that mergers which should take place are actually carried out. In this sense, the paper relates to the theorem of Myerson and Satterthwaite (1983) which shows the impossibility of efficient bilateral trade of indivisible goods under private valuations.

We consider a general mechanism design approach in a setting where two merger candidates each possess private information. Each party's information pertains to its own profits as well as the other's if no merger occurs, and it is also relevant to

[1] See, for instance, Ravenscraft and Scherer (1989), Agrawal et al. (1992), and Gugler et al. (2003). For a dissenting view, see Healy et al. (1992).

the profits of the merged entity in the event that a merger occurs. For instance, information may concern a firm's productivity. Typically, both a firm's stand-alone profits and the post-merger profits of an entity to which it belongs should be higher for more productive types. Moreover, if the firms compete in the same market, each firm's productivity will usually have a negative impact on the competitor's stand-alone profits. As another example, information could represent a firm characteristic that is only relevant to post-merger profits. For instance, one might think of private information as some measure of organizational culture such that higher merger profits require cultures to be more similar.

In this environment, we start by considering merger mechanisms which use a combination of share allocations and transfers to elicit agents' private information (i.e., their type) and decide on the merger. More specifically, a *direct merger mechanism* consists of

(i) a *merger decision* for each combination of type reports,
(ii) a *share function* allocating to each agent a report-dependent share in the merged entity,
(iii) a *transfer function* which demands a transfer from each agent, again dependent on type reports.

We furthermore restrict attention to mechanisms which, in equilibrium, give neither party any reason for *ex-post* regret.

Under a natural budget balance condition, we obtain the following results. First, we show that a fundamental problem of any such mechanism is that shares and transfers must essentially be invariant to agents' reports, thus limiting their usefulness for incentive design. Second and consequently, we show that except for special environments, only trivial merger decisions of the type 'always merge' or 'never merge' are implementable. Third, the additional requirement of voluntary participation in the mechanism further reduces the set of feasible mechanisms to ones which merely reproduce the status quo. Fourth, we show that, for a wide class of merger environments, the essential impossibility of implementing mergers without *ex-post* regret carries over to more general mechanisms where each partner's payoff can be conditioned on the merged firm's total profit in an arbitrary fashion. Finally, we show that our main conclusions crucially rely on the simultaneous presence of asymmetric information on *both* parties' outside options, as any remaining environments permit implementation of the efficient merger decision in a regret-free manner—in spite of asymmetric information on merger profits.

Classical mechanism design literature has long dealt with the issue of whether and how mechanisms can implement efficient decisions when information concerning the benefits and costs constitute individuals' private knowledge (see e.g. Milgrom, 2004, for an overview). A major finding has been that in quite general settings, Vickrey-Clarke-Groves (VCG) mechanisms solve this problem.[2]

[2]In fact, it has further been shown that VCG-mechanisms represent the only class of mechanisms

For two reasons, however, this result is not directly applicable to our setting. First, standard VCG mechanisms fail to implement the efficient decision when valuations are interdependent (see, e.g., Mezzetti, 2004).[3] Interdependent valuations constitute a core feature of our model, however, since private information held by the other—on own stand-alone profits and joint profits in the event of a merger—makes each agent's payoff naturally dependent on the other's information. Though there is work suggesting that in general, there are limits to implementation when valuations are interdependent,[4] encouraging findings have been made for specific frameworks, *auctions* in particular, by designing *extended* VCG mechanism which implement the efficient decision in spite of interdependent valuations.[5]

Second, in the private-valuations framework, VCG mechanisms are usually not budget balanced (cf. Holmstrom, 1977; Green and Laffont, 1979), a key requirement of merger mechanisms. Again, a generalization of this result to interdependent valuations is not available, and analogies from the aforementioned auction literature cannot be drawn, as budget balance is not an issue in auction contexts due to the auctioneer being a natural third-party residual claimant.[6]

In sum, therefore, it is not possible to directly relate our main finding that only trivial decision rules are implementable to the previous mechanism-design literature. However, there is an important feature of merger mechanisms that nonetheless makes this particularly strong result somewhat surprising: the availability of shares in addition to transfers for incentive design. Like transfers, shares provide a means of transferring utility between agents without affecting joint surplus and thereby the efficiency of the decision. In principle, one should expect this additional degree of freedom in incentive design to improve the prospects of efficient implementation.[7] As

to implement the efficient decision in dominant strategies (Green and Laffont, 1977, 1979; Walker, 1980).

[3]In addition to providing a nice account of why VCG mechanisms will fail, Mezzetti (2004) also shows how the problem can be overcome by by extending the standard mechanism by a further stage *after* the decision has been implemented, in which parties report their resulting gross decision payoffs and transfers can be made contingent on these reports.

[4] Jehiel et al. (2005) show in a general interdependent-values framework that, except for special cases, only trivial choice rules are implementable when private information is *multi-dimensional*.

[5]Crémer and McLean (1985), Maskin (1992), Dasgupta and Maskin (2000), Eső and Maskin (2000), Jehiel and Moldovanu (2001), Bergemann and Välimäki (2002), Perry and Reny (2002), Krishna (2003), and Ausubel (2004) all construct efficient, *ex-post* efficient mechanisms in settings with interdependent valuations.

[6]However, Chung and Ely (2003) provide largely negative results concerning implementation of *efficient* decisions both in the context of bilateral trade and the provision of a public good, each with interdependent valuations.

[7]Brusco et al. (2005) make a similar point. In the context of private-value auctions, this additional freedom in incentive design by auctioning discrete *shares* in a good has already been noted (see Hansen, 1985, and the comments in Crémer, 1987, and Samuelson, 1987). However, the focus has been on the use of such 'contingent payments' on increased surplus extraction by the auctioneer, rather than its impact on the set of implementable decisions—which is the focus of this paper.

a key step in our analysis shows, however, this intuition is largely invalid, as budget balance keeps tight limits on how shares and transfers can respond to information revealed.

A second way in which our results should appear particularly stark is that, not only do we show the general inability to implement the *efficient* decision rule, but we in fact exclude *any nontrivial* decision rule.[8] We establish a rationale for this in a separate building block of our analysis, by showing that budget balance produces a tight connection between implementability and efficiency in the sense that not much more than efficient decisions can be expected to be implementable in the first place.

Finally, an obvious fundamental ingredient to our result is the requirement of *ex-post* incentive compatibility instead of weaker notions such as Bayesian-Nash implementation. Our motivation for this is manyfold. First, a convenient feature of *ex-post* implementable mechanisms is their independence of priors on the type distribution, making the analysis more tractable and results more transparent. In arguments related to this property, the recent literature has furthermore put forward interesting theoretical justifications for this approach by investigating the robustness of Bayesian Mechanisms with respect to beliefs about the type distribution when valuations are interdependent.[9] More importantly, however, our use of *ex-post* implementability is less a hypothesis about participants' behavior than it is a restriction on the mechanism itself, imposed with the sole purpose of answering this paper's main question: whether *ex-post* regret is an inevitable feature of merger decisions or not. That is, even if we believe that firms entering merger deals optimize in a Bayesian sense, it is interesting to investigate the class of 'regret-free' mechanisms. Finally, even though the *ex-post* concept is strong, it is not so strong as to have prohibited efficient implementation in other settings with interdependent values, such as in the aforementioned applications to auction theory.

The paper most closely related to ours is Brusco et al. (2005). These authors use a mechanism design approach to analyze under which circumstances efficient merger decisions might come about in the presence of two-sided asymmetric information. Their setting differs considerably from ours, however. On the one hand, the mechanisms they consider are more general than ours. First, their mechanisms not only determine whether a merger takes place, but also which two out of several firms merge. Second, Brusco et al. (2005) consider Bayesian mechanisms. On the other

[8]This contrasts with results in settings with *private* valuations, where a fruitful way to address the problem that VCG-mechanisms are typically not budget balanced has been to consider the class of *second-best*, or *incentive efficient* mechanisms. Unfortunately, most investigations of *interdependent*-values frameworks so far have not gone beyond showing the ability or inability of implementing the efficient decision. A noteworthy except is Jehiel et al. (2005), who, like our paper, characterize the *full* set of implementable decisions in the context of *multi-dimensional* information.

[9]See Bergemann and Morris (2005) and Chung and Ely (2003) in particular, as well as the nice survey given in Jehiel et al. (2005).

hand, the merger environments they consider are in some ways more restrictive. For instance, the private information of each party on its stand-alone productivity is independent of its private information on synergies, so there is no sense in which inherently 'better' firms can also be better merger partners. Also, each potential partner's private information is uninformative about the other firm's stand-alone profit. This is problematic because whenever firms are engaged in some form of market interaction, a firm's private signal on own profitability will translate into a signal on the other firm's profitability.

Interestingly, there has recently been quite some investigation into the efficient *dissolution* of partnerships under interdependent valuations, which in some ways represents the inverse to the problem considered in this paper. This literature works out conditions under which it is possible to efficiently dissolve joint ownership of a firm when valuations for that firm are interdependent.[10] In contrast to our paper, whether it is efficient to dissolve or not is generally not an issue (the social value of the partnership is simply the share-weighted average of agents' individual valuations, so dissolution will be efficient whenever agents' valuations differ). Rather, the question is only one of allocating the single indivisible item owned by the partnership to the agent with the highest valuation. Thus, the problem is strongly related to auction design, with some additional difficulties implied by budget-balance requirements. In our setting, on the other hand, the main question is whether the partnership should be formed or not. In contrast, how shares in the partnership are to be allocated in the event of a merger does not affect efficiency, as agents have pure common valuations concerning the partnership itself.

There is also a small number of papers which address the issue of mergers under two-sided information by restricting attention to more specific mechanisms. Like this paper, Borek et al. (2004) consider an environment where firm types affect pre- and post-merger profits. In situations where the gains or losses from a merger are split according to some simple predetermined scheme,[11] they investigate circumstances under which both firms are willing to consent to a merger in a Bayesian game. The equilibrium is typically associated with *ex-post* regret. Banal-Estañol and Seldeslachts (2005) also consider such a merger consent game; however (i) asymmetric information concerns only the potential post-merger profits and (ii) additional complexity arises from the idea that both partners may or may not choose costly 'cultural adjustment' decisions after the merger.

The paper is organized as follows. Section 2 introduces the model with share-transfer mechanisms. Section 3 derives restrictions on share and transfer functions stemming exclusively from incentive compatibility. Section 4 works out additional

[10] See Fieseler et al. (2003), Jehiel and Pauzner (2003), Ornelas and Turner (2004) and the survey in Moldovanu (2002).

[11] For instance, each firm might earn a fixed share of profits under a merger, or one firm is up for sale at some fixed price.

restrictions which derive from budget balance and presents our main results. Section 5 analyzes additional constraints imposed by requiring voluntary participation in the mechanism. Section 6 discusses generalized mechanisms that allow each partner's payoff to depend on the merged firm's profit in a more general manner. Section 7 illustrates how the results developed apply to more specific classes of merger environments. Finally, Section 8 concludes.

2 The Model

This section presents the basis for our analysis of implementable merger mechanisms. Section 2.1 introduces the general setup and basic terminology. We motivate our framework with specific examples in Section 2.2.

2.1 Merger Mechanisms

We consider merger mechanisms in an environment of the following type:

Definition 2.1. A *merger environment* $\mathcal{E}$ is a tuple $(T_1, T_2, \pi_1, \pi_2, \pi^{\mathrm{M}})$ with the following components:

(i) $T_i = [0,1]$, $i = 1, 2$, is the *type space* for firm i;

(ii) $\pi_i : T_i \times T_j \to \mathbb{R}_{\geqslant 0}$ is the *stand-alone profit function* for firm i;

(iii) $\pi^{\mathrm{M}} : T_1 \times T_2 \to \mathbb{R}_{\geqslant 0}$ is the *merger profit function.*

Thus, an environment describes how each firm's stand-alone profits as well as joint profits in the event of a merger depend on firms' private information. We should point out that the dependence of firms' stand-alone profits on other firms' information is not crucial to our results, but represents a natural feature in many conceivable applications (see Section 2.2 below).[12] For expositional simplicity, unless stated otherwise, we will assume that π_i and π^{M} are twice continuously differentiable. Moreover, note that the above definition assumes that each agent's private information, denoted by $t_i \in T_i$, is univariate.[13]

For such an environment, we investigate mechanisms which use a system of transfers and—provided that a merger occurs—share allocations in the merged entity to implement a merger decision. Anticipating the usual revelation argument, we restrict our attention to *direct* merger mechanisms:

[12] This feature is absent in Brusco et al. (2005).

[13] Given the largely negative results concerning efficient implementation under interdependent valuations with *multi-dimensional* as opposed to *univariate* information in auction settings (Jehiel and Moldovanu, 2001) and in more general settings (Jehiel et al., 2005), we assume scalar-valued information to improve the scope for efficient implementation in our merger setting.

Definition 2.2. A *(direct) merger mechanism* $\mathcal{M}$ is a tuple (m, s, p) consisting of a merger decision function m, share functions $s = (s_1, s_2)$, and transfer functions $p = (p_1, p_2)$, which are defined as follows:

(i) The *merger decision function* $m : T_1 \times T_2 \to \{0, 1\}$ maps a combination of reports $(\tilde{t}_1, \tilde{t}_2)$ by players 1 and 2 about their type into a merger decision, with $m(\tilde{t}_1, \tilde{t}_2) = 1$ if and only if the merger takes place as a result of the reports. We let M^1 denote the subset of type space $T_1 \times T_2$ for which mergers are implemented, and let M^0 denote the subset for which mergers are not implemented.

(ii) The *share functions* $s_i : M^1 \to [0, 1]$, $i \in \{1, 2\}$, map a combination of reports to a share in the merged entity's profit.

(iii) The *transfer functions* $p_i : T_1 \times T_2 \to \mathbb{R}$, $i \in \{1, 2\}$, assign a transfer payment from player i to the mechanism operator as a function of reported types.

Given any merger environment $\mathcal{E}$ and any merger mechanism $\mathcal{M}$, we let $u_i : (T_i \times T_j)^2 \to \mathbb{R}$ denote the *individual payoff functions*, which assign profits

$$\begin{aligned} u_i(\tilde{t}_i, \tilde{t}_j; t_i, t_j) \\ &= [1 - m(\tilde{t}_1, \tilde{t}_2)] \cdot \pi_i(t_i, t_j) + m(\tilde{t}_1, \tilde{t}_2) \cdot s_i(\tilde{t}_i, \tilde{t}_j) \cdot \pi^{\mathrm{M}}(t_1, t_2) - p_i(\tilde{t}_i, \tilde{t}_j) \end{aligned} \quad (1)$$

to player i as a function of reported and true types.

For reasons given in the introduction, we shall confine ourselves to finding *ex-post* implementable mechanisms, that is, mechanisms for which truthful reporting is a best response to any given type of the competitor, *given* truthful reporting by the other type. Formally, letting $U_i(\tilde{t}_i; t_i, t_j) = u_i(\tilde{t}_i, t_j; t_i, t_j)$ denote firm i's payoff given truthful reporting by the other firm, the requirement of *ex-post* incentive compatibility can be compactly formulated as follows:

Definition 2.3. A merger mechanism $\mathcal{M}$ is *(ex-post) incentive compatible (IC)* if

$$U_i(t_i; t_i, t_j) \geqslant U_i(\tilde{t}_i; t_i, t_j) \qquad \text{for all } \tilde{t}_i, t_i \in T_i,\ t_j \in T_j. \quad (2)$$

We shall call a merger decision function m *implementable* if there exist share and transfer functions s and p such that the mechanism (m, s, p) is incentive compatible.

Efficiency of a merger decision function is defined as follows:

Definition 2.4. A merger decision function m is *efficient* if

$$m(t_1, t_2) = \begin{cases} 0 & \text{for } (t_1, t_2) \text{ s.t. } \pi^{\mathrm{M}} < \pi_1 + \pi_2, \\ 1 & \text{for } (t_1, t_2) \text{ s.t. } \pi^{\mathrm{M}} > \pi_1 + \pi_2. \end{cases} \quad (3)$$

A merger mechanism $\mathcal{M} = (m, s, p)$ is efficient if and only if m is efficient. Moreover, we shall call a merger decision m *trivial* if it is constant over the entire type space $T_1 \times T_2$.

Since it is natural to assume that no third party should benefit from or pay for the mechanism, we introduce the following condition:[14]

Definition 2.5. A merger mechanisms satisfies *budget balance (BB)* if

$$\begin{aligned} &1.\ p_1(\tilde{t}_1, \tilde{t}_2) + p_2(\tilde{t}_2, \tilde{t}_1) = 0 \text{ for all } (\tilde{t}_1, \tilde{t}_2) \in T_1 \times T_2, \text{ and} \\ &2.\ s_1(\tilde{t}_1, \tilde{t}_2) + s_2(\tilde{t}_2, \tilde{t}_1) = 1 \text{ for all } (\tilde{t}_1, \tilde{t}_2) \in M^1. \end{aligned} \tag{4}$$

(BB) is equivalent to requiring (i) that agents' utilities $u_1 + u_2$ sum to realized profits (π^{M} if a merger occurs and $\pi_1 + \pi_2$ if it does not), or (ii) that the surplus generated by the mechanism, $m \cdot (1 - s_1 - s_2) \cdot \pi^{\mathrm{M}} + p_1 + p_2$, is zero, with each requirement applying for any combination of types and reports.[15]

Finally, mechanisms in which agents will voluntarily participate must satisfy the following requirement:

Definition 2.6. A merger mechanism satisfies *individual rationality (IR)* if, for any $i \in \{1, 2\}$,

$$U_i(t_i; t_i, t_j) \geqslant \pi_i(t_i, t_j) \qquad \text{for all } (t_i, t_j) \in T_i \times T_j. \tag{5}$$

Note that there is one particular mechanism which—albeit typically being inefficient—satisfies all of the above constraints, namely the mechanism which reproduces the status quo by prescribing (i) never merge, and (ii) always require zero payments. Such a mechanism satisfies (IC), (BB), and (IR).

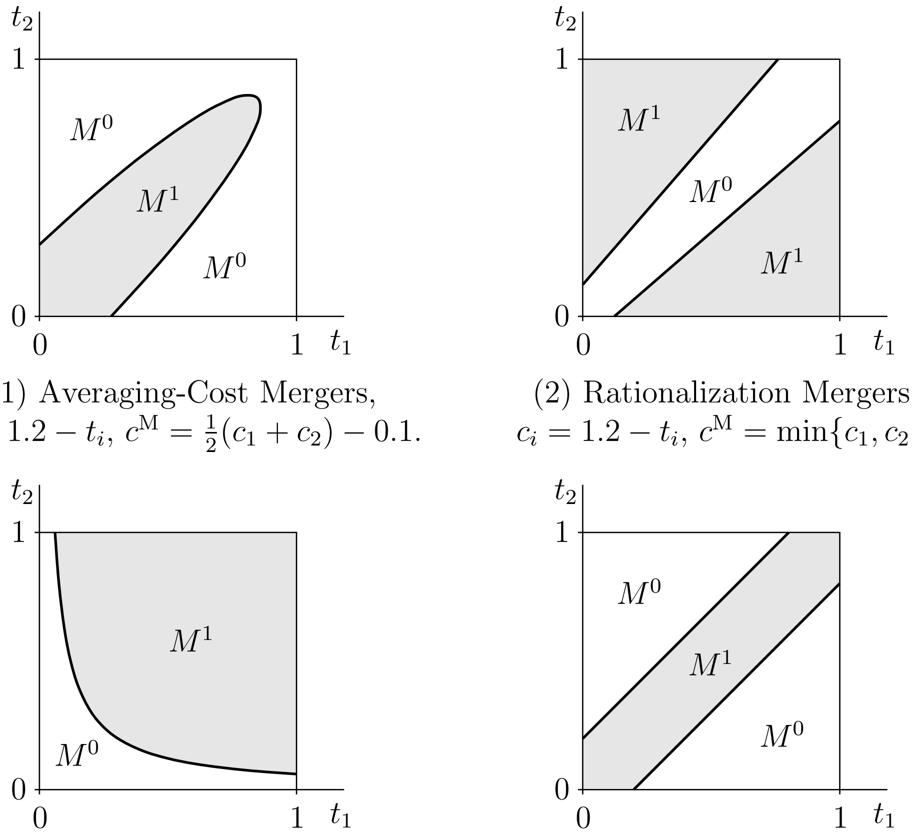
2.2 Examples of Merger Environments

We illustrate our set-up with a few exemplary merger environments. In all these examples, we consider two firms $i = 1, 2$ in a Cournot oligopoly with $n \geqslant 3$ firms facing linear demand, where each firm has constant marginal costs c_i of production. If firms 1 and 2 decide to merge, the new entity has marginal costs of c^{M} in the resulting $(n-1)$-firm oligopoly. The examples differ in their interpretation of private information t_1, t_2 and in the way that this information affects stand-alone costs c_1, c_2 and the merged entity's costs c^{M}.

Example 1 (Averaging-Cost Mergers). Assume $c_i = \overline{c} - t_i$ for $i = 1, 2$, $\overline{c} \in \mathbb{R}_{>0}$, and $c^{\mathrm{M}} = \frac{1}{2}(c_1 + c_2) - \gamma$, $\gamma \in \mathbb{R}_{>0}$. Thus, stand-alone costs are private information, and the merged entity's costs are determined by the mean of the constituent firms'

[14] Such a condition has typically been invoked also in the aforementioned literature on partnership dissolution.

[15] It is quickly seen that (BB) implies (ii) and that statements (i) and (ii) are equivalent. Moreover, it is easily checked that (ii) implies (BB) whenever merger profits π^{M} are not globally constant in types, which establishes the claim for non-constant merger profits. Equivalence does not hold whenever π^{M} is constant over $T_1 \times T_2$, but in this case, there is in fact no point in distinguishing shares from payments in our model.



(1) Averaging-Cost Mergers, $c_i = 1.2 - t_i$, $c^{\mathrm{M}} = \frac{1}{2}(c_1 + c_2) - 0.1$.

(2) Rationalization Mergers, $c_i = 1.2 - t_i$, $c^{\mathrm{M}} = \min\{c_1, c_2\}$.

(3) Synergies, $c_i = 1$, $c^{\mathrm{M}} = 1 - t_1 \cdot t_2$.

(4) Matching Corporate Culture, $c_i = 1$, $c^{\mathrm{M}} = 0.9 + (t_1 - t_2)^2$.

Figure 1: Efficient Merger Decisions for Examples of Section 2.2.

stand-alone costs, minus some known synergy effect.[16] In this case, $\frac{\partial}{\partial t_i}\pi_i > 0$ and $\frac{\partial}{\partial t_j}\pi_i < 0$, as each firm's stand-alone profit is decreasing in its own marginal costs and increasing in those of the potential merger partner. Furthermore, $\frac{\partial}{\partial t_i}\pi^{\mathrm{M}} > 0$ for each $i = 1, 2$ since the merged firm's marginal costs increase in the constituent parties' marginal costs. The top-left panel in Figure 1 illustrates the efficient merger decision for specific parameter values.[17] Observe that merging is only profitable if firms' pre-merger costs are sufficiently similar and sufficiently high.

Example 2 (Rationalization Mergers). As above, assume that $c_i = \overline{c} - t_i$ for $i = 1, 2$, $\overline{c} \in \mathbb{R}_{>0}$, but now let $c^{\mathrm{M}} = \min\{c_1, c_2\}$. Thus, the merger allows for rationalization in the sense that the merged entity fully implements the more efficient

[16] Note that without such a synergy effect, merging will never be profitable under linear demand (see Salant et al., 1983).

[17] All illustrations in Figure 1 assume demand to be given by $P(Q) = 3 - Q$, that the market is (originally) served by $n = 3$ firms, and that marginal costs of the third firm are fixed at $c_3 = 1$. All remaining parameters are noted in the corresponding captions.

party's technology. Contrary to Example 1, π^{M} will only be *strictly* decreasing in an agent's type if that agent is more efficient, whereas a change in the less efficient type will have no impact on merger profits, so $\frac{\partial}{\partial t_i}\pi^{\mathrm{M}} = 0$ for $t_i > t_j$. Figure 1 illustrates the efficient merger decision. Note that there are *two* separate regions where merging is profitable, each one where types are sufficiently dissimilar.[18]

Example 3 (Synergies). Assume now that $c_1 = c_2 = c$ for some $c \in \mathbb{R} > 0$ and that $c^{\mathrm{M}} = c - t_1 \cdot t_2$. This example may be thought of as a situation in which there are synergies to be achieved by merging and each party possesses private information concerning its synergy potential, where this potential in turn is (i) unrelated to stand-alone profits, but (ii) complementary to the potential merger partners' synergy potential. The resulting efficient merger decision is illustrated in Figure 1.

Example 4 (Matching Corporate Culture). Many case studies attribute failed mergers to incompatible corporate cultures of the merging firms.[19] To represent such a problem in our model, we let $c_1 = c_2 = c$ for some $c \in \mathbb{R} > 0$, but we now let $c^{\mathrm{M}} = c - \gamma + (t_1 - t_2)^2$ and interpret t_i as some measure of firm i's corporate culture. Then the merged firm's profits π^{M} increase with the similarity of the merged firms' corporate cultures. Moreover, if different corporate cultures are not *per se* better or worse, each firms' stand-alone profits will be independent of its own cultural measure and that of the other firm, which justifies the type-independent c_i's. The corresponding plot is shown in Figure 1.

3 General Incentive Compatibility Results

This section's goal is to present a set of restrictions on permissible share and transfer functions which are implied by the incentive compatibility requirement alone. We will largely confine ourselves to restrictions which derive from the condition that *marginal* deviations from truthful reporting be unprofitable. In Section 3.1, we deal with deviations that do not affect the merger decision; in Section 3.2 we address deviations that do. These results will provide the basis for Section 4, where the further incorporation of the budget-balance requirement (BB) will lead to our main results, Propositions 4.3 and 4.9. Proposition 4.3 will show that essentially, only efficient merger decision functions are implementable. Proposition 4.9 will argue that on the boundary between M^0 and M^1, profit functions must satisfy a geometric condition described as 'proportional variation'. These two results are then combined to argue that generally, only trivial merger decisions are implementable.

[18] See Barros (1998).

[19] See Larsson and Finkelstein (1999).

For expository simplicity, it will often be useful to restrict attention to certain well-behaved classes of mechanisms. Denoting the restrictions of p_i to M^0 and M^1 by p_i^0 and p_i^1, respectively, we define a *differentiable* mechanism as follows:

Definition 3.1. A mechanism $\mathcal{M} = (m, s, p)$ is *differentiable* if for each $i = 1, 2$, the functions s_i, p_i^0 and p_i^1 are each differentiable over the interior of their respective domains of definition.

The alternative requirements of a mechanism being *continuous*, *twice differentiable*, etc. are to be read analogously. It should be pointed out that none of these requirements place any restrictions on the original transfer function p_i at points on the boundary between M^0 and M^1.

3.1 Deviations without Effect on the Merger Decision

We first consider deviations which do not affect the merger decision and as such can be profitable only by leading to a decrease in transfers or an increase in shares (or both). As the next proposition shows, the need for a direct mechanism to prohibit such deviations leads to strong restrictions on how transfers and shares may vary over M^0 and M^1, respectively, given any merger decision function m.

Proposition 3.2. *Any differentiable merger mechanism $\mathcal{M} = (m, s, p)$ must be such that*

(a) for any $(t_i', t_j), (t_i'', t_j) \in M^0$,

$$p_i^0(t_i', t_j) = p_i^0(t_i'', t_j),$$

(b) for any $(t_1, t_2) \in M^1$,

$$\pi^{\mathrm{M}}(t_1, t_2) \cdot \tfrac{\partial}{\partial \tilde{t}_i} s_i(t_i, t_j) = \tfrac{\partial}{\partial \tilde{t}_i} p_i^1(t_i, t_j) \tag{6}$$

$$\tfrac{\partial}{\partial t_i} \pi^{\mathrm{M}}(t_1, t_2) \cdot \tfrac{\partial}{\partial \tilde{t}_i} s_i(t_i, t_j) \geqslant 0. \tag{7}$$

See the Appendix for the proof.

By Proposition 3.2(a), given the other firm's type t_j, any firm i's transfer must be the same for any report which does not lead to a merger. The intuition for this result is obvious, as any deviation from truthful reporting over M^0 will affect payoffs only through transfers, so that transfers being non-constant in own report over M^0 would inevitably provoke false reports by certain types.

Over M^1, on the other hand, the mechanism has two instruments at its disposal, giving it the freedom to let transfers and shares respond to changes in reported types. As Proposition 3.2(b) shows, incentive compatibility puts some limits on this freedom, however: By condition (6), the response of shares and transfers to own reports are linked in that any increase in shares must be matched by a concomitant increase in transfers, the factor of proportionality given by $\pi^{\mathrm{M}}(t_1, t_2)$. This condition simply arises from the requirement that marginal deviations from truth-telling be unprofitable. Additionally, condition (7) restricts the *direction* in which shares may respond to reports by requiring higher merger profits due to a marginal change in an agent's type to be complemented by (weakly) higher shares in the merged entity for that agent.

Condition (7) is best understood as an application of the 'Positive Association of Differences' (PAD) condition in Roberts (1979): If a change in agent i's signal from t_i to t_i' makes some alternative A relatively more preferable for agent i than some other alternative B (gross of transfers), then an incentive compatible mechanism must respect these preferences by not choosing A at signal t_i and B at signal t_i'.[20] In our case, a change in agent i's signal t_i which increases merger profits π^{M} will obviously make a larger share s_i more preferable to agent i, and condition (7) says that an implementable mechanism must respect this by not allocating the agent a smaller share in response to this change.[21]

3.2 Deviations which Affect the Merger Decision

Having derived important restrictions on how shares and transfers may vary over M^0 and M^1, we now derive related conditions for the *boundary* between M^0 and M^1, that is, at type profiles where the merger decision m is discontinuous in at least one agent's report. We let $\overline{M^0}$ and $\overline{M^1}$ denote the closure of M^0 and M^1, respectively, so that the relevant boundary is given by $\overline{M^0} \cap \overline{M^1}$. Further, it is useful to introduce the following notation:

Definition 3.3. For any differentiable mechanism, let $\overline{p}_i^0 : \overline{M^0} \to \mathbb{R}$ denote the continuous extension of p_i onto $\overline{M^0}$, and let $\overline{p}_i^1 : \overline{M^1} \to \mathbb{R}$ and $\overline{s}_i : \overline{M^1} \to [0,1]$ denote the continuous extensions of p_i^1 and s_i, respectively, onto $\overline{M^1}$, where the continuous extensions are all taken with respect to $\tilde{t}_i$.[22]

[20] For readers not familiar with this result, it may be useful to note that in the standard model of procurement under asymmetric information, where a principal buys a certain quantity of a good produced by an agent possessing private information on his marginal costs of producing this good, this condition produces the familiar requirement that the principal cannot procure a strictly larger quantity of the good from an agent with higher costs than from one with lower costs.

[21] As shown in the proof in the Appendix, there is a more general, global version of condition (7) which asserts that for any $t_i', t_i'' \in T_i$ and $t_j \in T_j$ such that $(t_i', t_j), (t_i'', t_j) \in M^1$, the share function satisfy $[s_i(t_i'', t_j) - s_i(t_i', t_j)] \cdot [\pi^{\mathrm{M}}(t_i'', t_j) - \pi^{\mathrm{M}}(t_i', t_j)] \geqslant 0$. However, this will be largely irrelevant to our results.

[22] Formally, $\overline{p}_i^0(\tilde{t}_i, \tilde{t}_j) = p_i^0(\tilde{t}_i, \tilde{t}_j)$ for any $(t_i, t_j) \in M^0$, $\overline{p}_i^1(\tilde{t}_i, \tilde{t}_j) = p_i^1(\tilde{t}_i, \tilde{t}_j)$, and $\overline{s}_i(\tilde{t}_i, \tilde{t}_j) = s_i(\tilde{t}_i, \tilde{t}_j)$ for any $(\tilde{t}_i, \tilde{t}_j) \in M^1$, whereas for any $(\tilde{t}_i, \tilde{t}_j) \in \overline{M^0} \cap \overline{M^1}$,

$$\overline{p}_i^0(\tilde{t}_i, \tilde{t}_j) = \lim_{\{(t_i', \tilde{t}_j)\}_{\in M^0} \to (\tilde{t}_i, \tilde{t}_j)} p_i^0(\tilde{t}_i', \tilde{t}_j),$$

$$\overline{p}_i^1(\tilde{t}_i, \tilde{t}_j) = \lim_{\{(\tilde{t}_i', \tilde{t}_j)\}_{\in M^1} \to (\tilde{t}_i, \tilde{t}_j)} p_i^1(\tilde{t}_i', \tilde{t}_j), \qquad \overline{s}_i^1(\tilde{t}_i, \tilde{t}_j) = \lim_{\{(\tilde{t}_i', \tilde{t}_j)\}_{\in M^1} \to (\tilde{t}_i, \tilde{t}_j)} s_i^1(\tilde{t}_i', \tilde{t}_j).$$

Using this notation, any mechanism must satisfy the following conditions at points where the merger decision is locally sensitive to an agent's report:

Proposition 3.4. *Any differentiable incentive compatible mechanism* $\mathcal{M} = (m, s, p)$ *must be such that, at any* $\mathbf{t} = (t_1, t_2) \in \overline{M^0} \cap \overline{M^1}$ *where agent* i *can change the merger decision* m *by a marginal deviation from truth-telling,*

(a) $\overline{s}_i(\mathbf{t}) \cdot \pi^{\mathrm{M}}(\mathbf{t}) - \pi_i(\mathbf{t}) = \overline{p}_i^1(\mathbf{t}) - \overline{p}_i^0(\mathbf{t})$,

(b) $\overline{s}_i(\tilde{\mathbf{t}}) \cdot \pi^{\mathrm{M}}(\mathbf{t}) - \pi_i(\mathbf{t})$ *must be weakly increasing in* t_i *locally at* $\tilde{\mathbf{t}} = \mathbf{t}$ *whenever* m *is locally increasing in* $\tilde{t}_i$*, and weakly decreasing in* t_i *if* m *is locally decreasing in* $\tilde{t}_i$*.*

See the Appendix for the proof.

Proposition 3.4(a) states that shares and transfers must be designed so that types (t_1, t_2) located at the boundary of a merger region (i.e., types for which a marginal deviation from truthful reporting will change the merger decision) will be indifferent between their payoff under a merger and under no merger. Put differently, whenever an agent can change the merger decision with a marginal deviation from truth-telling, any resulting *gross benefits to merging*, $\overline{s}_i(\mathbf{t}) \cdot \pi^{\mathrm{M}}(\mathbf{t}) - \pi_i(\mathbf{t})$ must be matched by a corresponding increase in transfers, $\overline{p}_i^1(\mathbf{t}) - \overline{p}_i^0(\mathbf{t})$.

Moreover, Proposition 3.4(b) shows that at any such point on the boundary, agents' relative valuations of these gross benefits to merging must be aligned with the preferences implicit in the merger decision. Like the sign-condition on shares in Proposition 3.2(b), this result can be interpreted as an implication of the PAD-property, this time concerning the merger decision m rather than the share function s. To see this, note that gross of transfers, payoffs to agent i will be $s_i \cdot \pi^{\mathrm{M}}$ if a merger is executed, and π_i if it is not. Hence, if $s_i \cdot \frac{\partial}{\partial t_i}\pi^{\mathrm{M}} - \frac{\partial}{\partial t_i}\pi_i \geqslant 0$, a merger will become relatively more preferable to agent i with increasing t_i, and the PAD-property requires the decision function m to respect this by not implementing a merger for low t_i and no merger for higher t_i (locally). The converse holds if $s_i \cdot \frac{\partial}{\partial t_i}\pi^{\mathrm{M}} - \frac{\partial}{\partial t_i}\pi_i \leqslant 0$.

4 Imposing Budget Balance

Our analysis thus far has provided several necessary conditions on transfer and share functions which must hold for any incentive-compatible merger decision function. This section investigates additional restrictions imposed by budget balance.

To repeat, our goal is not just to check for implementability of *efficient* merger decisions, but to describe the *full set* of implementable decisions. In a first step,

These extensions will always be well-defined and exist since p_i^0 is bounded w.r.t. $\tilde{t}_i$ over M^0 by Proposition 3.2(a), and s_i is bounded between 0 and 1 over M^1 by assumption, which in turn implies that p_i^1 is bounded w.r.t. $\tilde{t}_i$ by condition (6) in Proposition 3.2(b).

Section 4.1 establishes a strong link between efficiency and implementability under budget balance. We show essentially that type profiles such that a decision rule m is sensitive to both agents' information must lie on the boundary of the set where merging is efficient. In a second step, Section 4.2 shows that under budget balance, incentive compatible share and transfer schemes must largely ignore reports. Next, Section 4.3 combines this latter result with previous findings from Section 3 to show that each agent's gross benefits to merging, $s_i \cdot \pi^{\mathrm{M}} - \pi_i$, must be constant along lines separating M^0 from M^1, with s_i also constant along each such line. These results are brought together in Section 4.4, which leads to our central result that in general, only trivial merger decisions will be implementable.

4.1 The Link between Efficiency and Implementability

We now show that budget balance severely limits the design of incentive-compatible mechanisms, making it impossible to implement certain merger decisions m which (i) are responsive to both agents' reports and (ii) do not satisfy efficiency. To allow for more compact formulation of our results, we introduce the following notation:

Definition 4.1. For any merger decision function m and any agent $i \in \{1, 2\}$, let $I_i(m)$ denote the set of all interior type combinations $(t_1, t_2) \in (T_1 \times T_2)^\circ \equiv T^\circ$ such that m is discontinuous in t_i at (t_1, t_2). Furthermore, let $I(m) \equiv I_1(m) \cup I_2(m)$, and let $\tilde{I}(m) \equiv I_1(m) \cap I_2(m)$.[23]

Definition 4.2. Let $I^* \equiv \{(t_1, t_2) \in T^\circ : \pi^{\mathrm{M}} = \pi_1(t_1, t_2) + \pi_2(t_2, t_1)\}$ denote the set of all interior type combinations such that merging and not merging yield the same joint profits.

Graphically, the set $I(m)$ collects the boundaries between the merger- and no-merger regions M^1 and M^0 in type space $T_1 \times T_2$. The subsets I_1, I_2 denote the sections where the decision is locally dependent on agent 1 or 2's report, respectively. Finally, $\tilde{I}(m)$ collects parts of the boundary for which both agents can affect the merger deviation by a marginal deviation from truthtelling. Note that parts of the boundary where the merger function is discontinuous only in agent i's type, that is, the parts given by $I_i(m) \setminus \tilde{I}(m)$, will be perpendicular to the t_i-axis.

Using this notation, this section's main result may be stated as follows:

Proposition 4.3. *Any incentive compatible mechanism (m, s, p) satisfying (BB) must be such that $\tilde{I}(m) \subseteq I^*$.*

[23] The notation I for this set borrows from Jehiel et al. (2005), who call $I(m)$ the 'indifference set'. What gives it this name is that in equilibrium, by Proposition 3.4(a), at any $(t_1, t_2) \in I_i(m) \subseteq I$, agent i will be indifferent between merging and not merging.

Proposition 4.3 says that any connected subset of $I(m)$ which lies outside the boundary I^* of the set of types for which merging is efficient must be either horizontal or vertical. In other words, locally, any merger decision must be either efficient, or independent of at least one agent's private information. This is an immediate consequence of combining the following two independent results:

Lemma 4.4. *For any incentive compatible mechanism satisfying (BB),* $\pi^{\mathrm{M}} \geqslant \pi_1 + \pi_2$ *at any* $(t_1, t_2) \in M^1$ *such that both agents can unilaterally inhibit the merger by a deviation from truthful reporting.*

Lemma 4.5. *For any incentive compatible mechanism satisfying (BB),* $\pi^{\mathrm{M}} \leqslant \pi_1 + \pi_2$ *at any* $(t_1, t_2) \in \tilde{I}(m)$.

See the Appendix for the proofs.

In words, Lemma 4.4 asserts that whenever a mechanism seeks to implement a merger for type combinations such that either agent could unilaterally avoid the merger by misreporting, the merger must be efficient. This is because by (BB), the corresponding deviation payoffs must sum to $\pi_1 + \pi_2$, whereas equilibrium payoffs sum to π^{M}. Hence, if $\pi^{\mathrm{M}} < \pi_1 + \pi_2$, deviating must be profitable for at least one agent. A similar argument can be made for type profiles in M^0 close to $\tilde{I}(m)$ by considering deviations which induce a merger, which leads to Lemma 4.5.

4.2 The Limited Usefulness of Shares in Incentive Design

As a second building block of our analysis, this subsection in turn derives restrictions on share and transfer functions s and p which hold for any m, thus extending our previous analysis in Section 3 by the requirement of budget balance. The next result shows that under budget balance (BB), incentive compatible share and transfer schemes must essentially be independent of reports:

Proposition 4.6. *Any twice differentiable incentive compatible mechanism* (m, s, p) *satisfying (BB) must be such that*

(a) for any $(t_i, t_j), (t'_i, t'_j) \in M^0$, $i \in \{1, 2\}$, *we must have* $p_i^0(t_i, t_j) = p_i^0(t'_i, t'_j)$,

(b) for any $(t_1, t_2) \in M^1$, $i \neq j \in \{1, 2\}$, *such that* $\frac{\partial}{\partial t_j}\pi^{\mathrm{M}} \neq 0$, *we must have* $\frac{\partial}{\partial \tilde{t}_i} s_i = \frac{\partial}{\partial \tilde{t}_i} s_j = \frac{\partial}{\partial \tilde{t}_i} p_i^1 = \frac{\partial}{\partial \tilde{t}_i} p_j^1 = 0$.

See the Appendix for the proofs.

Part (a) is straightforward. Recall that by Proposition 3.2(a), transfers over M^0 must be independent of an agent's own reports. Together with budget balance, this trivially implies that transfers must also be independent of the other's report.

In marginal terms, part (a) implies that transfer functions must satisfy $\frac{\partial}{\partial t_i} p_i^0 = \frac{\partial}{\partial t_i} p_j^0 = 0$ over M^0. Over M^1, on the other hand, the mechanism has *two* incentive devices at its disposal: transfers p_i^1 and shares s_i. As argued before, one might expect this added degree of freedom to give the mechanism operator the possibility of simultaneously varying shares and transfers in an incentive compatible manner. Part (b) of the proposition shows, however, that budget balance similarly keeps both shares and transfers from responding to an agent's private information over M^1 whenever merger profits are responsive to the other's type. Essentially, this is because the manner in which shares and transfers must jointly vary to respect (IC) by Proposition 3.2(b) is compatible with (BB) only if they are constant.

To facilitate a more precise understanding of this key result, assume that joint merger profits π^{M} are strictly increasing in both types—which by Proposition 3.2(b) implies that both agents' shares, s_1 and s_2, must be non-decreasing in reports.[24] Now assume that, starting from an equilibrium situation, agent 2's type t_2 increases while his report $\tilde{t}_2$ is held fixed. The resulting change in agent 1's reporting incentives due to a change in agent 2's type is captured by the term $\frac{\partial^2}{\partial \tilde{t}_1 \partial t_2} u_1 = \frac{\partial}{\partial \tilde{t}_1} s_1 \cdot \frac{\partial}{\partial t_2} \pi^{\mathrm{M}}$, which will be non-negative by Proposition 3.2(b) when joint profits are increasing in both types, and *strictly* positive if agent 1's share depends nontrivially on his report: the increase in agent 2's type inflates the cake of available merger profits, of which agent 1 attempts to obtain a larger slice by distorting his type report upward.

In equilibrium, however, a change in agent 2's type will always be accompanied by a corresponding change in his report. Thus, the mechanism could in principle compensate this *ceteris paribus* incentive to misreport by letting an increase in agent 2's type report cause a *downward* readjustment in agent 1's reporting incentives, that is, by designing the mechanism such that

$$\frac{\partial^2}{\partial \tilde{t}_1 \partial \tilde{t}_2} u_1 + \frac{\partial^2}{\partial \tilde{t}_1 \partial t_2} u_1 = 0. \tag{8}$$

In the specific setting sketched, this would require $\frac{\partial^2}{\partial \tilde{t}_1 \partial \tilde{t}_2} u_1$ to be non-positive, and *strictly* negative whenever agent 1's report affects his share nontrivially. Moreover, by symmetry, exactly the same argument applies to agent 2 upon a change in agent 1's type, thus similarly requiring $\frac{\partial^2}{\partial \tilde{t}_2 \partial \tilde{t}_1} u_2$ to be non-positive, and strictly negative whenever agent 2's share depends nontrivially on his report.

In other words, to compensate for distortionary direct reporting incentives due to a change in the other agent's type, agents' reports must (i) be substitutes in both agents' payoff functions, and (ii) be *strict* substitutes whenever the concerned agent's share depends nontrivially on his own report. Now by budget balance, however, joint payoffs $u_1 + u_2$ must always sum to joint merger profits π^{M} (even out of equilibrium), where the latter of course depends only on true types, not on reports. This in turn implies that

$$\frac{\partial^2}{\partial \tilde{t}_1 \partial \tilde{t}_2} u_1 + \frac{\partial^2}{\partial \tilde{t}_2 \partial \tilde{t}_1} u_2 = 0, \tag{9}$$

[24] While we present the verbal argument for the case in which merger profits are increasing in both agent's types, it should be clear that the same rationale carries over to other constellations after appropriate sign changes.

meaning that *if* reports are to be substitutes in agent 1's payoff, they must be complements in agent 2's payoff. Combined with our previous argument, this immediately implies that reports can never be strictly complementary for any agent under (IC) and (BB), which in turn implies that shares must be locally independent of reports, as stated in Proposition 4.6(b).

4.3 Implementation and Private Gross Benefits to Merging

We now combine the constant-share and -transfer result of the previous subsection with properties of the share- and transfer-functions derived in Proposition 3.4. As we had argued there, incentive compatibility requires shares and transfers to be such that on $I_i(m)$, the gross benefits to merging for agent i, $s_i \cdot \pi^{\mathrm{M}} - \pi_i$ will be exactly offset by the corresponding discrete change in transfers $p_i^1 - p_i^0$. Proposition 4.6 allows us to reinterpret this requirement in terms of level curves of $s_i \cdot \pi^{\mathrm{M}} - \pi_i$ in type space.

The following assumption simplifies application of Proposition 4.6(b):

Assumption 4.7. *The merged entity's profit function is strictly monotone in types, so that* $\frac{\partial}{\partial t_i}\pi^{\mathrm{M}} \neq 0$ *for all* $(t_1, t_2) \in T_1 \times T_2$ *and* $i \in \{1, 2\}$.

Furthermore, we introduce the following terminology:

Definition 4.8. For any merger decision function m, we say that a set $A \in T_1 \times T_2$ is a *merger region of m* if it is a maximal smoothly path connected subset of the closure of M^1, $\overline{M^1}$, in the sense that there exists no smoothly path connected $A' \subseteq \overline{M^1}$ such that $A \subsetneq A'$.[25]

Assumption 4.7 immediately implies that shares must be constant over any merger region by Proposition 4.6(b) and the connectedness property, and thereby (by continuity) also over any subset of $I(m)$ which borders this merger region.[26]

With this terminology in place, our next result may be formulated as follows:

Proposition 4.9. *If Assumption 4.7 holds, any twice differentiable merger mechanism* $\mathcal{M} = (m, s, p)$ *which is implementable under (BB) must be such that, for any* $i \in \{1, 2\}$ *and any connected* $I' \subseteq I_i(m)$*, there exists a* $s_i' \in [0, 1]$ *such that*

$$s_i' \cdot \left[\pi^{\mathrm{M}}(\mathbf{t}'') - \pi^{\mathrm{M}}(\mathbf{t}')\right] = \pi_i(\mathbf{t}'') - \pi_i(\mathbf{t}') \qquad (10)$$

[25] A set Θ is *smoothly path connected* if for every two points $\theta, \theta' \in \Theta$ there exists a differentiable function $f : [0, 1] \to \Theta$ such that $f(0) = \theta$ and $f(1) = \theta'$.

[26] In principle, this leaves open the possibility of implementing *different* shares over different (disjoint) merger regions in order to achieve the required alignment. As we noted in our discussion surrounding Proposition 3.2(b) (and as is shown in that result's proof in the Appendix) there is a global version of condition (7), however, which requires *global* alignment in type between merger profits and shares.

for any $\mathbf{t}', \mathbf{t}'' \in I'$.[27]

In words, π_i and π^{M} must vary proportionally on any connected subset of $I_i(m)$, in the sense that differences in π_i are proportional to differences in π^{M}.

To understand this result, note first that by Proposition 4.6, the continuous extensions $\overline{s_i}$, $\overline{p_i^0}$ and $\overline{p_i^1}$ are constant on I'. Then, by Proposition 3.4(a), $\overline{s_i} \cdot \pi^{\mathrm{M}} - \pi_i = \overline{p_i^1} - \overline{p_i^0}$ for all $(t_1, t_2) \in I'$ with constant values of $\overline{s_i}$, $\overline{p_i^0}$ and $\overline{p_i^1}$, from which the claim immediately follows. Technically, the indifference set $I(m)$ of any merger decision must therefore correspond to a collection of level sets of $s_i' \cdot \pi^{\mathrm{M}} - \pi_i$ for some s_i'.

More fundamentally, Proposition 4.9 shows that the mechanism must fully align agents' preferences over type profiles gross of transfers, $s_i \cdot \pi^{\mathrm{M}} - \pi_i$, and the preferences over type profiles implicit in the merger decision m. To see this, think of the boundary $I(m)$ as an 'indifference set' of the mechanism over types: for all type combinations $\mathbf{t} \in I(m)$, the mechanism is 'indifferent' between implementing a merger or not. By Proposition 4.9, at least one agent must similarly be indifferent between such type combinations in the sense that they yield the same gross benefits to merging, $s_i \cdot \pi^{\mathrm{M}} - \pi_i$. Thus, the share must be chosen such that on the boundary of a merger region, both the mechanism and at least one agent value *both* agents' private information on the same terms.[28]

This interpretation of Proposition 4.9 is further augmented by Proposition 3.4(b), by which agents' gross benefits to merging $s_i \cdot \pi^{\mathrm{M}} - \pi_i$ must be aligned with the merger decision on the boundary. Thus, not only must $I(m)$ coincide with indifference sets of gross benefits to merging $s_i \cdot \pi^{\mathrm{M}} - \pi_i$, but these benefits to merging must also be locally increasing in the direction in type space in which a merger is implemented.

Eventually, however, the problem is that all this must be achieved both by a share that is locally constant, and on a boundary that is already strongly restricted by Proposition 4.3. We consider the joint implication of these restrictions in the next section.

4.4 Why Non-Trivial Mechanisms are Generally not Implementable

We now argue that the conditions derived so far essentially preclude the implementation of non-trivial merger decisions by means of share-transfer mechanisms in quite

[27] Note that on any such subset of $\tilde{I}(m)$, where the decision function is discontinuous in *both* types, condition (10) must of course hold simultaneously for both agents $i = 1, 2$. However, due to the fact that $I' \subseteq I^*$ by Proposition 4.3, $s_i' \cdot \pi^{\mathrm{M}} - \pi_i$ being constant on I' for one agent i immediately implies the same for the other agent j with $s_j' = 1 - s_i'$.

[28] This part of our analysis has much in common with Jehiel et al. (2005), who derive similar restrictions in the context of ex-post implementation with *multidimensional* information. On the one hand, derivations are simplified in our case by the additional budget balance requirement and one-dimensional information. On the other, the presence of shares makes the argument slightly more involved (Jehiel et al., 2005, consider only binary decisions).

general environments. Our analysis builds from Proposition 4.9, which implies a geometric condition which profits function must satisfy on the boundary between M^0 and M^1. Proposition 4.10 will state the implications of this condition in marginal terms. To formulate the result, for any $f : \mathbb{R}^2 \to \mathbb{R}$ and $\mathbf{v} \in \mathbb{R}^2$, we let $\partial_{\mathbf{v}} f = \mathbf{v}^{\mathrm{T}} \nabla f$ denote the directional derivative of f along $\mathbf{v}$.

Proposition 4.10. *Under any incentive compatible mechanism satisfying (BB),*

(a) for any $\mathbf{t} \equiv (t_1, t_2) \in I' \subseteq \tilde{I}(m)$ *where* I' *is connected, either* $\nabla \pi^{\mathrm{M}}(\mathbf{t}) = 0$ *or the following conditions must hold simultaneously for any tangent* $\mathbf{r}(\mathbf{t})$ *of* I^* *in* $\mathbf{t}$[29]*:*

- *(a1)* $\partial_{\mathbf{r}(\mathbf{t})} \left(\partial_{\mathbf{r}(\mathbf{t})} \pi_i / \partial_{\mathbf{r}(\mathbf{t})} \pi^{\mathrm{M}} \right) = 0,$
- *(a2)* $\dfrac{\partial_{\mathbf{r}(\mathbf{t})} \pi_i}{\partial_{\mathbf{r}(\mathbf{t})} \pi^{\mathrm{M}}} \in [0, 1],$ *and*
- *(a3) letting* $\mathbf{q}(\mathbf{t}) = \nabla(\pi^{\mathrm{M}} - \pi_1 - \pi_2)$, $\dfrac{\partial_{\mathbf{r}(\mathbf{t})} \pi_i}{\partial_{\mathbf{r}(\mathbf{t})} \pi^{\mathrm{M}}} \cdot \partial_{\mathbf{q}(\mathbf{t})} \pi^{\mathrm{M}} - \partial_{\mathbf{q}(\mathbf{t})} \pi_i \geqslant 0.$

(b) for any $(t_1, t_2) \in I' \subseteq I_i(m) \setminus \tilde{I}(m)$ *where* I' *is connected, either* $\frac{\partial}{\partial t_j} \pi_i = \frac{\partial}{\partial t_j} \pi^{\mathrm{M}} = 0$*, or the following conditions must hold simultaneously:*

- *(b1)* $\frac{\partial}{\partial t_j} \left(\frac{\partial}{\partial t_j} \pi_i / \frac{\partial}{\partial t_j} \pi^{\mathrm{M}} \right) = 0,$
- *(b2)* $\dfrac{\frac{\partial}{\partial t_j} \pi_i}{\frac{\partial}{\partial t_j} \pi^{\mathrm{M}}} \in [0, 1],$ *and*
- *(b3)* m *is locally increasing (decreasing) in* t_i *if* $\dfrac{\frac{\partial}{\partial t_j} \pi_i}{\frac{\partial}{\partial t_j} \pi^{\mathrm{M}}} \cdot \pi^{\mathrm{M}} - \pi_i$ *is.*

See the Appendix for the proof.

Statements (a1), (a2) and (b1), (b2), respectively, are immediate implications of the proportional-variation requirement in (10). (a3) and (b3) rephrase the idea that the mechanism must give types that are meant to merge greater incentives to do so than types that are not meant to merge.

The power of Proposition 4.10 becomes clear in combination with Proposition 4.3: As the boundary of the merger set coincides with boundaries of the efficiency set or with horizontal or vertical lines by Proposition 4.3, Proposition 4.10 boils down to restrictions on the merger environment on the efficiency set I^* and the parallels to the axes. For instance, conditions 4.10 (a) and (b) together imply that $\partial_{\mathbf{r}(\mathbf{t})} \pi_i / \partial_{\mathbf{r}(\mathbf{t})} \pi^{\mathrm{M}}$ must be constant along the boundary of the efficiency set, or that $\frac{\partial}{\partial t_j} \pi_i / \frac{\partial}{\partial t_j} \pi^{\mathrm{M}}$ must be constant on a line parallel to the t_j-axis for some $i \in \{1, 2\}$, $j \neq i$. These conditions are obviously satisfied only by a small subclass of environments. Intuitively, these requirements can only be satisfied if the effects of modifying type profiles on stand-alone profits are very similar to those on joint profits. In principle,

[29] That is, for any $\mathbf{r}(\mathbf{t}) \in \{\mathbf{r}' \in \mathbb{R}^2 : \mathbf{r}' \perp \nabla \pi^{\mathrm{M}}(\mathbf{t})\}$

these insights could be used to show that non-trivial mechanisms without *ex-post* regret are not implementable in examples 1 and 2. However, we will show this using a simpler approach in Section 6.

We will see in Section 7.2 below, however, there is an important class of 'degenerate'—but not irrelevant—merger environments where this result does not apply: If at least one agent's stand-alone profits π_i are common knowledge, the efficient decision is indeed implementable.

5 Requiring Individual Rationality

We have argued above that in most general cases, the restrictions imposed by budget balance will be so severe as to permit only trivial merger decisions of the sort 'never merge' or 'always merge'. We now argue that the participation constraints will eliminate the latter option, so that under budget balance, an incentive compatible mechanism can only reproduce the status quo.

To establish this, we use the very simple result that under budget balance, the participation constraint prohibits the merger mechanism from ever implementing an inefficient merger:

Proposition 5.1. *Any incentive compatible merger mechanism satisfying (BB) and (IR) must be such that*

$$\pi^{\mathrm{M}}(t_1,t_2) \geqslant \pi_1(t_1,t_2) + \pi_2(t_2,t_2) \qquad \textit{for all } (t_1,t_2) \in M^1. \tag{11}$$

Proof. Incentive compatibility and individual rationality (i.e., participation) imply that for any (t_1,t_2) such that $m(t_1,t_2)=1$,

$$s_i(t_i,t_j)\cdot\pi^{\mathrm{M}}(t_1,t_2) - p_i^1(t_i,t_j) \geqslant \pi_i(t_i,t_j),$$

which, by adding up across both types $i=1,2$, implies

$$[s_1(t_1,t_2)+s_2(t_2,t_1)]\cdot\pi^{\mathrm{M}}(t_1,t_2) - p_1^1(t_1,t_2) - p_2^1(t_2,t_1) \geqslant \pi_1(t_1,t_2)+\pi_2(t_2,t_1).$$

Together with (BB), this in turn immediately implies (11). □

Proposition 5.1 captures the simple fact that if a merger is executed even though it is inefficient, then—unless the mechanism operator tosses in subsidies—there are simply not enough profits around to satisfy both parties' *ex-post* participation constraints. Thus, *ex-post* inefficiencies in an individually rational mechanism can never be due to mergers occurring even though they are inefficient, but only due to efficient mergers foregone.

A simple corollary to Proposition 5.1 is that in any of the cases in which the restrictions derived in Section 4 permit only trivial (i.e., constant) merger decisions, the only implementable decision under voluntary participation is in fact the 'never merge'-rule. Furthermore, it is easily seen that under voluntary participation, transfers if no merger occurs must be zero, so $p_i^0 = 0$, $i \in \{1, 2\}$ under (IR). Hence, such a mechanism will always preserve the status quo:

Corollary 5.2. *Under voluntary participation (BB) and (IR), if there exists any type constellation* (t_1, t_2) *such that merging is strictly inefficient (so* $\pi^{\mathrm{M}} < \pi_1 + \pi_2$*), then the only trivial implementable merger mechanism is that for which mergers never occur, and payments are always zero.*

6 Generalized Mechanisms

This section investigates implementation under more general mechanisms than the share-transfer mechanisms analyzed hitherto. We derive conditions for non-trivial implementation which, while not as generally exclusive as those derived in Section 4.4, nonetheless prevent non-trivial implementation in many merger environments of interest.

The main motivation for doing so is that real mergers often involve contracts that condition post-merger profits of the participants to the deal on total merger profits in a non-linear fashion, so that the contracts cannot be captured by share-transfer mechanisms.[30] An additional benefit of our generalization is that it sheds light on the share-transfer case: Though our approach to generalized merger mechanisms does not pertain to as many merger environments as our treatment of the share-transfer mechanisms in the preceding sections, it yields very simple non-existence results in those environments where it does apply. The environments are still fairly general, including for instance our examples 1 and 2.

Definition 6.1. A *generalized (direct) merger mechanism* $\mathcal{M}$ is a tuple $(m, \hat{\pi}^{\mathrm{M}}, p^0)$ consisting of a merger decision function m, merger-profit sharing rules $\hat{\pi}^{\mathrm{M}} = (\hat{\pi}_1^{\mathrm{M}}, \hat{\pi}_2^{\mathrm{M}})$, and no-merger transfer functions $p^0 = (p_1^0, p_2^0)$, which are defined as follows:

(i) The *merger decision function* $m : T_1 \times T_2 \to \{0, 1\}$ maps a combination of reports $\tilde{\mathbf{t}} = (\tilde{t}_1, \tilde{t}_2)$ into a merger outcome, resulting in sets M^0 and M^1 as in Definition 2.2.

[30] Such more general mechanisms are also analyzed by Brusco et al. (2005). Practical examples include collars (Officer, 2004), which use changes in stock prices to determine the partners remuneration, and contingent value rights (Hietala et al., 2003), where sellers obtain put options on the shares of the new entity.

(ii) The *merger-profit sharing rules* $\hat{\pi}_i^{\mathrm{M}} : M^1 \times \mathbb{R}_{\geqslant 0} \to \mathbb{R}$, $i \in \{1,2\}$, map a combination of reports $\tilde{\mathbf{t}}$ for which $m(\tilde{\mathbf{t}}) = 1$ and realized merger profits $\pi^{\mathrm{M}}(\mathbf{t})$ into a payoff to firm i.

(iii) The *no-merger transfer functions* $p_i^0 : M^0 \to \mathbb{R}$, $i \in \{1,2\}$, assign a transfer payment from player i to the mechanism operator for reported types $\tilde{\mathbf{t}}$ such that $m(\tilde{\mathbf{t}}) = 0$.

For any combination of types $\mathbf{t}$ and reports $\tilde{\mathbf{t}}$, the mechanism results in payoffs

$$u_i(\tilde{\mathbf{t}};\mathbf{t}) = \begin{cases} \pi_i(\mathbf{t}) - p_i^0(\tilde{\mathbf{t}}), & m(\tilde{\mathbf{t}}) = 0, \\ \hat{\pi}_i^{\mathrm{M}}[\tilde{\mathbf{t}};\pi^{\mathrm{M}}(\mathbf{t})], & m(\tilde{\mathbf{t}}) = 1 \end{cases} \tag{12}$$

for each agent i. Thus, the mechanisms considered in this section permit conditioning agents' payoffs on joint merger profits π^{M} in a more general fashion than is permitted by share-transfer mechanisms. Note that share-transfer mechanisms obtain in this more general setting by restricting $\hat{\pi}_i^{\mathrm{M}}$ to functions of the form $\hat{\pi}_i^{\mathrm{M}}[\tilde{\mathbf{t}},\pi^{\mathrm{M}}(\mathbf{t})] = s_i(\tilde{\mathbf{t}}) \cdot \pi^{\mathrm{M}}(\mathbf{t}) - p_i^1(\tilde{\mathbf{t}})$, that is, to functions which are affine in π^{M}.

Moreover, we introduce the following generalization of (BB):

Definition 6.2. A generalized merger mechanisms satisfies *budget balance (BB')* if

$$\sum_{i=1,2} u_i(\tilde{\mathbf{t}};\mathbf{t}) = \begin{cases} \sum_{i=1,2} \pi_i(\mathbf{t}), & m = 0, \\ \pi^{\mathrm{M}}(\mathbf{t}), & m = 1, \end{cases} \tag{13}$$

for all $\tilde{\mathbf{t}}, \mathbf{t} \in T_1 \times T_2$.

It is important to note that this generalization inherits from (BB) the property that the budget must be balanced also *off equilibrium*, that is, when $\tilde{\mathbf{t}} \neq \mathbf{t}$.

To simplify the argumentation, we will assume that the payoff functions $\hat{\pi}_i^{\mathrm{M}}$, $i = 1,2$, are differentiable in all arguments. Furthermore, we will assume that in addition to (IC) and (BB'), such a generalized mechanism must satisfy (IR). In straightforward extension to Proposition 5.1, it is quickly checked that generalized mechanisms satisfying (BB') and (IR) must be such that $V_i(\mathbf{t}) = \pi_i(\mathbf{t})$ for $\mathbf{t} \in M^0$ (i.e., $p_i^0(\tilde{\mathbf{t}}) = 0$ everywhere). Moreover, by an argument paralleling that in Proposition 3.4(a), V_i must again be continuous over $I_i(m)$. Combined with the fact that (IR) requires $V_i(\mathbf{t}) \geqslant \pi_i(\mathbf{t})$ over M^1, these insights immediately give rise to the following lemma:

Lemma 6.3. *For any* $\mathbf{t} \in I_i(m)$*,* $i \in \{1,2\}$*, and any vector* $\mathbf{r} \in \mathbb{R}^2 \setminus \{\mathbf{0}\}$ *pointing into the interior of* M^1 *at* $\mathbf{t}$*,* $\partial_{\mathbf{r}} V_i(\mathbf{t}) \geqslant \partial_{\mathbf{r}} \pi_i(\mathbf{t})$*.*

Next, over M^1, the gradient of the value function, ∇V_i, is related to the gradient of merger profits, $\nabla \pi^{\mathrm{M}}$, as follows:

Lemma 6.4. $\nabla V_i = \frac{\partial}{\partial \pi^{\mathrm{M}}} \hat{\pi}_i^{\mathrm{M}} \cdot \nabla \pi^{\mathrm{M}}$ *for* $i = 1, 2$ *over* M^1*.*

Proof. Since $\nabla V_i = \nabla_{\tilde{\mathbf{t}}} \hat{\pi}_i^{\mathrm{M}} + \frac{\partial}{\partial \pi^{\mathrm{M}}} \hat{\pi}_i^{\mathrm{M}} \cdot \nabla \pi^{\mathrm{M}}$ over M^1 by definition of the value function V_i^1, we need only show that $\nabla_{\tilde{\mathbf{t}}} \hat{\pi}_i^{\mathrm{M}} = \mathbf{0}$, which in turn follows directly from $\frac{\partial}{\partial \tilde{t}_i} \hat{\pi}_i^{\mathrm{M}} = 0$ due to (IC), and $\frac{\partial}{\partial \tilde{t}_j} \hat{\pi}_i^{\mathrm{M}} + \frac{\partial}{\partial \tilde{t}_j} \hat{\pi}_j^{\mathrm{M}} = 0$ by (BB'). □

From Lemma 6.4, it immediately follows that $\partial_{\mathbf{r}} V_i = \frac{\partial}{\partial \pi^{\mathrm{M}}} \hat{\pi}_i^{\mathrm{M}} \cdot \partial_{\mathbf{r}} \pi^{\mathrm{M}}$, which in turn implies that each agent's value function V_i must be constant on level sets of π^{M} over M^1. Combined with Lemma 6.3, this leads to the following key result:

Proposition 6.5. *Any general mechanism satisfying (IC), (BB'), and (IR) must be such that,*

(a) for any $\mathbf{t} \in I_i(m)$*,* $i \in \{1, 2\}$*, and any* $\mathbf{r} \in \mathbb{R}^2 \setminus \{\mathbf{0}\}$ *such that (i)* $\partial_{\mathbf{r}} \pi^{\mathrm{M}}(\mathbf{t}) = 0$ *and (ii)* $\mathbf{r}$ *points toward the interior of* M^1 *at* $\mathbf{t}$*, we must have* $\partial_{\mathbf{r}} \pi_i \leqslant 0$*;*

(b) for any $\mathbf{t} \in I_i(m)$*,* $i \in \{1, 2\}$ *such that no* $\mathbf{r}$ *satisfying (i) and (ii) in (a) exists, either the boundary* $\overline{M^0} \cap \overline{M^1}$ *must be kinked (i.e., non-smooth), or the level curves of* π^{M} *and* π_i *must have identical slopes.*

See the Appendix for the proofs.

By part (a), at any point $\mathbf{t}$ on the boundary of M^1 where (i) the boundary is intersected by a level curve of π^{M} and (ii) agent i can alter the merger decision by a marginal deviation from truthful reporting, π_i must be locally decreasing along the part of the level curve of π^{M} which penetrates M^1. Part (b) essentially serves to formalize the wide applicability of part (a): at points $\mathbf{t} \in I_i(m)$ where the level curve of π^{M} does not penetrate M^1, the level curve of π^{M} must be tangent to the boundary, which implies by the required continuity of the value functions that the level curve of π_i must be tangent to the boundary as well. Thus, unless π^{M} and π_i are themselves aligned in a rather narrow sense, part (a) is applicable.

The power of Proposition 6.5 is illustrated in the following corollary, which excludes non-trivial implementation in a large and relevant class of environments:

Corollary 6.6. *For merger environments such that (i)* $\frac{\partial}{\partial t_i} \pi_i > 0$*, (ii)* $\frac{\partial}{\partial t_j} \pi_i < 0$*, and (iii)* $\frac{\partial}{\partial t_i} \pi^{\mathrm{M}} > 0$ *for all* $i, j \in \{1, 2\}$*,* $i \neq j$*,* $\tilde{I}(m)$ *must be empty for all mechanisms satisfying (IC), (BB'), and (IR).*

See the Appendix for the proof.

Conditions (i) and (iii) are very natural whenever the type variable corresponds to higher productivity; with the additional requirement that firms are competing in the same market, condition (ii) will typically hold.

The intuition for Corollary 6.6 is quickly seen: Over $\tilde{I}(m) = I_1(m) \cup I_2(m)$, Proposition 6.5 is applicable for both agents $i = 1, 2$. Hence, on parts of the boundary of M^1 where *both* agents can unilaterally alter the merger decision by a marginal deviation from truth-telling, *both* π_1 and π_2 must be decreasing along any level curve of π^{M} running into M^1. Now, the merger environments described in Corollary 6.6 are such that, in a t_1/t_2-plane with t_1 on the ordinate, the level curves of π^{M} have a negative slope by condition (iii). Furthermore, by (i) and (ii), π_1 will be strictly increasing in any northwest direction, whereas π_2 will be strictly decreasing in any southeast direction. Hence, on any level curve of π^{M}, π_1 and π_2 will be non-constant, and increasing in opposite directions. Thus, $\tilde{I}(m)$ must be empty by Proposition 6.5.

7 Considering Specific Merger Environments

We now reconsider the specific examples of merger environments presented in Section 2.2 in light of our results. We first show that only trivial merger decisions are implementable in Examples 1 and 2 due to the results in Section 6. Next, we show more generally that the main obstacle to non-trivial implementation lies in private information on *stand-alone* profits π_i rather than private information on *post-merger* profits π^{M}, which permits efficient implementation in examples 3 and 4.

7.1 Reconsidering the Examples

We start by reconsidering Examples 1 and 2 presented in Section 2.2. As it turns out, any scope for non-trivial *general* mechanisms of the form presented in Section 6 is swiftly eliminated by Corollary 6.6 and Proposition 6.5, respectively, which in particular also rules out any non-trivial share/transfer mechanisms.

Example 1 continued (Averaging-Cost Mergers). Recall that in this example, merger profits π^{M} are strictly increasing in each type, whereas stand-alone profits are strictly increasing in own type and strictly decreasing in the other's type. Thus, the example falls into the class of environments covered by Corollary 6.6, implying that any *general* mechanism (and thereby also any share/transfer mechanism) can only implement merger decisions for which $\tilde{I}(m) = \emptyset$, that is, merger decisions which depend only trivially on at least one agent's report.

For share-transfer mechanisms, such 'semi-trivial' merger decision functions can further be ruled because this class of merger environments violates the constant-variation requirement of Proposition 4.9 on horizontal and vertical lines in $T_1 \times T_2$-space, as is quickly checked by means of of Proposition 4.10(b1). Hence, only share-transfer mechanisms of the form $m(\mathbf{t}) = 0$, $\forall \mathbf{t} \in T_1 \times T_2$, are implementable.

Example 2 continued (Rationalization Mergers). In this case, Corollary 6.6 is no longer directly applicable, as π^{M} will be constant in firm i's type for $t_i < t_j$,

$j \neq i$. However, the argument implicit in Corollary 6.6 can quickly be extended to this case by reconsidering Proposition 6.5.[31]

Recall that by Proposition 6.5, at any points on $\tilde{I}(m)$, level curves of π^{M}, π_1, and π_2 must either have identical slopes, or both π_1 and π_2 must be (locally) increasing (decreasing) in the same direction along level curves of π^{M}. Since level curves of π^{M} consist only of horizontal/vertical segments, the first possibility is trivially ruled out by the fact that level curves of π_1 and π_2 have strictly positive slopes. The second in turn is ruled out by $\frac{\partial}{\partial t_i}\pi_i$ and $\frac{\partial}{\partial t_i}\pi_i$ having opposite signs for any $i = 1, 2$, $j \neq i$, implying that π_1 and π_2 will be increasing in opposite directions on any π^{M}-level curve.[32] Hence, the set $\tilde{I}(m)$ must be empty for any general mechanism. Again, for share-transfer mechanisms, semi-trivial merger decisions making use of only *one* agent's report are furthermore quickly ruled out by Proposition 4.10(b1).

By having firms' stand-alone values independent of private information, Examples 3 and 4 in turn fall into a broader class of environments for which results presented in the next section spell good news concerning not just non-trivial, but in fact *efficient* implementation.

7.2 Implementation when One Firm's Standalone-Value is Common Knowledge

The results presented in Section 4.4 for share/transfer-mechanisms and in Section 6 for more general mechanisms crucially hinge on the assumption that neither agent's stand-alone profits are common knowledge. This is shown by the following result, pertaining to environments where one of the agents' stand-alone profits (w.l.o.g. agent 2's) are independent of t_1 and t_2.

Proposition 7.1. *If neither agent holds any private information on agent 2's stand-alone profits, so $\nabla\pi_2 = \mathbf{0}$, then the efficient merger decision can be implemented by a share/transfer-mechanism satisfying (BB) and (IR) by agent 1 obtaining the full share in the event of a merger and paying π_2 to agent 2, and zero transfers if no merger occurs. Moreover, if $\nabla\pi^{\mathrm{M}}$ and $\nabla\pi_1$ are not collinear on I^*, the shares and transfers required to implement any efficient merger decision under (BB) and (IR) are unique.*

[31] Alternatively, we may interpret Example 2 as the limiting case of a class of environments to which Corollary 6.6 *does* directly apply. Consider the class of environments for which the costs of the merged entity are given by the *generalized* mean of stand-alone-costs: $c^{\mathrm{M}} = (\frac{1}{2}c_1^{\nu} + \frac{1}{2}c_2^{\nu})^{\frac{1}{\nu}}$, $\nu \in \mathbb{R}$. It is quickly seen that Corollary 6.6 is applicable to any such environment, and that the environment of Example 2 obtains as $\nu \to -\infty$. Note that this generalized form for c^{M} nests other interesting cases such as the arithmetic mean (i.e., Example 1) for $\nu = 1$, the harmonic mean for $\nu = -1$, the geometric mean as $\nu \to 0$, and the maximum-function as $\nu \to \infty$.

[32] More formally, the proof of Corollary 6.6 is quickly adapted to cases where, at every $\mathbf{t} \in T_1 \times T_2$, $\frac{\partial}{\partial t_i}\pi^{\mathrm{M}}(\mathbf{t}) > 0$ need only hold for one agent.

See the Appendix for the proof.

It is easy to see why—in spite of our general result—this particular class of environments allows for implementation of the efficient decision: Given $\nabla\pi_2 = \mathbf{0}$, $\nabla(\pi^{\mathrm{M}} - \pi_1 - \pi_2) = \nabla\pi^{\mathrm{M}} - \nabla\pi_1$, so that for $s_1 = 1$, agent 1's gross benefits to merging $s_1 \cdot \pi^{\mathrm{M}} - \pi_1$ are perfectly aligned with social benefits to merging, whereas agent 2 is always indifferent between merging and not merging. It is also obvious that, for any other distribution of shares, private and social goals will generally be misaligned, which explains the uniqueness-result.

A trivial corollary of Proposition 7.1 is that implementation of the efficient merger decision is also possible if *both* agents' stand-alone profits are common knowledge, that is, if each agent's private information pertains only to profits under a merger.[33] Thus, both the 'Synergy'- and 'Corporate Culture'-examples discussed in Section 2.2 permit implementation of the efficient decision. However, in these settings, the uniqueness result in Proposition 7.1 no longer holds, and it is easily verified that in any environment with $\nabla\pi_1 = \nabla\pi_2 = \mathbf{0}$, the efficient decision can be implemented with *any* allocation of shares, so long as transfers are chosen accordingly. It is worth noting though that there is a unique allocation of shares which implements the efficient merger in a 'cash-free' way: By Proposition 3.4(a), $p_i^0 = p_i^1 = 0$ for $i = 1, 2$ requires $s_i = \pi_i/(\pi_1 + \pi_2)$, so that the allocation of shares must be proportional to firms' (commonly known) stand-alone values.[34]

8 Conclusion

This paper has shown that share-transfer mechanisms cannot implement *any* mergers in an *ex-post* incentive compatible fashion, if balance in shares and transfers and individual rationality are required. Even for more general mechanisms, the scope for implementation is limited. We have thereby shown that even if the potential merger candidates jointly possess all information necessary to assess the profitability of a merger, the dispersion of this information across parties poses a serious problem to the implementation of mergers.

This result stands in stark contrast to the encouraging results obtained in the literature on auctions with interdependent valuations (cf. Maskin, 1992; Dasgupta and Maskin, 2000; Jehiel and Moldovanu, 2001; Perry and Reny, 2002), where it has been shown that an efficient outcome is quite generally attainable. While the two problems of efficiently allocating a good to buyers and forming efficient partnerships certainly display inherently dissimilar structures, recall that we argued at the outset that one should expect the additional degree of freedom in incentive design which

[33] Proposition 2 in Brusco et al. (2005) is analogous; but recall that their setting differs from ours.

[34] Brusco et al. (2005) come to an analogous conclusion in their setting.

is offered by the possibility of designing share allocations in addition to transfers to improve prospects for an efficient mechanism in the merger context. As arguments in Section 7.2 show, the main hindrance to efficient implementation in our setting comes from private knowledge on reservation payoffs (i.e., on stand-alone profits).

The fact that merger deals are rather frequently observed in reality must of course lead us to conclude that at least one of the requirements we have imposed on the merger decision process will typically fail to hold in real world merger deals. Most likely, the requirement of *ex-post* incentive compatibility is excessively strong, meaning parties in actual merger negotiations choose strategies that are optimal only in *expected* terms. However, in such a context, our results nonetheless provide the interesting conclusion that in any such Bayesian mechanism, there must be some merger outcomes (be they efficient or not) involving *ex-post* regret on behalf of at least one of the parties. This insight provides an alternative explanation of the 'unhappy marriages' often alluded to, without invoking any agency arguments concerning self-interested managers.

An important caveat applies to this argument, however, as *ex-post* regret about the strategy played need not necessarily imply regret about the merger decision itself. Indeed, regret may concern an alternative strategy which would just as well have led to a merger, but which would have provided the agent with a larger slice of the merger cake. On the other hand, it is clear that any time a *Bayesian* mechanism implements an *ex-post inefficient* merger, at least one of the agents must indeed exhibit *ex-post* regret about the merger decision itself. An important future line of research should therefore lie in investigating the efficiency properties of Bayesian merger mechanisms.[35] What makes this task somewhat cumbersome is that in a *Bayesian* implementation context, the distribution of types crucially affects the class of implementable mechanisms and thereby the set of implementable merger decisions, which will likely make the analysis much less tractable than the *ex-post* approach.

Acknowledgements

The authors thank David A. Miller, Andrea Patacconi and Andrew Postlewaite for very helpful discussions. Financial support for this project from the Swiss National Science Foundation under Grant No. PBZH1-108484 is gratefully acknowledged.

[35] There is another argument in favor of *a priori* focussing attention on *efficient* mechanisms—even in our *ex-post* framework—based on a renegotiation-proofness argument. Indeed, any time a mechanism (be it *ex-post* or Bayesian incentive compatible) implements an inefficient merger decision and parties have somehow revealed their type through the mechanism, there is ample scope for renegotiation. The applicability of any inefficient mechanism therefore relies crucially on the mechanism operator's ability to commit. However, standard reputation arguments to justify the credibility of such a commitment are likely to fail in the merger context, where the problem to be solved by the mechanism is of an inherent one-shot nature.

Appendix: Proofs

Proof of Proposition 3.2. (a) Using requirement (IC) along with (1),

$$U_i(t_i';t_i'',t_j) - U_i(t_i'';t_i'',t_j) = p_i^0(t_i'',t_j) - p_i^0(t_i',t_j) \leqslant 0$$
$$U_i(t_i'';t_i',t_j) - U_i(t_i';t_i',t_j) = p_i^0(t_i',t_j) - p_i^0(t_i'',t_j) \leqslant 0,$$

from which the claim immediately follows.

(b) Equation (6) follows immediately from (IC), by which $\frac{\partial}{\partial \hat{t}_i} u_i(t_i;t_i,t_j) = 0$. To prove condition (7), we show more generally that for any $t_i', t_i'' \in T_i$ and $t_j \in T_j$ such that $(t_i',t_j),(t_i'',t_j) \in M^1$,

$$[s_i(t_i'',t_j) - s_i(t_i',t_j)] \cdot [\pi^{\mathrm{M}}(t_i'',t_j) - \pi^{\mathrm{M}}(t_i',t_j)] \geqslant 0. \tag{A.1}$$

To see this, note that in order to avoid type t_i'' claiming to be of type t_i' and vice versa, we must have

$$s_i(t_i'',t_j)\cdot\pi^{\mathrm{M}}(t_i'',t_j) - p_i^1(t_i'',t_j) \geqslant s_i(t_i',t_j)\cdot\pi^{\mathrm{M}}(t_i'',t_j) - p_i^1(t_i',t_j)$$
$$s_i(t_i',t_j)\cdot\pi^{\mathrm{M}}(t_i',t_j) - p_i^1(t_i',t_j) \geqslant s_i(t_i'',t_j)\cdot\pi^{\mathrm{M}}(t_i',t_j) - p_i^1(t_i'',t_j),$$

respectively. Jointly, these inequalities immediately imply (A.1), of which (7) is simply an implication in marginal terms. □

Proof of Proposition 3.4. (a) Consider any two sequences $(t_{i,n}^k)_{n=1}^\infty$, $k \in \{0,1\}$, such that $\lim_{n\to\infty} t_{i,n}^k = t_i^\circ$, and such that $(t_{i,n}^0,t_j) \in M^0$ and $(t_{i,n}^1,t_j) \in M^1$ for all n. Such sequences exist by assumption. By incentive compatibility (IC),

$$U_i(t_{i,n}^1;t_{i,n'}^0,t_j) - U_i(t_{i,n'}^0;t_{i,n'}^0,t_j) \leqslant 0 \qquad \text{and}$$
$$U_i(t_{i,n'}^0;t_{i,n}^1,t_j) - U_i(t_{i,n'}^1;t_{i,n'}^1,t_j) \leqslant 0$$

for all n, n'. Therefore,

$$s_i(t_{i,n}^1,t_j)\cdot\pi^{\mathrm{M}}(t_{i,n'}^0,t_j) - \pi_i(t_{i,n'}^0,t_j) \leqslant p_i^1(t_{i,n}^1,t_j) - p_i^0(t_{i,n'}^0,t_j)$$
$$\leqslant s_i(t_{i,n}^1,t_j)\cdot\pi^{\mathrm{M}}(t_{i,n}^1,t_j) - \pi_i(t_{i,n}^1,t_j).$$

Letting $n, n' \to \infty$, using the continuity conditions on π_i and π_i^{M} and rearranging gives the result.

(b) We prove this result by using properties of the *value function* collected in the following Lemma:

Lemma A.1 (Properties of the Value Function). *Letting* $V_i(t_i,t_j) \equiv U_i(t_i;t_i,t_j)$ *denote agent i's* equilibrium utility *or* value function, *under any differentiable incentive compatible mechanism* (m,s,p),

(a) $V_i(t_i,t_j)$ *must be continuous in* t_i *for any* $t_j \in T_j$,
(b) $\frac{\partial}{\partial t_i} V_i(t_i,t_j) = \frac{\partial}{\partial t_i}\pi_i(t_i,t_j)$ *for* $(t_1,t_2) \in M^{0^\circ}$,
(c) $\frac{\partial}{\partial t_i} V_i(t_i,t_j) = s_i(t_i,t_j)\cdot\frac{\partial}{\partial t_i}\pi^{\mathrm{M}}(t_1,t_2)$ *for* $(t_1,t_2) \in M^{1^\circ}$,

For interior regions of M^0 and M^1, part (a) immediately follows from (1) and the fact that profit, share, and transfer functions are all continuous by assumption. Continuity on the boundary is an immediate implication of Proposition 3.4(a). Parts (b) and (c) represent straightforward applications of the envelope theorem, by which $\frac{\partial}{\partial t_i} V_i = \frac{\partial}{\partial t_i} U_i$.

Returning to the proof of Proposition 3.4(b), assume first that $m(\tilde{t}_i, \tilde{t}_j)$ is increasing in $\tilde{t}_i$ at (t_1°, t_2°). Fixing $t_j = t_j^\circ$ and suppressing it for notational convenience, let $U_i(t_i^\circ; t_i^\circ)$ denote the equilibrium utility of an agent of type t_i°. Now, for $\varepsilon > 0$ small, by Lemma A.1(c), an agent of type $t_i^\circ + \varepsilon$ will have equilibrium utility

$$V(t_i^\circ + \varepsilon) = V(t_i^\circ) + \int_{t_i^\circ}^{t_i^\circ + \varepsilon} s_i(\tau) \cdot \frac{\partial}{\partial t_i} \pi^{\mathrm{M}}(\tau) d\tau. \tag{A.2}$$

By reporting $t_i^\circ - \delta$ for $\delta > 0$ small, using Lemma A.1, his utility will be

$$\begin{aligned} &\pi_i(t_i^\circ + \varepsilon) - p_i^0(t_i^\circ - \delta) = \pi_i(t_i^\circ + \varepsilon) - \pi_i(t_i^\circ - \delta) + \pi_i(t_i^\circ - \delta) - p_i^0(t_i^\circ - \delta) \\ &= \pi_i(t_i^\circ + \varepsilon) - \pi_i(t_i^\circ - \delta) + V_i(t_i^\circ - \delta) = \int_{t_i^\circ - \delta}^{t_i^\circ + \varepsilon} \frac{\partial}{\partial t_i} \pi_i(\tau) d\tau + V_i(t_i^\circ) - \int_{t_i^\circ - \delta}^{t_i^\circ} \frac{\partial}{\partial t_i} \pi_i(\tau) d\tau \\ &= V_i(t_i^\circ) + \int_{t_i^\circ}^{t_i^\circ + \varepsilon} \frac{\partial}{\partial t_i} \pi_i(\tau) d\tau. \end{aligned} \tag{A.3}$$

Combining (A.2) and (A.3), his deviation gain will be negative for small ε only if

$$\frac{\partial}{\partial t_i} \pi_i(t_i^\circ) \leqslant s_i(t_i^\circ) \cdot \frac{\partial}{\partial t_i} \pi^{\mathrm{M}}(t_i^\circ),$$

as was to be shown. Proving the claim for a locally *decreasing* m merely requires a switch in signs for ε and δ. □

Proof of Lemma 4.4. By assumption, for each $i = 1, 2$, there exists a t_i' such that $(t_i', t_j) \in M^0$. By (IC), we must have

$$V_i(t_i, t_j) \geqslant \pi_i(t_i, t_j) - p_i^0(t_i', t_j).$$

Adding the two inequalities for $i = 1, 2$ and using the fact that $p_i^0(t_i', t_j) = p_i^0(t_i, t_j')$ due to Proposition 3.2(a), and that $V_1(t_1, t_2) + V_2(t_1, t_2) = \pi^{\mathrm{M}}(t_1, t_2)$ and $p_1^0(t_1', t_2) + p_2^0(t_1', t_2) = 0$ under (BB) gives the result. □

Proof of Lemma 4.5. By assumption, there exist t_1', t_2' both arbitrarily close to t_1 and t_2, respectively, such that $(t_1', t_2), (t_1, t_2') \in M^0$. (IC) thus implies

$$\begin{aligned} V_i(t_i', t_j) = \pi_i(t_i', t_j) - p_i^0(t_j) &\geqslant s_i(t_i, t_j) \pi^{\mathrm{M}}(t_i', t_j) - p_i^1(t_i, t_j) \\ &= s_i(t_i, t_j) \pi^{\mathrm{M}}(t_i, t_j) - p_i^1(t_i, t_j) + s_i(t_i, t_j) \cdot [\pi^{\mathrm{M}}(t_i', t_j) - \pi^{\mathrm{M}}(t_i, t_j)]. \end{aligned}$$

for each $i = 1, 2$. Adding the inequalities for $i = 1, 2$ and using the fact that $p_1^0(t_2) + p_2^0(t_1) = 0$ and $[s_1(t_1, t_2) + s_2(t_2, t_1)] \cdot \pi^{\mathrm{M}}(t_1, t_2) - [p_1^1(t_1, t_2) + p_2^1(t_2, t_1)] = \pi^{\mathrm{M}}(t_1, t_2)$ under (BB), we obtain

$$\begin{aligned} \pi_1(t_1', t_2) + \pi_2(t_2', t_1) \geqslant \pi^{\mathrm{M}}(t_i, t_j) &+ s_1(t_1, t_2) \cdot [\pi^{\mathrm{M}}(t_1', t_2) - \pi^{\mathrm{M}}(t_1, t_2)] \\ &+ s_2(t_2, t_1) \cdot [\pi^{\mathrm{M}}(t_1, t_2') - \pi^{\mathrm{M}}(t_1, t_2)]. \end{aligned}$$

By continuity of the profit functions, letting $t_1' \to t_1$ and $t_2' \to t_2$ yields $\pi_1(t_1, t_2) + \pi_2(t_2, t_1) \geqslant \pi^{\mathrm{M}}(t_1, t_2)$. □

Proof of Proposition 4.6. (a) By Proposition 3.2(a), $p_i^0(t_i, t_j) = p_i^0(t_i', t_j)$ and $p_j^0(t_j, t_i') = p_j^0(t_j', t_i')$, where $j \neq i$. Under budget balance, $p_i^0(t_i', t_j) = -p_j^0(t_j, t_i')$ and $p_i^0(t_i', t_j') = -p_j^0(t_j', t_i')$, from which the claim immediately follows.

(b) Observe first that since $\frac{\partial}{\partial t_i} s_i = -\frac{\partial}{\partial t_i} s_j$ and $\frac{\partial}{\partial t_i} p_i^1 = -\frac{\partial}{\partial t_i} p_j^1$ due to (BB), and $\frac{\partial}{\partial t_i} p_i^1 = \pi^{\mathrm{M}} \cdot \frac{\partial}{\partial t_i} s_i$ due to Proposition 3.2(b), it suffices to show that $\frac{\partial}{\partial t_i} s_i = 0$ under the conditions stated.[36] To prove this, we first show that, for any $(t_1, t_2) \in M^1$, (IC) and (BB) imply

$$\frac{\partial}{\partial \tilde{t}_i} s_i(t_i, t_j) \cdot \frac{\partial}{\partial t_j} \pi^{\mathrm{M}}(t_1, t_2) + \frac{\partial}{\partial \tilde{t}_j} s_j(t_i, t_j) \cdot \frac{\partial}{\partial t_i} \pi^{\mathrm{M}}(t_1, t_2) = 0. \tag{A.4}$$

Differentiation of condition (6) in Proposition 3.2(b) with respect to t_j gives

$$\frac{\partial}{\partial \tilde{t}_j \partial \tilde{t}_i} s_i(t_i, t_j) \cdot \pi^{\mathrm{M}}(t_1, t_2) + \frac{\partial}{\partial \tilde{t}_i} s_i(t_i, t_j) \cdot \frac{\partial}{\partial t_j} \pi^{\mathrm{M}}(t_1, t_2) = \frac{\partial}{\partial \tilde{t}_j \partial \tilde{t}_i} p_i^1(t_i, t_j). \tag{A.5}$$

Swapping indices, we obtain

$$\frac{\partial}{\partial \tilde{t}_i \partial \tilde{t}_j} s_j(t_j, t_i) \cdot \pi^{\mathrm{M}}(t_1, t_2) + \frac{\partial}{\partial \tilde{t}_j} s_j(t_j, t_i) \cdot \frac{\partial}{\partial t_i} \pi^{\mathrm{M}}(t_1, t_2) = \frac{\partial}{\partial \tilde{t}_i \partial \tilde{t}_j} p_j^1(t_j, t_i). \tag{A.6}$$

Using budget balance (BB),

$$\begin{aligned}
&\frac{\partial}{\partial \tilde{t}_i \partial \tilde{t}_j} s_j(t_j, t_i) \cdot \pi^{\mathrm{M}}(t_1, t_2) - \frac{\partial}{\partial \tilde{t}_i \partial \tilde{t}_j} p_j^1(t_j, t_i) \\
&\qquad = -\left(\frac{\partial}{\partial \tilde{t}_i \partial \tilde{t}_j} s_i(t_i, t_j) \cdot \pi^{\mathrm{M}}(t_1, t_2) - \frac{\partial}{\partial \tilde{t}_i \partial \tilde{t}_j} p_i^1(t_i, t_j) \right) \\
&\qquad\qquad = -\left(\frac{\partial}{\partial \tilde{t}_j \partial \tilde{t}_i} s_i(t_i, t_j) \cdot \pi^{\mathrm{M}}(t_1, t_2) - \frac{\partial}{\partial \tilde{t}_j \partial \tilde{t}_i} p_i^1(t_i, t_j) \right).
\end{aligned}$$

Inserting this condition in (A.5) and (A.6) yields (A.4).

Now, using (A.4), the claim in Proposition 4.6(b) is immediately obvious if $\frac{\partial}{\partial t_i} \pi^{\mathrm{M}} = 0$. Hence, assume that $\frac{\partial}{\partial t_i} \pi^{\mathrm{M}} \neq 0$. Then, we may rewrite condition (A.4) as

$$\frac{\frac{\partial}{\partial \tilde{t}_i} s_i}{\frac{\partial}{\partial t_i} \pi^{\mathrm{M}}} = -\frac{\frac{\partial}{\partial \tilde{t}_j} s_j}{\frac{\partial}{\partial t_j} \pi^{\mathrm{M}}}. \tag{A.7}$$

Assume now that, in contradiction to Proposition 4.6(b), $\frac{\partial}{\partial t_i} s_i \neq 0$. Then the left-hand side of (A.7) must be strictly positive by Proposition 3.2(b). Similarly, however, the right-hand side is non-positive by Proposition 3.2(b), a contradiction. □

Proof of Proposition 4.10. (a) By Proposition 4.9, I' belongs to a level set of $s_i' \cdot \pi^{\mathrm{M}} - \pi_i$ for some $s_i' \in [0, 1]$, which can be assumed to have measure zero w.l.o.g. Since π^{M} and π_i are twice continuously differentiable, I' must then be smoothly path connected

[36] Strictly speaking, the final conclusion concerning $\frac{\partial}{\partial \tilde{t}_i} p_i$ only follows directly from Proposition 3.2(b) if $\pi^{\mathrm{M}} \neq 0$. However, if $\pi^{\mathrm{M}} = 0$, it is quickly seen that any transfer function $\frac{\partial}{\partial \tilde{t}_i} p_i \neq 0$ cannot be incentive compatible by a more basic argument paralleling that in Proposition 3.2(a) for no-merger outcomes.

and parameterizable by differentiable functions $(t_1(x), t_2(x))$ of $x \in [0,1]$. Because $s_i' \cdot \pi^{\mathrm{M}}(t_1(x), t_2(x)) - \pi_i(t_1(x), t_2(x))$ is constant in x,

$$s_i' \cdot \left(\frac{\partial \pi^{\mathrm{M}}}{\partial t_1} \frac{\partial t_1}{\partial x} + \frac{\partial \pi^{\mathrm{M}}}{\partial t_2} \frac{\partial t_2}{\partial x} \right) - \left(\frac{\partial \pi_i}{\partial t_1} \frac{\partial t_1}{\partial x} + \frac{\partial \pi_i}{\partial t_2} \frac{\partial t_2}{\partial x} \right) = s_i' \cdot \partial_{\mathbf{r}(\mathbf{t})} \pi^{\mathrm{M}} - \partial_{\mathbf{r}(\mathbf{t})} \pi_i = 0. \quad \text{(A.8)}$$

(a1) and (a2) immediately follow.[37] (a3) is an immediate implication of (10) and (A.8). To see (a3), recall from Proposition 3.4(b) that for both agents i, $s_i' \cdot \pi^{\mathrm{M}} - \pi_i$ must be increasing in t_i in the direction in which m is increasing. We know furthermore from Lemma 4.4 that m must be locally increasing with $\pi^{\mathrm{M}} - \pi_1 - \pi_2$, i.e. it must be increasing in the direction $\mathbf{q}(\mathbf{t})$. Combining, this implies that on $\tilde{I}(m)$, $s_i \cdot \pi^{\mathrm{M}} - \pi_i$ must be increasing in the direction $\mathbf{q}(\mathbf{t})$, which is stated in (a3) using $s_i(\mathbf{t}) = \partial_{\mathbf{r}(\mathbf{t})}\pi_i / \partial_{\mathbf{r}(\mathbf{t})}\pi^{\mathrm{M}}$ by our previous results to substitute for s_i.

(b) The proof of part (b) is analogous. As $I_i(m) \setminus \tilde{I}(m)$ is perpendicular to the t_i-axis, the parametrization $t_1(x), t_2(x)$ satisfies $\frac{\partial}{\partial x} t_i = 0$. (A.8) thus becomes

$$s_i' \cdot \left(\frac{\partial \pi^{\mathrm{M}}}{\partial t_j} \frac{\partial t_j}{\partial x} \right) - \frac{\partial \pi_i}{\partial t_j} \frac{\partial t_j}{\partial x} = 0, \quad \text{(A.9)}$$

which implies (b1)–(b2). (b3) again follows from Proposition 3.4(b) because $s_i(\mathbf{t}) = \frac{\partial}{\partial t_j}\pi_i / \frac{\partial}{\partial t_j}\pi^{\mathrm{M}}$ by (A.9). □

Proof of Proposition 6.5. (a) Because $\mathbf{r}$ points into M^1, $\partial_{\mathbf{r}} V_i^1 \geqslant \partial_{\mathbf{r}} \pi_i$ by Lemma 6.3. Moreover, $\partial_{\mathbf{r}} \pi^{\mathrm{M}} = 0$ implies $\partial_{\mathbf{r}} V_i^1 = 0$ by Lemma 6.4. Thus, $\partial_{\mathbf{r}} \pi_i \leqslant 0$, as was to be shown.

(b) If the subset of the boundary $I_i(m)$ is not kinked, then it is only possible for an $\mathbf{r}$ as specified in part (a) not to exist if the level curve of π^{M} in $\mathbf{t}$ is tangent to the boundary. But then V_i must be locally constant on the boundary by Lemma 6.4, which in turn is compatible with the required continuity of V_i *across* the boundary only if π_i (the value of V_i over M^0) is locally constant on the boundary as well. □

Proof of Corollary 6.6. We first show that at any $\mathbf{t} \in T_1 \times T_2$ and $\mathbf{r} \in \mathbb{R}^2$ such that $\partial_{\mathbf{r}} \pi^{\mathrm{M}}(\mathbf{t}) = 0$, we cannot have $\partial_{\mathbf{r}} \pi_i(\mathbf{t}) \leqslant 0$ for both $i = 1, 2$. To see this, let $\mathbf{r} = (r_1, r_2)$, and note that $r_1 \cdot r_2 \leqslant 0$ by assumption (iii) in the corollary. Then,

$$\begin{aligned} \partial_{\mathbf{r}} \pi_1 \cdot \partial_{\mathbf{r}} \pi_2 &= \left(r_1 \cdot \tfrac{\partial}{\partial t_1} \pi_1 + r_2 \cdot \tfrac{\partial}{\partial t_2} \pi_1 \right) \cdot \left(r_1 \cdot \tfrac{\partial}{\partial t_1} \pi_2 + r_2 \cdot \tfrac{\partial}{\partial t_2} \pi_2 \right) \\ &= r_1^2 \cdot \tfrac{\partial}{\partial t_1} \pi_1 \cdot \tfrac{\partial}{\partial t_1} \pi_2 + r_2^2 \cdot \tfrac{\partial}{\partial t_2} \pi_1 \cdot \tfrac{\partial}{\partial t_2} \pi_2 + r_1 \cdot r_2 \cdot \left(\tfrac{\partial}{\partial t_1} \pi_1 \cdot \tfrac{\partial}{\partial t_2} \pi_2 + \tfrac{\partial}{\partial t_2} \pi_1 \cdot \tfrac{\partial}{\partial t_1} \pi_2 \right), \end{aligned}$$

which is quickly checked to be strictly negative by assumptions (i)—(iii). Thus, $\partial_{\mathbf{r}} \pi_1 \cdot \partial_{\mathbf{r}} \pi_2 < 0$, implying that we cannot have $\partial_{\mathbf{r}} \pi_i(\mathbf{t}) \leqslant 0$ for both $i = 1, 2$.

Hence, by Proposition 6.5, $\tilde{I}(m) = I_1(m) \cup I_2(m)$ must either be empty or contain only points $\mathbf{t}$ where the level curves of π^{M}, π_1, and π_2 all have identical slopes, which in turn is trivially ruled out by assumptions (i)—(iii). □

[37] The formulation of part (a) implicitly assumes w.l.o.g. that $\partial_{\mathbf{r}(\mathbf{t})} \pi^{\mathrm{M}} \neq 0$. Note that if $\partial_{\mathbf{r}(\mathbf{t})} \pi^{\mathrm{M}} = 0$, then (A.8) implies $\partial_{\mathbf{r}(\mathbf{t})} \pi_i = 0$, so that $\nabla \pi^{\mathrm{M}}$ and $\nabla \pi_i$ would need to be collinear to begin with.

Proof of Proposition 7.1. We first show that, as claimed, a mechanism with $s_1 = 1$, $p_1^1 = -p_1^1 = \pi_2$, $p_1^0 = p_2^0 = 0$ will implement any efficient decision. Under such a mechanism, agent 2 will always get a payoff of $U_2(\tilde{t}_2; t_2, t_1) = \pi_2$ no matter what signal $\tilde{t}_2$ he sends, so that truth-telling will always be weakly optimal. Concerning agent 1, observe that transfers and shares are unaffected by any deviations which leave the merger decision unaffected, so that we only need to show incentive compatibility for deviations which affect the merger decision. Given any (t_1, t_2), agent 1's payoff will be $\pi^{\mathrm{M}}(t_1, t_2) - \pi_2$ from merging, and $\pi_1(t_1, t_2)$ from not merging. Hence, agent 1's gains to merging are $\pi^{\mathrm{M}}(t_1, t_2) - \pi_1(t_1, t_2) - \pi_2$, which will be positive for any type constellation such that merging is efficient, and negative for any type constellation such that merging is inefficient. Hence, under any efficient mechanism, a type such that merging is efficient (given the other type) will loose from deviating from truth-telling, and so will a type such that no merger occurs. It is trivially checked that the described mechanism satisfies (BB) and (IR).

To prove uniqueness, consider any efficient merger decision m. Since $\nabla\pi^{\mathrm{M}}$ and $\nabla\pi_1$ are by assumption not collinear on I^*, Proposition 4.10(a) is applicable with

$$\mathbf{r}(\mathbf{t}) = a \cdot [\tfrac{\partial}{\partial t_2}\pi^{\mathrm{M}} - \tfrac{\partial}{\partial t_2}\pi_1, -(\tfrac{\partial}{\partial t_1}\pi^{\mathrm{M}} - \tfrac{\partial}{\partial t_1}\pi_1)]^{\mathrm{T}}$$

for some $a \neq 0$. Calculating the directional derivatives of π^{M} and π_1 on $I(m) \subseteq I^*$ along $\mathbf{r}(\mathbf{t})$ then immediately shows that

$$\partial_{\mathbf{r}(\mathbf{t})}\pi^{\mathrm{M}} = \mathbf{r}(\mathbf{t})^{\mathrm{T}}\nabla\pi^{\mathrm{M}} = a \cdot [-\tfrac{\partial}{\partial t_2}\pi_1 \tfrac{\partial}{\partial t_1}\pi^{\mathrm{M}} + \tfrac{\partial}{\partial t_1}\pi_1 \tfrac{\partial}{\partial t_2}\pi^{\mathrm{M}}] = \mathbf{r}(\mathbf{t})^{\mathrm{T}}\nabla\pi_1 = \partial_{\mathbf{r}(\mathbf{t})}\pi_1,$$

so that $\partial_{\mathbf{r}(\mathbf{t})}\pi^{\mathrm{M}} = \partial_{\mathbf{r}(\mathbf{t})}\pi_1 \neq \mathbf{0}$, where the latter inequality again follows from $\nabla\pi^{\mathrm{M}}$ and $\nabla\pi_1$ not being collinear. Hence, we must have $s_1(t_1, t_2) = 1$ on $I(m) \subseteq I^*$, and thereby on M^1 by Proposition 4.6. □

References

Agrawal, Anup, Jeffrey F. Jaffe, and Gershon N. Mandelker (1992) 'The post-merger performance of acquiring firms: A re-examination of an anomaly.' *Journal of Finance* 46(4), 1605–1621

Ausubel, Lawrence R. (2004) 'An efficient ascending-bid auction for multiple objects.' *American Economic Review* 94(5), 1452–1475

Banal-Estañol, Albert, and Jo Seldeslachts (2005) 'Merger failures.' Working Paper, Wissenschaftszentrum Berlin

Barros, Pedro Pita (1998) 'Endogenous mergers and size asymmetry of merger participants.' *Economics Letters* 60(1), 113–119

Bergemann, Dirk, and Juuso Välimäki (2002) ‘Information acquisition and efficient mechanism design.’ *Econometrica* 70(3), 1007–1033

Bergemann, Dirk, and Stephen Morris (2005) ‘Robust mechanism design.’ *Econometrica* 73(6), 1771–1813

Borek, Thomas, Stefan Buehler, and Armin Schmutzler (2004) ‘Mergers under asymmetric information—Is there a lemons problem?’ Working Paper, University of Zürich

Brusco, Sandro, Giuseppe Lopomo, David T. Robinson, and S. Viswanathan (2005) ‘Efficient mechanisms for mergers and acquisitions.’ Mimeo, Duke University

Chung, Kim-Sau, and Jeffrey C. Ely (2003) ‘Ex-post incentive compatible mechanism design.’ Mimeo, Northwestern University

Crémer, Jacques (1987) ‘Auctions with contingent payments: Comment.’ *American Economic Review* 77(4), 746

Crémer, Jacques, and Richard P. McLean (1985) ‘Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent.’ *Econometrica* 53(2), 345–361

Dasgupta, Partha, and Eric Maskin (2000) ‘Efficient auctions.’ *Quarterly Journal of Economics* 115(2), 341–388

Eső, Péter, and Eric Maskin (2000) ‘Multi-good efficient auctions with multidimensional information.’ mimeo, Harvard University

Fieseler, Karsten, Thomas Kittsteiner, and Benny Moldovanu (2003) ‘Partnerships, lemons, and efficient trade.’ *Journal of Economic Theory* 113(2), 223–234

Green, Jerry R., and Jean-Jacques Laffont (1977) ‘Characterization of satisfactory mechanisms for the revelation of preferences for public goods.’ *Econometrica* 45(2), 427–438

— (1979) *Incentives in Public Decision Making* (Amsterdam: North-Holland)

Gugler, Klaus, Dennis C. Mueller, B. Burcin Yurtoglu, and Christine Zulehner (2003) ‘The effects of mergers: An international comparison.’ *International Journal of Industrial Organization* 21(5), 625–653

Hansen, Robert G. (1985) ‘Auctions with contingent payments.’ *American Economic Review* 75(4), 862–865

Healy, Paul M., Krishna G. Palepu, and Richard S. Ruback (1992) ‘Does corporate performance improve after mergers?’ *Journal of Financial Economics* 31(2), 135–175

Hietala, Pekka, Steven N. Kaplan, and David T. Robinson (2003) 'What is the price of hubris? Using takeover battles to infer overpayments and synergies.' *Financial Management* 32(3), 5–31

Holmstrom, Bengt (1977) 'On incentives and control in organizations.' PhD dissertation, Stanford University

Jehiel, Philippe, and Ady Pauzner (2003) 'Partnership dissolution with interdependent valuations.' Mimeo, Tel Aviv University, forthcoming in *RAND Journal of Economics*

Jehiel, Philippe, and Benny Moldovanu (2001) 'Efficient design with interdependent valuations.' *Econometrica* 69(5), 1237–1259

Jehiel, Philippe, Moritz Meyer-ter-Vehn, Benny Moldovanu, and William R. Zame (2005) 'The limits of ex-post implementation.' Discussion Paper, University of Bonn

Jensen, Michael C. (1986) 'Agency costs of free cash flow, corporate finance, and takeovers.' *American Economic Review* 76(2), 323–329

Krishna, Vijay (2003) 'Asymmetric english auctions.' *Journal of Economic Theory* 112(2), 261–288

Larsson, Rikard, and Sidney Finkelstein (1999) 'Integrating strategic, organizational, and human resource perspectives on mergers and acquisitions: A case study of synergy realization.' *Organizational Science* 10(1), 1–26

Maskin, Eric (1992) 'Auctions and privatization.' In *Privatization,* ed. Horst Siebert (Institut für Weltwirtschaft an der Universität Kiel) pp. 115–136

Mezzetti, Claudio (2004) 'Mechanism design with interdependent valuations: Efficiency.' *Econometrica* 72(5), 1617–1626

Milgrom, Paul (2004) *Putting Auction Theory to Work* (Cambridge: Cambridge University Press)

Moldovanu, Benny (2002) 'How to dissolve a partnership.' *Journal of Institutional and Theoretical Economics* 158(1), 66–80

Myerson, Roger B., and Mark A. Satterthwaite (1983) 'Efficient mechanisms for bilateral trading.' *Journal of Economic Theory* 29(2), 265–281

Officer, Micah (2004) 'Collars in merger negotations.' *Journal of Finance* 59(6), 2719–2743

Ornelas, Emanuel, and John L. Turner (2004) 'Efficient dissolutions of parterships and the structure of control.' Mimeo, University of Georgia

Perry, Motty, and Philip Reny (2002) 'An efficient auction.' *Econometrica* 70(3), 1199–1212

Ravenscraft, David J., and F. M. Scherer (1989) 'The profitability of mergers.' *International Journal of Industrial Organization* 7(1), 101–116

Roberts, Kevin W. (1979) 'The characterization of implementable choice rules.' In *Aggregation and Revelation of Preferences,* ed. Jean-Jacques Laffont (Amsterdam: North-Holland) pp. 321–348

Salant, Stephen W., Sheldon Switzer, and Robert J. Reynolds (1983) 'Losses from horizontal merger: The effects of an exogenous change in industry structure on cournot-nash equilibrium.' *Quarterly Journal of Economics* 98(2), 185–199

Samuelson, William (1987) 'Auctions with contingent payments: Comment.' *American Economic Review* 77(4), 740–745

Shleifer, Andrei, and Robert W. Vishny (1988) 'Value maximization and the acquisition process.' *Journal of Economic Perspectives* 2(1), 7–20

Walker, Mark (1980) 'On the nonexistence of a dominant strategy mechanism for making optimal public decisions.' *Econometrica* 48(6), 1521–1540

Chapter 3

Are There Waves in Merger Activity After All?

Dennis Gärtner and Daniel Halbheer

1 Introduction

There is broad consensus that mergers occur in waves. Since the seminal contribution by Nelson (1959), many studies have reported a wave-like pattern in merger activity, pointing out the merger waves of the mid 1980s and mid 1990s in the US in particular.[1] Guided by these observations, a vast empirical literature has sought to identify potential causes and triggers for merger waves.[2] This empirical strand has more recently been complemented by efforts to explain the phenomenon of merger waves in the theoretical literature.[3]

While the general notion of mergers occurring in waves is practically undisputed, there is no clear consensus on how to operationalize the concept of a 'merger wave' in a time series context. The empirical literature has put forward three distinct approaches to modeling and identifying such waves. First, Golbe and White (1993) have sought to identify waves by fitting a sine curve to historic merger data. Second, merger series have been modeled by autoregressive processes capable of producing

[1]Studies discussing the 1980s merger wave include Ravenscraft (1987), Golbe and White (1988, 1993), and Mitchell and Mulherin (1996). Mueller (1997), Andrade et al. (2001), and Harford (2005) describe the 1990s merger wave. An extensive review of earlier merger waves is provided in Scherer and Ross (1990, pp. 154–159).

[2]See e.g. Ravenscraft (1987), Shleifer and Vishny (1990), and Holmstrom and Kaplan (2001). Gugler et al. (2005) examine hypotheses that have been put forward as explanations of merger waves.

[3]Examples include Fauli-Oller (2000), Jovanovic and Rousseau (2002), Rhodes-Kropf and Viswanathan (2004), and Toxvaerd (2004).

wave-like behavior (Shughart and Tollison, 1984; Clark et al., 1988; Chowdhury, 1993; Barkoulas et al., 2001). Third, and finally, merger series have been modeled by means of parameter-switching models where waves in activity are caused by discrete parameter switches (Town, 1992, Linn and Zhu, 1997).

This paper reexamines the case for detecting waves in merger activity in a time series context using more recent, consistent data and refined estimation techniques. Following Town (1992) and Linn and Zhu (1997), we employ a Markov regime switching model to describe the stochastic behavior of merger activity. We provide a thorough motivation for this approach, starting from Nelson's (1959, p. 126) observation that aggregate merger series are characterized by "large bursts of activity separated by lengthy intervals of very low activity," which we take to suggest the presence of two distinct unobserved states of merger activity, 'high' and 'low'. By letting mean and variance of the autoregressive model be determined by realizations of the Markov process governing the evolution of the two states, waves are triggered by switches in the unobserved state. While this approach borrows from Town (1992) and Linn and Zhu (1997), we propose a slightly modified formal specification in which the autoregressive processes' inertia persists also *across* state switches, leading merger activity to react less abruptly to such switches. More importantly, we use new and consistent quarterly time series data covering merger activity both in the US and the UK, extending from 1973:IV through 2003:IV and from 1969:I through 2003:IV, respectively.[4]

In this paper, we challenge the notion of the much-discussed 1980s merger wave in the US. We argue that the discrepancy between our findings and previous econometric identifications of this wave is driven by a further distinguishing feature of our analysis: the use of more recent estimation techniques. To address the central issue of wave identification, we conduct inference on the regime indicator within a Bayesian framework employing Gibbs sampling techniques (Gelfand and Smith, 1990; Albert and Chib, 1993). In contrast, the aforementioned studies by Town (1992) and Linn and Zhu (1997) base wave identification on Maximum Likelihood techniques (Hamilton, 1989, 1993). In this latter approach, inference consists in first estimating the model's unknown parameters via Maximum Likelihood, and then conducting inference on the unobserved state conditional on the parameter estimates. Bayesian analysis, on the other hand, avoids this two-step procedure by treating both the model parameters and state variable as random variables and basing inference on states on a joint distribution of parameters and states rather than on a conditional distribution. This methodological difference can lead to quite different conclusions regarding the likely path of the unobserved regime indicator if parameter uncertainty is sufficiently high, as the uncertainty on parameter estimates does not

[4] Previous empirical studies examined the merger wave hypothesis using an assemblage of separate series differing in coverage and inclusion criteria. For a general discussion of available historical time series merger data and their limitations, see e.g. Golbe and White (1988).

feed into uncertainty on states when employing a two-step estimation procedure.

Our main results are as follows. *First*, we find that the US have witnessed only the beginning of a wave in merger activity, this wave starting in 1995:IV. This result is consistent with the observations in Mueller (1997), Holmstrom and Kaplan (2001), and Andrade et al. (2001), all of which report an upsurge in merger activity in the mid 1990s. However, our investigation of industry level data does not support the prominent notion that waves in aggregate merger activity represent the clustering of surges within one or a few industries (Mitchell and Mulherin, 1996; Mulherin and Boone, 2000; Andrade et al., 2001; Harford, 2005). *Second*, even when fitting the model only to the data prior to the estimated break date, we fail to identify the much discussed 1980s merger wave. To explain our difference in findings, we argue that if there is sufficient uncertainty surrounding the model's parameters, then the two-step Maximum Likelihood estimation procedure can convey a deceptive degree of certainty about state inference. *Third*, the UK has witnessed two merger waves, the first starting in 1971:I and ending in 1973:IV and the second lasting from 1986:III to 1989:IV. The dating of these merger waves is close to the evidence reported in Hughes (1993, p. 16).

The remainder of the paper is organized as follows. Section 2 briefly comments on the data employed. Section 3 provides a thorough motivation of the model. In Section 4, we describe the inference problem and give a brief introduction to the Gibbs sampling approach. Section 5 presents the main results of our estimation both for the US and the UK series. Section 6 concludes.

2 The Data

Our paper follows the majority of previous studies, particularly Town (1992) and Linn and Zhu (1997), in using the number of transactions as the measure of historical merger activity.[5] Specifically, we investigate the following series:

(i) *The US merger series, covering 1973:I–2003:IV.* The time series data are taken from various issues of *Mergerstat Review*, a publication by FactSet Mergerstat LLC. The series reports publicly announced mergers, acquisitions and unit divestitures involving (i) at least one US company, (ii) a transaction volume exceeding $1 million, and (iii) a purchase price exceeding 10% of the acquired company's equity (i.e., an interest exceeding 10% of the acquired firm's equity).

(ii) *The UK merger series, covering 1969:I–2003:IV.* These data are published by

[5]Other prominent measures of merger activity suggested in the literature are the dollar value of merger transactions (see e.g. Golbe and White, 1988, Scherer and Ross (1990) and the number of transactions relative to appropriate population totals (see e.g. Hughes, 1993, p. 16; Gugler et al., 2005).

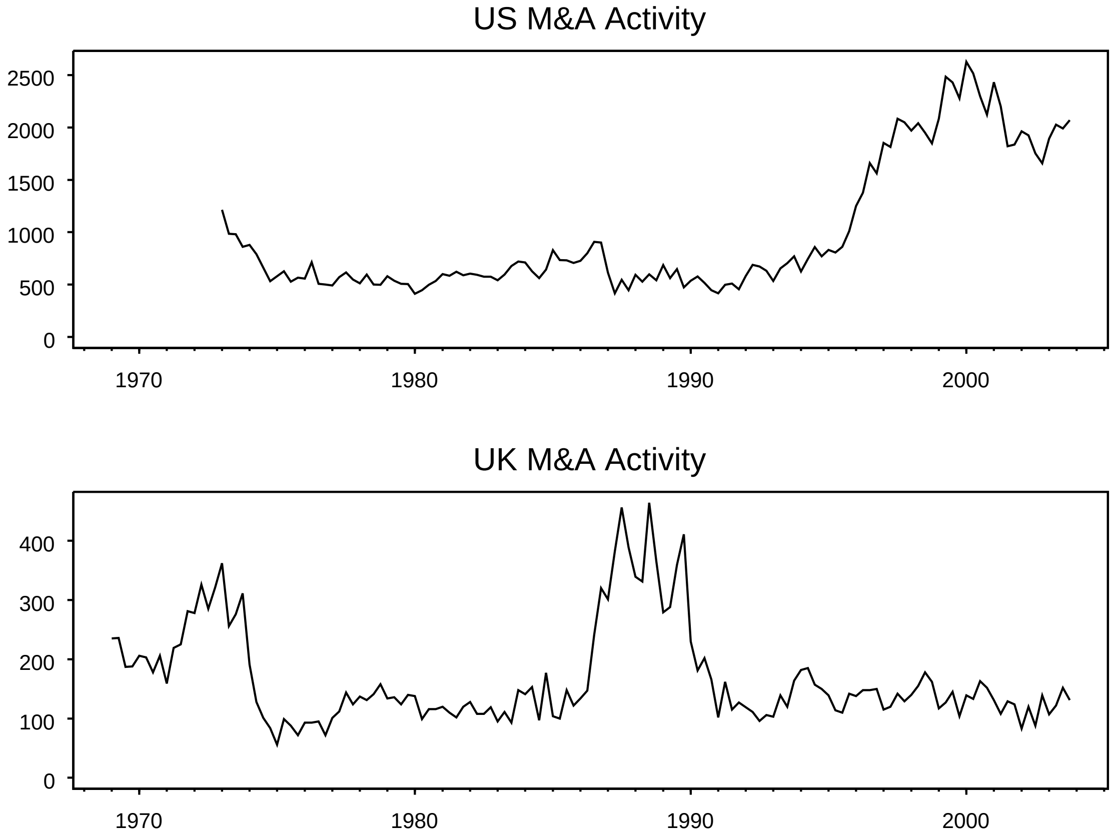


Figure 1: The Number of Merger Transactions, US Series, 1973:I–2003:IV (in Top Panel) and UK Series, 1969:I–2003:IV (Bottom Panel).

the Office for National Statistics on a quarterly basis. The series consists of publicly announced mergers and acquisitions involving UK companies only. In contrast to the US data, there is no explicit cut-off bias relating to the value of the transaction, but the deal has to aim at gaining de jure control of the acquired company (i.e., a controlling interest exceeding 50% of the acquired firm's equity).

Plots of these series are presented in Figure 1.

3 A Markov-Switching Model of Merger Waves

As outlined above, the literature has advanced the idea that mergers follow a wave pattern. We take this casual impression to suggest the presence of two distinct states of merger activity, high and low, as follows:[6]

Assumption 3.1. *Each period t is associated with an unobserved latent state variable $S_t \in \{1, 2\}$, where $S_t = 1$ implies that period t is a* low-activity period *and $S_t = 2$ denotes a* high-activity period.

The basic idea is then to let unobserved switches between states of high and low activity feed into observed merger activity—in a sense to be made precise shortly—so as to induce the alleged wave-like behavior. Hence, given Assumption 3.1, the remaining key questions concerning our description of mergers are: (1) What determines the unobservable state S_t in any period t, and (2) how exactly do the unobservable states feed into observed merger activity y_t?

The general framework in which we deal with these questions is the Markov regime switching model originally proposed by Hamilton (1989). In a nutshell, this approach treats both the sequence of observations y_t and the sequence of states S_t as (interdependent) random variables, specifies a model which jointly generates the two sequences, and then estimates the model using the observed series y_t while treating the sequence of states as 'missing data'. This framework offers several advantages over more traditional approaches to break-point analysis which typically rely on casual determination of candidate break-dates or *ad hoc* restrictions on the number of break dates (see e.g. Chow, 1960, 1984; Andrews, 1993): First and foremost, a major goal of our analysis is not only the estimation of regime-dependent structural model parameters, but dating the waves (i.e. conducting inference on the break dates themselves). This in turn requires modelling the probability law governing changes in regime rather than imposing particular break-dates *a priori*. Through the probability model, we can then let the data itself speak about the likely incidence of significant changes. Second, we would like to propose a unified structural process capable of describing various merger series (such as across countries or industries) with apparently different frequency and timing of waves, which requires that wave dates be determined endogenously by the process.

More specifically, concerning the determination of states, we shall assume that states follow an independent first-order Markov process. Thus, in any period t, the probability of switching to a certain state in the next period $t + 1$ depends only on the state in period t. Specifically, we assume the following:

Assumption 3.2. *The unobserved state variable S_t follows a first-order Markov process with transition probabilities from any period t to period $t + 1$ given by*

$$\Pr(S_{t+1} = 1|S_t = 1) = p_{11} \qquad \textit{and} \qquad \Pr(S_{t+1} = 2|S_t = 2) = p_{22}, \tag{1}$$

[6]We shall comment on the idea of using more than two states when discussing our estimation results further below. Let us just note for now that rasing the number of attainable states invokes the usual trade-off between achieving a better fit to the data and overparameterizing the model. As a consequence, we suggest using the minimal number of states capable of producing the described behavior in mergers.

with $p_{11}, p_{22} \in [0,1]$. *In any period* t*, these transition probabilities are independent of past (log) merger realizations* $(y_t, y_{t-1}, \ldots)$.

It is important to note that 'merger waves' as we understand and model them need not display a highly regular periodic pattern. Indeed, the first-order Markov specification implies that the process governing the states displays very little memory. This low-memory approach seems justified by the aforementioned literature giving little impression that the documented bursts of high activity display a highly regular periodic pattern.[7] Some structure is of course nonetheless implied by our Markov specification, such as the expected duration of a high state being $p_{22}/(1-p_{22})$ and the expected duration of a low state being $p_{11}/(1-p_{11})$, but these durations generally display a rather high variability. Furthermore, due to the first order Markov property, the *remaining* expected duration of a certain state is independent of how long the process has already been in that state, which again reflects the low-memory quality of the process.

Finally, note that the Markovian model encompasses the extreme possibility of a state being 'absorbing' in the sense that, once the process reaches a certain state, it remains in that state indefinitely (so that the regime switch is permanent rather than transitory). This is the case for the low-activity state if $p_{11} = 1$ and for the high-activity state if $p_{22} = 1$. Conversely, whenever this is not the case, so $p_{11}, p_{22} < 1$ and if in addition $p_{11} + p_{22} > 0$ (so there is no completely deterministic alternation between states), then the Markov chain turns out to be *ergodic* (see e.g. Hamilton, 1989). Then, a further key characteristic of the state switching model is given by the ergodic regime probabilities $\Pr(S_t = i)$, i.e. the unconditional probability of state $i \in \{1, 2\}$.[8] These can be shown to be given by $\Pr(S_t = 1) = (1-p_{11})/(2-p_{11}-p_{22})$ and $\Pr(S_t = 2) = 1 - \Pr(S_t = 1)$.

Let us now turn to the second question concerning the exact form of the state's impact on merger activity. We start from the idea that merger activity follows some sort of mean reverting autoregressive process and augment this by assuming that both the mean and the variance of this process are time-varying and dependent on the states.[9] Specifically, we make the following assumption:

[7] Although Golbe and White (1993) do report evidence of a sine wave pattern in US merger activity based on data up to 1989, by inspection of the plots in Figure 1 we strongly suspect that their model would no longer provide a very good fit to our more recent series.

[8] A convenient way to think of the ergodic probabilities is in terms of the fraction of high and low states observed in an infinitely long realization of the Markov chain.

[9] A previous study by Shughart and Tollison (1984) reports little success in describing waves in merger activity as a standard autoregressive process with *constant* mean and variance, $y_t - \mu = \sum_{i=1}^{k} \phi_i (y_{t-i} - \mu) + \varepsilon_t$ with $\varepsilon_t \sim$ i.i.d. $\mathrm{N}(0, \sigma^2)$, where the wave property would be reflected solely by some of the higher-order autoregressive coefficients $\phi_2, \ldots, \phi_k$ being nonzero. However, such a specification can produce only rather 'tame', linear wave-like oscillations, while we suspect that the large bursts of activity separated by long intervals of low activity identified in the aforementioned literature can only be reconciled with a nonlinear model such as ours.

Assumption 3.3. *Conditional on the sequence of unobserved states S_t, (log) mergers y_t follow the* AR(k) *process*

$$y_t - \mu_{S_t} = \sum_{i=1}^{k} \phi_i (y_{t-i} - \mu_{S_t}) + \varepsilon_t, \tag{2}$$

where (i) the ε_t are independently $\mathrm{N}(0, \sigma_{S_t}^2)$ *and independent of previous merger realizations $(y_{t-1}, y_{t-2}, \ldots)$, (ii) $\mu_{S_t} \in \{\mu_1, \mu_2\}$ and $\sigma_{S_t} \in \{\sigma_1, \sigma_2\}$ are determined by the state in period t, (iii) $\mu_2 \geqslant \mu_1$, and (iv) the autoregressive coefficients $\phi_1, \phi_2, \ldots, \phi_k$ are restricted so that the roots of the associated lag polynomial, $\phi(L) \equiv 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_k L^k$, lie outside the complex unit circle.*

We let the idea that $S_t = 2$ entails higher activity impose the normalization $\mu_2 \geqslant \mu_1$, while σ_1 and σ_2 are left unrestricted (except for the obvious nonnegativity requirement). Furthermore, the familiar condition on the autoregressive parameters $\phi_1, \phi_2, \ldots, \phi_k$ ensures that the process is in some sense mean reverting, where this mean however depends on the state sequence. Put somewhat differently, the condition ensures that the only source of non-stationarity is through switches in the regime (i.e. given a constant sequence of states, $S_t = i$ for all t and $i \in \{1, 2\}$, y_t follows a stationary process).[10]

Specification (2) differs in a small but important way from Hamilton's (1989) original specification,

$$y_t - \mu_{S_t} = \sum_{i=1}^{k} \phi_i (y_{t-i} - \mu_{S_{t-i}}) + \varepsilon_t, \tag{3}$$

which has been the workhorse model in the literature on mean and variance switching Markov models and also happens to be the model used by Town (1992) and Linn and Zhu (1997) to describe mergers in particular. The subtle but important difference is that specification (2) assumes 'sluggish' adjustments of the merger series to a state switch, whereas by specification (3), state switches cause an immediate full shift in activity. To see this, observe that in specification (2), what systematically affects today's deviation from the mean, $y_t - \mu_{S_t}$, is a weighted sum of past deviations from the *current* mean, $y_{t-i} - \mu_{S_t}$, whereas in specification (3), today's deviation is determined by past deviations from the *contemporaneous* mean, $y_{t-i} - \mu_{S_{t-i}}$. Thus,

[10] The literature has also proposed non-mean reverting processes such as random walks to describe merger activity (see e.g. Chowdhury, 1993). Even though standard tests reject the unit root hypothesis for our UK merger series, this is indeed not the case for the US series. However, it is a well understood fact that in general, unit-root tests have very little power over Markov-switching alternatives (see e.g. Nelson et al., 2001), so that such tests do not invalidate our proposed model. Furthermore, if we perform the unit root tests using only US data prior to 1995:III (which amounts to discarding little more than a quarter of the data), the unit root hypothesis is clearly rejected. We take this as evidence favoring our Markov-switching model over the random walk hypothesis.

the two models imply rather different dynamic consequences of a shift in regime. This is most effectively illustrated by setting $k = 1$ in both (2) and (3) above and considering a permanent shift from state 1 to state 2 between dates t and $t+1$. According to specification (3), the switch to state 2 at date t raises the value of any subsequent y_{t+j} ($j \geqslant 0$) by $\mu_2 - \mu_1$ over its respective value if no state-switch had occurred. In model (2), on the other hand, the impact of the state switch at t only raises subsequent y_{t+j} by $(1 - \phi_1^j)(\mu_2 - \mu_1) \leqslant \mu_2 - \mu_1$ for any $j \geqslant 0$.[11] Hence, model (3) suggests that the merger series immediately jumps toward the new mean after a state switch, whereas model (2) describes a more gradual, 'sluggish' gravitation toward the new mean. Note however that the difference of the state switch's impact between the models disappears as j rises, so that the models differ most markedly during the adjustment period.

We favor specification (2) over (3) for two somewhat interrelated reasons. First, casual inspection of real merger data suggests that the transition to a significantly higher (or lower) level of merger activity is indeed sluggish rather than immediate. Second, perhaps contrary to other common applications of mean switching models, there seems to be no intuitive reason to suggest that the merger process does not display the same amount of inertia when switching to a high or low activity state as within a given state. Indeed, if for instance we suspect the sluggishness in merger series to be a consequence of the fact that real world mergers may take considerable time to process (due to preparation, approval, etc.), thereby causing sluggish adjustment to any unobserved structural shocks, then this sluggishness should persist also when the economy moves to a generally higher or lower level of activity (i.e. when it is hit by a 'large' shock).[12] For these reasons, we shall employ specification (2) for the remainder of our analysis.[13]

As a final remark, we should point out that more generally, mean and variance switching is not the only way in which high and low activity states may be thought to affect mergers. For instance, an alternative specification might have states impact only the growth rate rather than the mean level of the merger series.[14] However, we view the mean-switching specification as closest in spirit to the wave notion developed in the literature. What may nonetheless seem somewhat extreme at first sight is that our mean switching model appears to posit that waves always have the

[11] Recall that in an AR(1) model, $|\phi_1| < 1$ by the restriction on the autoregressive coefficients in Assumption 3.3.

[12] This argument can be formalized by noting that specification (2) can be interpreted as a standard AR(k) model where the state only affects the distribution of the error term. This can be seen by rewriting model (2) as $y_t = \sum_{i=1}^{k} \phi_i y_{t-i} + \tilde{\varepsilon}_t$, $\tilde{\varepsilon}_t \sim N(\tilde{\mu}_{S_t}, \sigma^2_{S_t})$, where $\tilde{\mu}_{S_t} = \big(1 - \sum_{i=1}^{k} \phi_i\big)\mu_{S_t}$.

[13] As shown in the Appendix, a nice technical side-effect of using specification (2) is that inference on states does not involve an approximation (see e.g. Kim and Nelson, 1999, pp. 68–70).

[14] For instance, Owen (2004) pursues such an idea for UK mergers by fitting a three-state mean switching Markov model to the differenced level data $\Delta y_t \equiv y_t - y_{t-1}$.

same magnitude (or, stated differently, the described 'bursts of activity' always have the same magnitude).[15] However, we would like to argue that empirically, a major task in identifying waves is being able to tell actual waves from smaller 'ripples', and a straightforward way to accomplish this is to posit that waves always have a certain height. We will return to this point in our discussion of the 80s merger wave in Section 5.3.

4 Estimation Techniques

We estimate the model parameters and the path of the latent Markov switching regime indicator within a Bayesian framework employing Markov chain Monte Carlo simulation methods. Letting $\boldsymbol{\beta} \equiv (\mu_1, \mu_2, \sigma_1, \sigma_2, \phi_1, \phi_2, \ldots, \phi_k, p, q)$ denote the model's parameters, letting $\mathcal{Y}_T = (y_1, y_2, \ldots, y_T)$ denote the data observed, and letting $\mathcal{S}_T \equiv (S_1, S_2, \ldots, S_T)$ denote the unobserved sequence of states, Bayesian inference in our model takes the form of using the data $\mathcal{Y}_T$ and the model specified in Section 3 to map a given prior distribution of parameters, $p(\boldsymbol{\beta})$, into a joint posterior distribution of states and parameters, $p(\mathcal{S}_T, \boldsymbol{\beta}|\mathcal{Y}_T)$.

Rather than investigating $p(\mathcal{S}_T, \boldsymbol{\beta}|\mathcal{Y}_T)$ analytically, Markov chain Monte Carlo methods provide a simple way of simulating draws from this distribution. We use a particular form of these methods, the Gibbs sampling technique, which is an iterative scheme based on simulating successive draws from the conditional posterior distributions of the state vector $\mathcal{S}_T$ and the appropriately partitioned parameter vector $\boldsymbol{\beta}$:

(i) $p(\mathcal{S}_T|\boldsymbol{\beta}, \mathcal{Y}_T)$

(ii) $p(\mathbf{p}|\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\phi}, \mathcal{S}_T, \mathcal{Y}_T)$

(iii) $p(\boldsymbol{\mu}|\boldsymbol{\sigma}, \boldsymbol{\phi}, \mathbf{p}, \mathcal{S}_T, \mathcal{Y}_T)$

(iv) $p(\boldsymbol{\sigma}|\boldsymbol{\mu}, \boldsymbol{\phi}, \mathbf{p}, \mathcal{S}_T, \mathcal{Y}_T)$

(v) $p(\boldsymbol{\phi}|\boldsymbol{\mu}, \boldsymbol{\sigma}, \mathbf{p}, \mathcal{S}_T, \mathcal{Y}_T)$,

where $\boldsymbol{\mu} \equiv (\mu_1, \mu_2)$, $\boldsymbol{\sigma} \equiv (\sigma_1, \sigma_2)$, $\boldsymbol{\phi} \equiv (\phi_1, \phi_2, \ldots, \phi_k)$, and $\mathbf{p} \equiv (p_{11}, p_{22})$.[16]

In contrast to the *full* posterior $p(\mathcal{S}_T, \boldsymbol{\beta}|\mathcal{Y}_T)$, each of the *marginal* posterior distributions (i)–(v) can be handled analytically. Simulated draws from (i)–(v) are thus easily obtained, and the Gibbs sampler provides a way of iterating on such draws to simulate draws from the full posterior $p(\mathcal{S}_T, \boldsymbol{\beta}|\mathcal{Y}_T)$. We offer the details of the

[15]Note that in our model with sluggish adjustment to the mean, this is only strictly true in expectation for waves having the same duration. If the process is sluggish, then the shorter a wave, the lower its peak.

[16]For a general introduction to Gibbs sampling, readers are referred to Geman and Geman (1984) and Gelfand and Smith (1990). A textbook treatment of the method can also be found in Kim and Nelson (1999).

sampling procedure and the straightforward derivation of the posterior distributions (i)–(v) in the Appendix.

Given the tool to generate representative sets of draws from $p(\mathcal{S}_T, \boldsymbol{\beta}|\mathcal{Y}_T)$, properties of this distribution such as individual parameters' marginal distributions and moments are easily characterized by use of their sample equivalents. To address the central issue of wave identification, we will be particularly interested in $\Pr(S_t = 2|\mathcal{Y}_T)$, the posterior probability of being in a high merger activity state at any date t. These probabilities are obtained by averaging across the simulated paths for the states, each simulated while cycling through the above posterior distributions.

Basing state inference on $\Pr(S_t = 2|\mathcal{Y}_T)$, while natural in the Bayesian framework of Gibbs Sampling, elegantly overcomes a potential pitfall to classical inference. In a classical setting, inference on states is obtained through a two-step procedure by first obtaining a Maximum Likelihood parameter estimate $\hat{\boldsymbol{\beta}}$, and then calculating $\Pr(S_t = 2|\mathcal{Y}_T, \hat{\boldsymbol{\beta}})$, the probability of $S_t = 2$ in any period t under the assumption that $\hat{\boldsymbol{\beta}}$ corresponds to the *true* parameter values.[17] From a Bayesian perspective, the derived inference on states is thus to be read as contingent on the econometrician having full confidence in his parameter estimate $\hat{\boldsymbol{\beta}}$. But only rarely will this correspond to the econometrician's true confidence in $\hat{\boldsymbol{\beta}}$. Moreover, alternative conceivable values of $\boldsymbol{\beta}$ will typically lead to different values of $\Pr(S_t = 2|\mathcal{Y}_T, \boldsymbol{\beta})$, so that uncertainty about $\boldsymbol{\beta}$ will feed into uncertainty on states. As a result, basing state inference on $\Pr(S_t = 2|\mathcal{Y}_T, \hat{\boldsymbol{\beta}})$ rather than on $\Pr(S_t = 2|\mathcal{Y}_T)$ can convey a false degree of certainty about states by neglecting uncertainty about parameters.

To make this important point more transparent, note that $\Pr(S_t = 2|\mathcal{Y}_T)$ and $\Pr(S_t = 2|\mathcal{Y}_T, \hat{\boldsymbol{\beta}})$ are related by

$$\Pr(S_t = 2|\mathcal{Y}_T) = \int \Pr(S_t = 2|\mathcal{Y}_T; \boldsymbol{\beta}) p(\boldsymbol{\beta}|\mathcal{Y}_T) d\boldsymbol{\beta}, \tag{4}$$

where $p(\boldsymbol{\beta}|\mathcal{Y}_T)$ denotes the posterior density of the parameter vector. By (4), our Bayesian inference based on $\Pr(S_t = 2|\mathcal{Y}_T)$ can be interpreted as considering $\Pr(S_t = 2|\mathcal{Y}_T; \boldsymbol{\beta})$ for *any* conceivable parameter constellation $\boldsymbol{\beta}$—that is, for the maximum likelihood estimate $\hat{\boldsymbol{\beta}}$ in particular, but also for any other conceivable $\boldsymbol{\beta}$—, weighting it with the respective posterior probability of $\boldsymbol{\beta}$, and adding up across $\boldsymbol{\beta}$ to produce $\Pr(S_t = 2|\mathcal{Y}_T)$. It is straightforward to see from (4) that if posterior parameter uncertainty is sufficiently high and conditional inference on S_t is sufficiently sensitive to $\boldsymbol{\beta}$, then $\Pr(S_t = 2|\mathcal{Y}_T, \hat{\boldsymbol{\beta}})$ can differ substantially from $\Pr(S_t = 2|\mathcal{Y}_T)$. We will provide an impressive illustration of this difference below in our investigation of US merger activity, where we shall in fact argue that this methodological point is likely to have played a substantial role in previous identifications of a 1980s merger wave.

[17] For details on standard maximum likelihood methods, refer to Hamilton (1989, 1993).

5 Estimation Results

This section reports the results of fitting our lagged mean and variance switching model as given by Assumptions 3.1 through 3.3 to US and UK log merger series.[18] In a first step, Section 5.1 analyzes the entire quarterly US merger series from 1973 through 2003. Its main finding, the identification of a wave beginning in the mid 1990s, is augmented by a look at industry level data in Section 5.2. In a second step, to investigate the lack of evidence for an 1980s merger wave in more detail, Section 5.3 reestimates the model using only data up to 1995. Section 5.4 argues that in so doing, the estimation techniques described in Section 4 play a decisive role. Finally, Section 5.5 offers a shift of focus to the UK by fitting the model to the quarterly UK merger series spanning 1969 through 2003.

As outlined, inference is conducted in a Bayesian context using the Gibbs sampling method to derive posterior distributions of the model parameters and to assess the sequence of unobserved states. In all cases, Gibbs sampling involved 20,000 iterations, where a burn-in sequence consisting of the first 1,000 draws was discarded prior to any inference. Output of the sampler was closely monitored to ensure proper convergence of the filter.

5.1 The Tidal Wave of the Mid 1990s: US Mergers, Full Series 1973:I through 2003:IV

First, we consider the full series on US mergers over the entire available time span from 1973:I through 2003:IV, as presented in Figure 1. Preliminary estimation of the model with various lag lengths k suggested setting $k = 4$.

For all model parameters, Table 1 gives summary statistics both on the priors used and on the marginal posterior distributions obtained. The priors on all parameters where chosen to be relatively uninformative within the class of admissible conjugate priors (which are: normal distributions for the parameters μ_{S_t} and ϕ_i, inverted gamma distributions for $\sigma^2_{S_t}$, and beta distributions for p_{11} and p_{22}). To give an impression of the posterior parameter distributions beyond the mere summary statistics, Figure 2 plots histograms representing the estimated marginal posterior distributions of the parameters. Despite our focus on Bayesian inference, Table 1 also gives maximum likelihood point estimates of all parameters in the last column.

The first feature obvious from Table 1 is that the data indeed leads to signifi-

[18] The main reason for using log rather than level merger data is that all series considered are nonnegative by construction. Strictly speaking, the model as defined by (1) and (2) therefore provides no valid description of the level series due to its capability of producing negative observations. However, all subsequent inference was nonetheless also conducted after fitting the level merger data to the model. Qualitative results, specifically concerning inference on states, differ only little from those obtained for the log merger series. Any remaining noteworthy differences are explicitly pointed out in the subsequent discussion.

	priors		posteriors				ML
	mean	std	mean	median	std	95%-band	MLE
μ_1	6.7	1000.00	6.386	6.390	0.083	$(6.204, 6.538)$	6.3959
μ_2	6.7	1000.00	7.687	7.677	0.103	$(7.510, 7.924)$	7.6653
σ_1	0.2	0.28	0.137	0.136	0.008	$(0.123, 0.153)$	0.1320
σ_2	0.2	0.28	0.096	0.095	0.010	$(0.079, 0.118)$	0.0834
ϕ_1	0.0	1.00	0.683	0.683	0.098	$(0.488, 0.874)$	0.6578
ϕ_2	0.0	1.00	0.063	0.063	0.113	$(-0.159, 0.287)$	0.0538
ϕ_3	0.0	1.00	-0.114	-0.114	0.112	$(-0.334, 0.103)$	-0.1136
ϕ_4	0.0	1.00	0.146	0.145	0.085	$(-0.021, 0.313)$	0.1386
p_{11}	0.9	0.21	0.987	0.991	0.012	$(0.955, 1.000)$	0.9886
p_{22}	0.9	0.21	0.997	1.000	0.010	$(0.969, 1.000)$	1.0000

Note: Result of 20,000 Gibbs-Sampling iterations, iterations 1,000 through 20,000 used for inference. 95%-band refers to 95% posterior probability bands.

Table 1: Estimation Results for US Merger Activity, 1973:I through 2003:IV.

cant updates in the priors on all parameters, as shown by direct comparison of the standard deviations of priors and posteriors. Furthermore, Figure 2 reveals that the corresponding marginal posterior distributions are single peaked and well behaved. As may have been expected from a glance at the original data, comparing the posteriors for μ_1 and μ_2 reveals that mean log merger activity in the high state 2 significantly exceeds that in the low state 1.

Although not literally interpretable as state-contingent means of the untransformed series due to the nonlinear log transformation, the median values of $\exp(\mu_1)$ and $\exp(\mu_2)$ suggest that in level terms, merger activity in the low and high activity state are in the region of 596 and 2,160 mergers per quarter, respectively. Furthermore, log mergers seem to be significantly less volatile in the high activity state, as shown by a comparison of the posteriors on σ_1 and σ_2. Again however, caution is called for when drawing conclusions concerning the volatility of the *level* merger series, as the log transformation compresses differences at higher absolute levels of activity. Indeed, fitting the model to the level merger data suggests that in level terms, mergers are significantly more volatile in the high-activity state. Next, estimates on the autoregressive coefficients ϕ_i show that mergers display a considerable degree of inertia also *within* states. Moreover, mean and median of the largest root across samples both figure at 0.84, suggesting a significant degree of intertia in the merger series. Finally, the posteriors on the transition probabilities p_{11} and p_{22} let us conclude that switches in regime are rather unlikely. Specifically, the median expected duration of a low activity state, $p_{11}/(1 - p_{11})$ is approximately 27 years (mean expected duration is heavily influenced by the skewness in the posterior on q

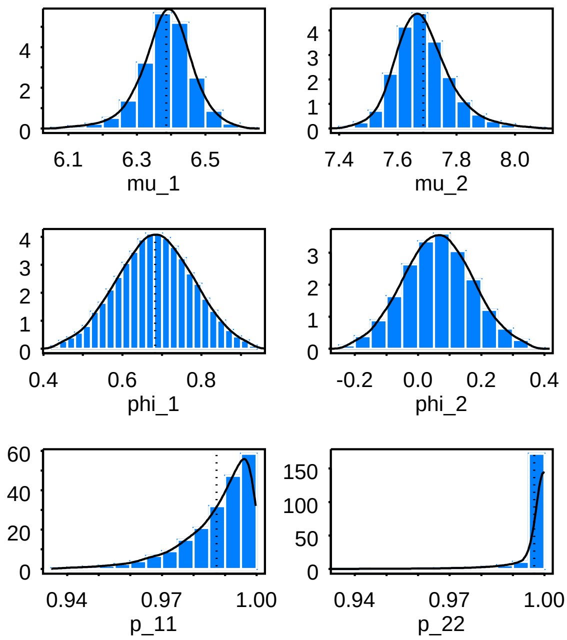


Figure 2: Marginal Posterior Distributions of Parameters for US Log Merger Series 1973:I–2003:IV.

and lies around 133 years). Conversely, estimates on p_{22} show that the high activity state seems to be essentially absorbing, so that a regime of high merger activity is unlikely ever to be left again—this result being driven, of course, by the fact that there does not seem to have been a single reversion from high to low activity in the series so far. Finally, for the ergodic probability of being in a high state (unconditional on the data), $\Pr(S_t = 2) = (1 - p_{22})/(2 - p_{11} - p_{22})$, the mean posterior is 86.5%, whereas the median is 99.7%.

With these results on the model's parameters in mind, let us now return to our main objective, the identification of waves in mergers. Figure 3 plots the probability of being in a high state in any period t given the observed US merger data, $\Pr(S_t = 2|\mathcal{Y}_T)$. This probability plot is shown in the top panel of Figure 3, whereas the bottom panel reproduces the underlying log merger series (along with the posterior median of μ_1 and μ_2 as dashed horizontal lines) for convenience. To highlight the most likely state for any period t, periods for which $\Pr(S_t = 2|\mathcal{Y}_T) > 0.5$ are shaded. However, any interpretation of this 'best guess' should take account of the underlying value of $\Pr(S_t = 2|\mathcal{Y}_T)$ as a straightforward measure of confidence in this guess: The closer $\Pr(S_t = 2|\mathcal{Y}_T)$ to 0.5, the more uncertainty surrounds the best guess.

Figure 3 shows that our estimation produces strikingly clear inference concerning the unobserved state. First, we find compelling evidence that over the entire period between 1973 and 2003, US merger activity has in fact experienced only a single regime switch, that switch being from low to high activity. Cast into the wave

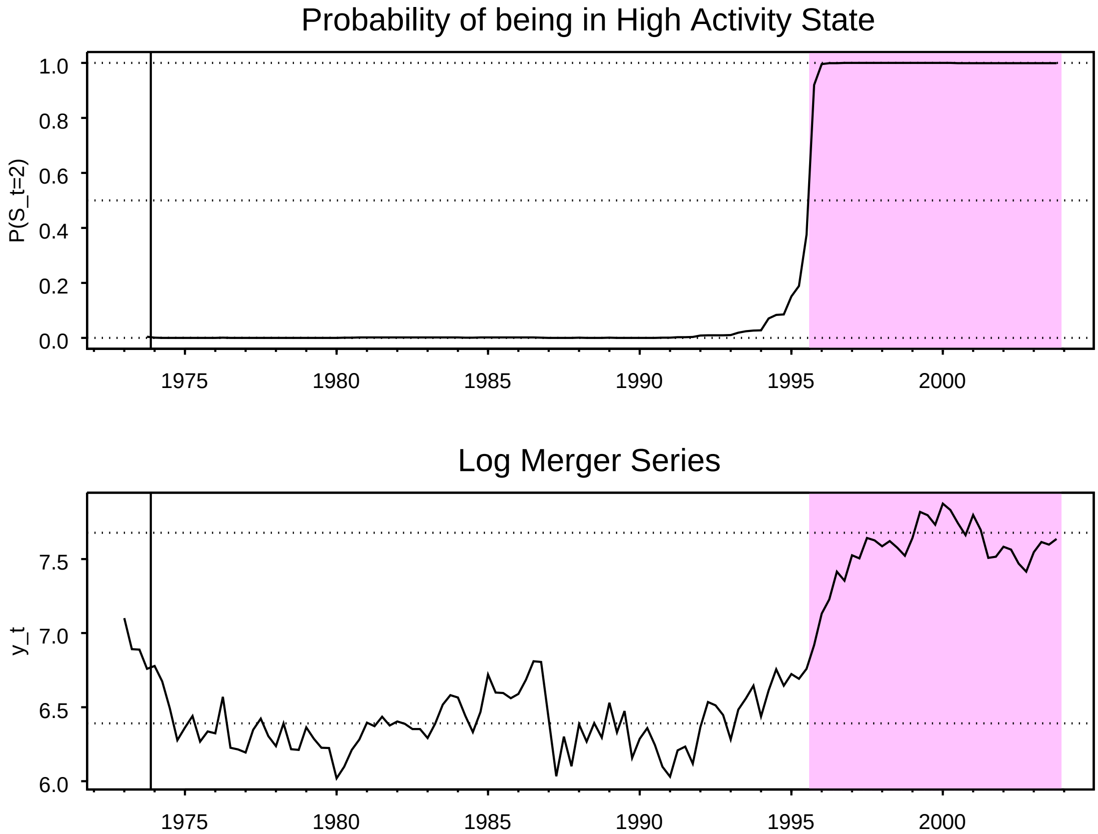


Figure 3: Estimated Probability of Being in State of High Merger Activity, US Log Merger Series 1973:I–2003:IV (in Top Panel, Bottom Panel displays Log Merger Series).

terminology, this suggests that since 1973, the US have so far witnessed only the beginning of a single 'tidal wave' in merger activity. Second, while this observation alone may come as no major surprise given a glance at the log merger plot, the clear-cut jump in the probability plot also allows us to date the beginning of this tidal wave rather precisely. Specifically, the assessed probability of a high state of merger activity jumps from 0.189 in 1995:II and 0.375 in 1995:III to a value of 0.920 in 1995:IV. We may thus conclude that the wave in US merger activity is very likely to have been triggered between the third and fourth quarter of 1995.[19]

Rounding up, we should stress three points relevant to the interpretation of these results. First, as pointed out in Section 4, using $\Pr(S_t = 2|\mathcal{Y}_T)$—rather than $\Pr(S_t = 2|\mathcal{Y}_T, \hat{\boldsymbol{\beta}})$ for a point estimate of the parameters $\hat{\boldsymbol{\beta}}$—means taking account of uncertainty about the model's structural parameters for the inference on states.

[19]Specifically, the probability of the US having witnessed a *single* state switch from low to high between 1995:III and 1995:IV (rather than at any other date) can be calculated at 54.2%, which is contrasted by the probability of a corresponding single switch one quarter earlier (i.e. between 1995:II and 1995:III) of 18.5%, of 3.7% two quarters earlier, and 7.5% one quarter later.

It is all the more noteworthy that Figure 3 conveys an appreciably clear message concerning the likely sequence of states. Second, although we shall more thoroughly investigate the failure to identify the hump in merger activity around the mid 1980s as a wave in a moment in Section 5.3, let us note for now that this result is even more clear-cut when fitting the state-switching model to the level rather than the log-merger data. Intuitively, this is due to the simple fact that the log transformation exaggerates the mid 1980s hump in merger activity relative to the level data. Finally, we should point out our estimation procedure's weakness in producing inference on states at the very beginning of the series: For technical reasons, inference on states actually begins only in 1974:I (rather than in 1973:I), and inference on states in these early periods is generally somewhat sensitive to starting values used.[20]

5.2 Reflections of the US 1990s Wave at the Industry Level

As a prominent explanation of the clustering of merger activity in time, the literature has advanced the idea that surges in aggregate merger activity represent firms' optimal responses to industry-level shocks.[21] According to this hypothesis, waves in merger activity at the aggregate level will be the result of temporary surges in merger activity in one or a few industries. Concerning our above findings, this naturally raises the question of whether the marked increase in US merger activity in the mid 1990s was dominated by any specific industries. While the available data do not permit us to estimate our model at the industry level, casual investigation of annual industry level data suggests that the mid 1990s wave is hardly attributable to one or a few industries alone.

To this end, of the 50 industries identified by the *Mergerstat Review*, Figure 4 plots *annual* merger data for those eleven US industries with the strongest merger activity between 1990 and 2003, where industries were ranked according to overall activity (in terms of numbers of mergers) for that period. While Figure 4 shows that industries were certainly not uniformly hit by a wave in 1995, it is nevertheless

[20]To understand the first point, note that our estimation of an AR(k) model takes the first k observations of y_t as exogenously given, which is why inference on states begins in 1974:I. Concerning the second point, inferring states by means of the Hamilton filter requires specifying the initial exogenous probability of being in a high activity state, $\Pr(S_0 = 2|\boldsymbol{\beta}; y_{-k+1}, \ldots, y_0)$. All inference shown was produced using an uninformative but nonetheless somewhat arbitrary initial probability of 0.5, although simulations with alternative initial probabilities were run to check the results for robustness. Specifically, for the full US merger series considered here, results turn out to be very robust despite the relatively high level of merger activity at the very beginning of the series. Intuitively, this stems from the fact that, even though the data suggests that the US may have been in a high level of merger activity immediately prior to 1973, the strong downward trend at the beginning of the sampling period clearly leads us to conclude that the US economy must nevertheless have already found itself in a low state of merger activity at the time our inference on states begins.

[21]See for instance Mitchell and Mulherin (1996), Mulherin and Boone (2000), Andrade et al. (2001), and Harford (2005).

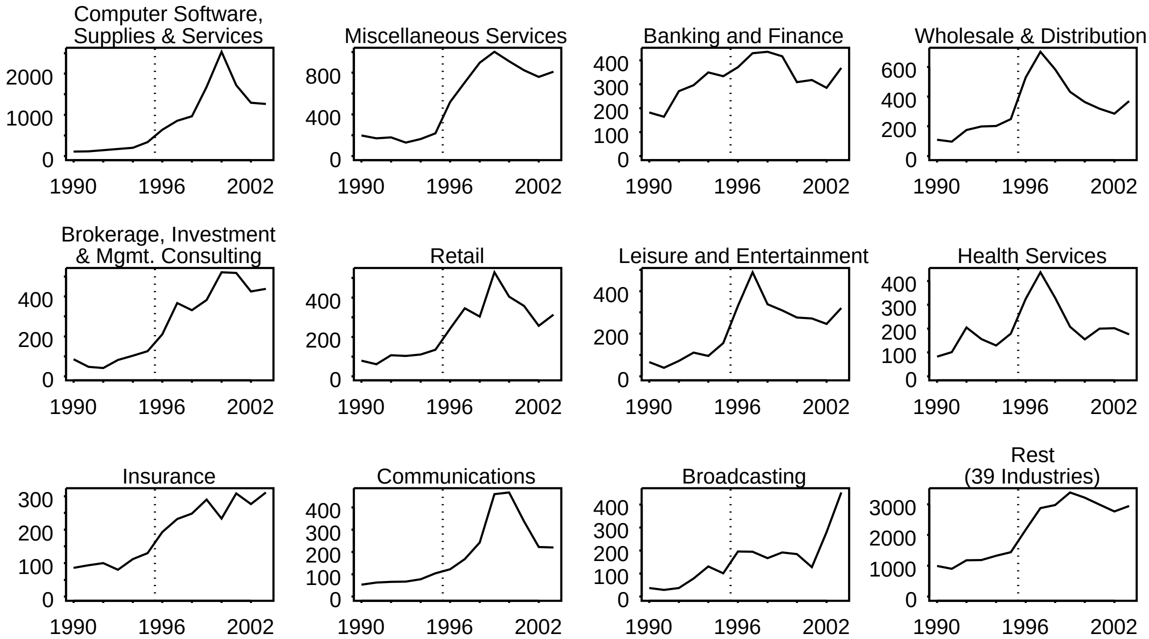


Figure 4: Annual US Industry Level Merger Activity, 1990 through 2003 (Source: *Mergerstat Review*).

apparent that the resulting aggregate wave is anything but the result of a single industry level burst. For instance, the largest industry level share in overall annual merger activity was in Computer Software, Supplies and Services at its pronounced peak in 2000, with a share of 26%. Although significant, such shares still make it impossible to explain the pronounced jump in aggregate mergers—from around 600 per quarter before the wave to over 2,000 thereafter—as caused only by a small subset of industries. Furthermore, as the bottom right plot in Figure 4 shows, even after removing the 11 most active industries (which account for 52% to 66% of annual merger activity between 1990 and 2003), residual merger activity still gives a strong impression of a mid 1990s merger wave.

5.3 What ever happened to the 1980s Merger Wave? US Mergers 1973:I through 1995:III

One of the most striking findings in Section 5.1 has been our failure to identify even the faintest hint of the much-discussed 1980s merger wave. Indeed, a simple look at the log merger series depicted in Figure 3 may indeed raise questions about there being something less pronounced—but nonetheless 'wave-like'—about the visible hump in merger activity in the mid 1980s.

As pointed out previously, an arguably rather strong assumption implicit in our two-state model is that waves always come in similar sizes (where 'size' refers to a wave's peak height). Along these lines, a perfectly valid reservation with our result

	priors		posteriors				ML
	mean	std	mean	median	std	95%-band	MLE
μ_1	6.4	1000.00	5.804	6.274	1.179	(2.467, 6.451)	6.3129
μ_2	6.4	1000.00	6.604	6.588	0.206	(6.323, 7.037)	6.7007
σ_1	0.1	0.14	0.125	0.128	0.036	(0.053, 0.200)	0.1222
σ_2	0.1	0.14	0.105	0.105	0.030	(0.055, 0.161)	0.0753
ϕ_1	0.0	1.00	0.704	0.728	0.136	(0.399, 0.932)	0.4655
p_{11}	0.9	0.21	0.736	0.890	0.292	(0.078, 1.000)	0.9495
p_{22}	0.9	0.21	0.793	0.907	0.252	(0.115, 1.000)	0.8518

Note: Result of 20,000 Gibbs-Sampling iterations, iterations 1,000 through 20,000 used for inference. 95%-band refers to 95% posterior probability bands.

Table 2: Estimation Results for US Merger Activity, 1973:I through 1995:III.

might be that, even if US merger activity between 1973 and 2003 was dominated by a single gigantic tidal wave in the mid 1990s, the assumption of similarly-sized waves downplays the importance of underlying, less gigantic, but nonetheless significant and perhaps more regular wave-like merger activity.

As a straightforward way to investigate this possibility, this section presents estimates for the two-state model using only the data *prior* to the estimated break date which started the tidal wave, i.e. from 1973:I through 1995:III. Note that this corresponds to discarding little more than a quarter of the full series, which should leave us with sufficient data points to identify waves. Table 2 gives the parameter estimates resulting from fitting an AR(1) model. Comparison with estimation results from the full series in Table 1 shows that parameter inference is rather imprecise for the subsample. Figure 5 again shows the inferred probabilities of being in a high activity state for this particular subsample period. By comparison with the clear-cut result presented in Figure 3 for the entire sample period, Figure 5 suggests that waves are rather hard to identify in the merger data up to 1995:III. Although the plot does indeed hint at a somewhat increased probability of a high activity state around the mid 1980s (as well as around the mid 1990s as a warm-up to the ensuing large wave), this hint remains very faint due to the fact that, except for a short period around 1987, the probability of a high activity state stays in a rather tight band around 0.5. Overall, the fact that the inferred probability of a high activity state, $\Pr(S_t = 2|\mathcal{Y}_T)$, stays far clear of either 0 or 1 implies that the data reaches no clear conclusion concerning a likely sequence of states.[22]

[22] An alternative approach would be to extend our two-state model by introducing a third state, thereby explicitly allowing for both 'medium' and 'high' waves. However, casual inspection of the series strongly suggests that in such a three-state model, all quarters following 1995:III would rather clearly be associated with the 'high wave state', leading to inference on 'low' vs. 'medium' state comparable to the two-state analysis presented in this section.

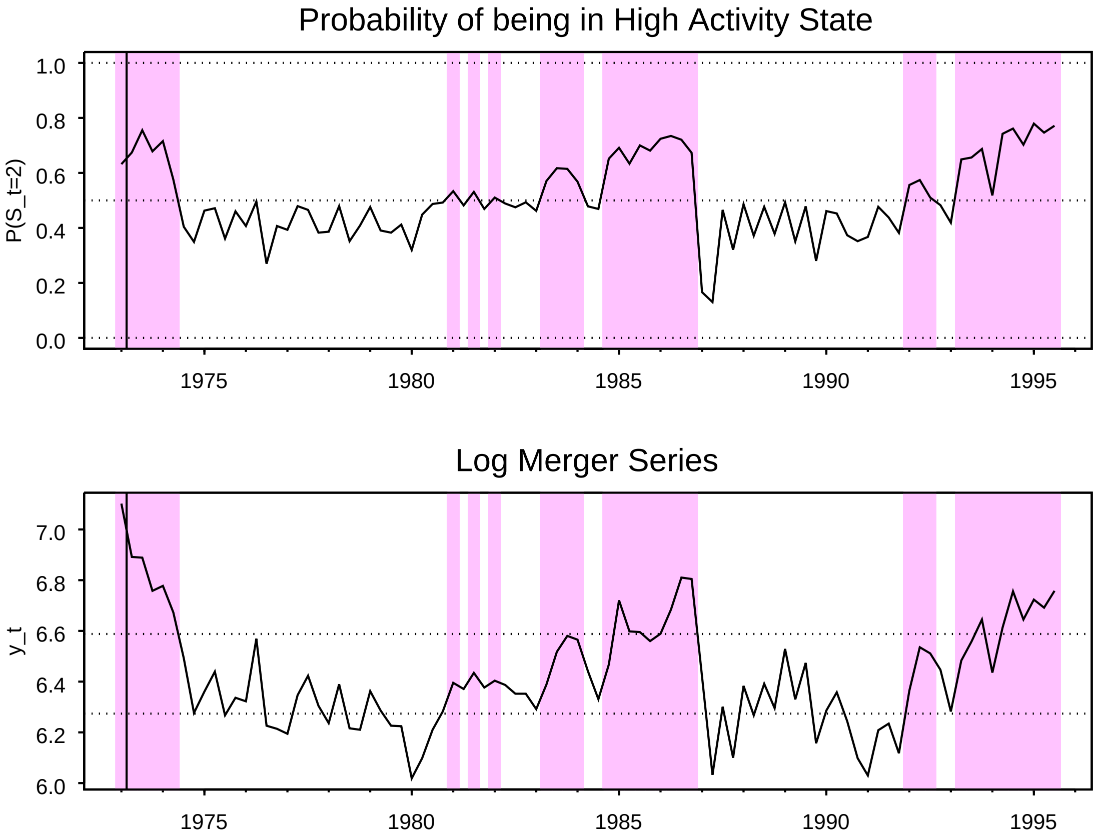


Figure 5: Estimated Probability of Being in State of High Merger Activity, US Log Merger Series 1973:I–1995:III (in Top Panel, Bottom Panel displays Log Merger Series).

This indecision in state inference is a first clear indication that the US have not witnessed a change in the mean level of merger activity at all during the 1980s.[23] A second striking feature of the probability plot in Figure 5 is its strong qualitative similarity with the underlying time series of log mergers, reproduced in the bottom panel of Figure 5. Indeed, the high-activity estimate in the top panel comes quite close to representing a positive affine transformation of the log merger series in the bottom panel. This is something we should expect to see from fitting a regime switching model to a series generated by a process with no actual switch in regime. Third, Albert and Chib (1993) point out that fitting a Markov switching model to non-switching data results in rather large posterior bands on parameters, particu-

[23]Note that in finite samples, given the data $\mathcal{Y}_T$, $\mathrm{P}(S_t = 2|\mathcal{Y}_T)$ will of course generally deviate from 0.5 for any t even if the data were indeed generated by a stationary autoregressive process with no change in regime (i.e., a process with $\mu_1 = \mu_2$ and $\sigma_1 = \sigma_2$) due to remaining posterior uncertainty about the model parameters. Sample runs of the Gibbs Sampler on simulated data involving no regime change (and parameter values similar to those inferred for US mergers) revealed pictures quite similar to Figure 5.

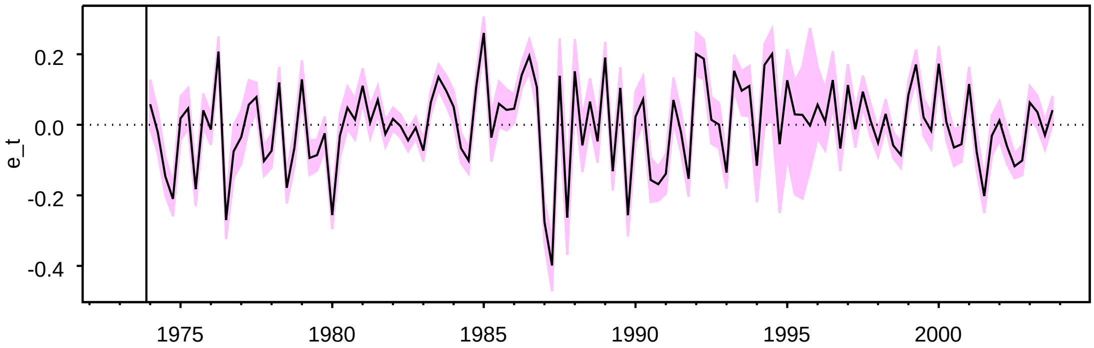


Figure 6: Mean Residuals (with 95% Posterior Bands) for US Log Merger Series 1973:I–2003:IV.

larly concerning the means μ_1 and μ_2 and the transition probabilities p_{11} and p_{22}, which is reflected in our results. Finally and perhaps most importantly, inference runs with simulated data show the posterior distributions of $\mu_2 - \mu_1$ for our US subsample to be quite comparable to the posterior distributions of $\mu_2 - \mu_1$ resulting from simulated series with similar but constant model parameter values (i.e. a series with no state switch).

In sum, therefore, we find little evidence in support of the notion of a 1980s merger wave even in the truncated series. It is important however to be clear about the exact meaning of this result, as it does not contradict of course the 1980s having witnessed somewhat increased merger activity. What our result fleshes out is that this increased activity is rather unlikely to have been associated with a nonlinear shift in regime to the underlying autoregressive process (i.e. a 'burst in activity', such as the boom following 1995:III). Rather, it appears to be quite compatible with regular, well-behaved random shocks hitting a stationary linear autoregressive process. This is perhaps best illustrated by Figure 6, which plots the mean residuals (i.e. estimates of ε_t) resulting from the original estimation of the model over the full data range from 1973:I through 2003:IV.[24] Overall, these residuals appear well-behaved and compatible with our model assumptions of serial independence and normality. Nonetheless, one might indeed see the first half of the 1980s as having been hit by a sequence of slightly above-average shocks which—amplified by the processes' strong positive autocorrelation—have given rise to a period of somewhat increased merger activity. We argue, however, that this is compatible with the usual behavior of a stationary autoregressive process rather than signifying a non-linear burst such as a discrete switch in mean over that period. In other words, in terms

[24]Recall that in a Bayesian estimation context, we consider posterior (i.e. 'updated') *distributions* of the parameters rather than particular point estimates. Thus, the resulting residuals themselves are random not only due to uncertainty about the unobservable state, but also due to posterior parameter uncertainty. Figure 6 displays both the mean and 95% posterior probability bands for the residual in any period.

of our regime switching model of waves, we would like to suggest that the 1980s merger wave symbolized a 'ripple' rather than a real wave.

5.4 Parameter Uncertainty Matters: (Mis)identifying the 1980s Wave

The preceding section has thoroughly discussed our finding that, in contrast to the 1990s merger wave, the increased merger activity in the 1980s constituted no extraordinary burst in activity. This leaves unexplained, however, why the aforementioned studies by Town (1992) and Linn and Zhu (1997), both of which similarly fit a two-state mean-switching Markov model to US merger data, have identified an 1980s merger wave nonetheless.

Candidate reasons for this difference are manifold, as the studies differ in the particular series used, the time span considered, and details in model specification. As this section argues, however, the main difference in our interpretation of the 1980s merger wave is likely to stem from a more subtle, methodological reason: the refined methods of inference offered by the Gibbs sampling approach.

Recall that by means of this approach, our inference on regimes as presented in Figures 3 and 5 is based on $\Pr(S_t = 2|\mathcal{Y}_T)$. This contrasts with Town's (1992) and Linn and Zhu's (1997) inference based essentially on $\Pr(S_t = 2|\mathcal{Y}_T; \hat{\boldsymbol{\beta}})$, the likelihood of a high state of merger activity while treating the underlying model parameters as given by the result of a preceding maximum likelihood estimation. Put differently, previous studies have asked the following question: Given that we believe the obtained parameter estimates to correspond to the true parameter values, what is the likelihood of a high-activity state in any period? What this question obviously forgets to ask is just how reliable these parameter estimates actually are. Thus, if parameter estimates involve a high degree of uncertainty, answers to this question can severely overstate the evidence in favor of a high state of activity, and this appears to be highly relevant to the discussion of the 1980s merger wave.

To drive this point home, using our data spanning 1973:I through 1995:III, we have replicated the procedure in Town (1992) and Linn and Zhu (1997) by calculating the probability of a high activity state while holding the model parameters fixed at their maximum likelihood estimates reported in Table 2. Figure 7 plots the results. Clearly, by comparison with Figure 5, neglecting parameter uncertainty both leads to considerably more clear-cut inference on regimes and accentuates evidence for a high state of activity in the 1980s. Interestingly, the resulting sharp identification of a merger wave lasting from late 1984 to late 1986 is not very far from findings in Town (1992), who identifies 1986:IV as a period of intense merger activity, and findings in Linn and Zhu (1997), who identify the 'mid-to-late 1980s' as a merger wave.

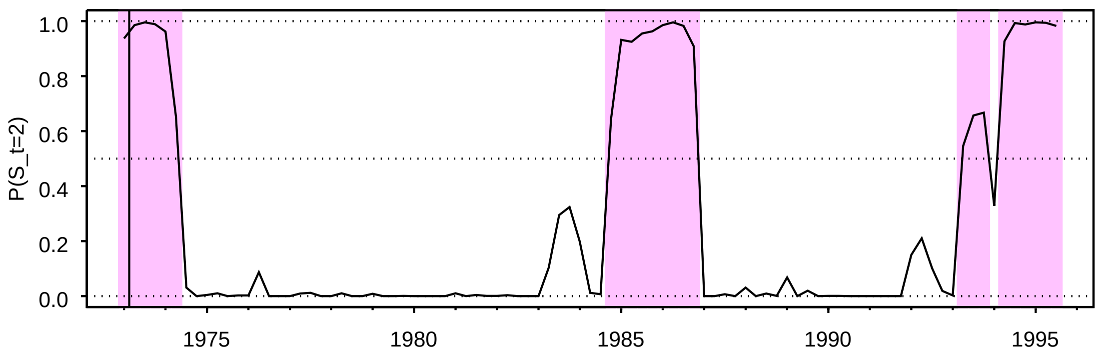


Figure 7: Probability of Being in High State of Merger Activity, US Log Merger Series 1973:I–1995:III, Calculated at Maximum Likelihood Parameter Estimates.

5.5 But Waves Do Exist: The UK Merger Wave Experience

Next, we inspect the UK merger series, shown in the bottom panel of Figure 1, for its wave-like behavior. Analyzing the UK series turns out to be rewarding not only from the point of view of understanding UK merger activity, but also as a more general validation of our proposed Markov switching merger model and the wave hypothesis in particular. Indeed, the preceding analysis of the US merger data may be seen as somewhat disappointing as regards the wave hypothesis: While the proposed model itself does seem to provide a very good description of the US data, it does so in a fashion that hardly reflects the *repeated* bursts of activity attributed to mergers by the literature—namely by identifying only a single switch from low to high activity around the mid 1990s, but no subsequent reversion to low activity, let alone a second or even a third wave. Given the limitation of our analysis to the last 30 years of US merger activity, this finding is of course not an invalidation of the wave hypothesis *per se*. However, the idea of a wave-like process governing US mergers would certainly be reinforced by observing more complete and possibly repeated wave cycles in other series such as the UK's.

To analyze the UK data, we estimate an AR(1)-version of the model in Section 3. Table 3 reports summary statistics on both the priors used and on the estimated posterior distributions of the parameters. As with the full US series, the UK series permits a significant update on the model parameters' priors, shown by the marginal posteriors' standard deviation and posterior bands. Again, the inferred high activity mean μ_2 significantly exceeds the low activity mean μ_1. Interestingly, the ratio of high to low activity (in log terms) seems comparable across the US and the UK, as the high activity mean μ_2 exceeds the low activity mean μ_1 by an average (and mean) 20% for both series. Also, shocks to UK mergers again appear to be less volatile in the high activity state, although this is much less significantly so than for the full US series (as would be expected, fitting the model to the untransformed data reveals much higher volatility in the high than in the low activity state). The

	priors		posteriors				ML
	mean	std	mean	median	std	95%-band	MLE
μ_1	5.00	1000.00	4.789	4.823	0.342	(4.655, 4.957)	4.8328
μ_2	5.00	1000.00	5.756	5.781	0.325	(5.057, 6.037)	5.8324
σ_1	0.15	0.21	0.193	0.193	0.020	(0.141, 0.225)	0.1915
σ_2	0.15	0.21	0.173	0.169	0.033	(0.126, 0.255)	0.1497
ϕ_1	0.00	1.00	0.603	0.583	0.105	(0.455, 0.896)	0.5511
p_{11}	0.90	0.21	0.973	0.985	0.067	(0.892, 0.999)	0.9810
p_{22}	0.90	0.21	0.908	0.928	0.094	(0.724, 0.991)	0.9061

Note: Result of 20,000 Gibbs-Sampling iterations, iterations 1,000 through 20,000 used for inference. 95%-band refers to 95% posterior probability bands.

Table 3: Estimation Results for UK Merger Activity, 1969:I through 2003:IV.

estimate of ϕ_1 suggests that autocorrelation in the UK series is significantly positive and again quite comparable to the degree observed in the US data. Finally, however, the unobserved state is more likely to change in any period for the UK series, as shown by the posteriors on p_{11} and p_{22}. Particularly, the median expected duration of a low activity state is 16 years whereas the median expected duration of a high activity state is 13 quarters in the UK. Concerning the ergodic probability of being in a high state (unconditional on the data), $\Pr(S_t = 2)$, the mean posterior is 22.9%, whereas the median is 17.0%.

Next, Figure 8 shows the estimated probabilities of being in a high activity state for the UK merger series. The probability plot shows strong evidence that the UK has witnessed two distinct merger waves between 1969 and 2003. The first wave seems to have lasted from 1971:I through 1973:IV. Due to the aforementioned inference problems at the beginning of the series as well as because the merger series seems somewhat 'undecided' prior to 1971, we cannot precisely date the beginning of the first wave. Indeed, our inference leaves open to some extent whether the UK started off in a high or low activity state in 1969.[25] It appears quite clear, however, that this first wave found its end in 1973:IV, as $\Pr(S_t = 2|\widetilde{y}_T)$ drops from 0.836 in 1973:IV to 0.066 in 1874:I. The second wave in turn is reliably estimated to have started in 1986:III and ended in 1989:IV (the probability of a high state jumps from 0.320 in 1986:II to 0.955 in 1986:III and dips from 0.936 in 1989:IV to 0.109 in 1990:I).

[25] In fact, by investigating annual data, Hughes (1993, p. 17) argues that merger activity culminates in twin peaks of activity in the late 1960s and early 1970s.

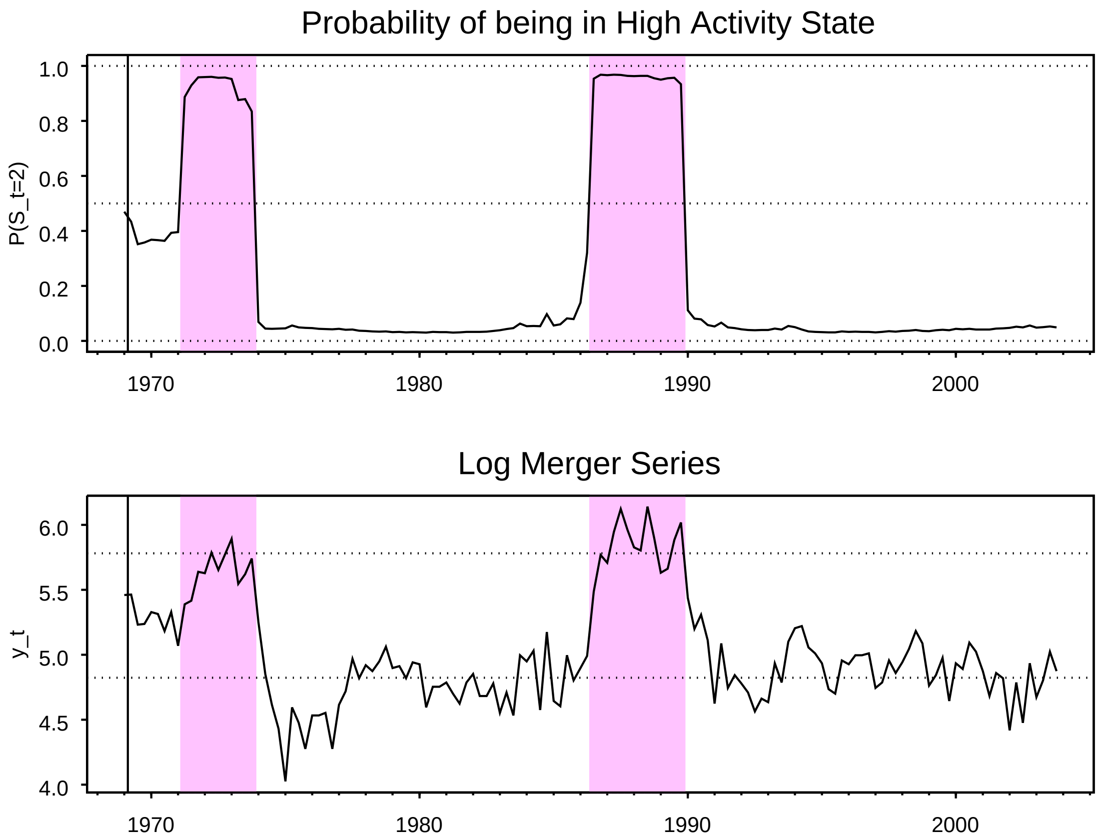


Figure 8: Estimated Probability of Being in State of High Merger Activity, UK Log Merger Series 1969:I–2003:IV (in Top Panel, Bottom Panel displays Log Merger Series).

6 Conclusion

The goal of this paper has been to revisit quantitative evidence on the merger wave hypothesis. Using a model of Markovian parameter switching, recent merger data and refined methods of inference, we have sought to identify and date waves in merger activity. A key finding has been that, concerning merger activity in the US and the UK over the past 30 years, the interpretation of merger activity as a mean and variance switching autoregressive process provides a promising quantitative operationalization of the wave hypothesis.

Moreover, fitting such a model to the data has produced the following merger wave chronology: First, since the beginning of our US series in 1973, the US appear to have witnessed only the beginning of a single large merger wave, this wave having been kicked off between the third and fourth quarter of 1995. Particularly, as a second major result and in contrast to other recent empirical work, we find very little evidence for the much discussed 1980s merger wave. We have argued that there is a methodological reason for this discrepancy in findings, as less refined inference methods which neglect parameter uncertainty are likely to have played a significant role in the econometric misidentification of 1980s merger activity as a wave. Third, fitting our model to UK merger activity between 1969 and 2003 clearly identifies two UK merger waves, one in the early 70s and a second in the late 1980s.

We hope that these findings will serve as a sound basis for a further discussion and investigation of possible underlying causes for merger waves. Particularly, the rather precise identification and precise timing of distinct states of merger activity based on our Markov-switching model openly calls for an economic interpretation and explanation of these states. Investigating one such hypothesis, our brief digression in Section 5.2 concerning industry-level data for the US has argued that, whatever the trigger for the US 1990s wave, industries seem to have been rather uniformly affected by it. Resende (1999), using an industry-level Markov switching model, reaches a similar conclusion concerning UK merger activity. Interestingly, while this leads us to conclude that nationally, industries in each the US and the UK seem to have been similarly affected by the observed waves, comparison of our results in Sections 5.1 and 5.5 quickly reveals that there is no sign of a similar 'coordination of waves' across countries over the last 30 years.

Acknowledgements

The authors are indebted to Stefan Buehler, James Hamilton, Armin Schmutzler, Konrad Stahl, Ulrich Woitek, Michael Wolf, and seminar audiences at the University of Zurich, the 2004 EC2 conference on the econometrics of industrial organization, and the 2005 annual meeting of the Swiss Society of Economics and Statistics for most helpful comments and suggestions.

Appendix: Implementing the Gibbs Sampler

This appendix demonstrates the application of Gibbs sampling to Bayesian estimation of the Markov switching model as given by Assumptions 3.1 through 3.3.

A.1 The Sampling Scheme

Collect the state variables and the model parameters in the vector $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3, \boldsymbol{\theta}_4, \boldsymbol{\theta}_5)$, where $\boldsymbol{\theta}_1 = (S_1, S_2, ..., S_T)$, $\boldsymbol{\theta}_2 = (p_{11}, p_{22})$, $\boldsymbol{\theta}_3 = (\mu_1, \mu_2)$, $\boldsymbol{\theta}_4 = (\sigma_1^{-2}, \sigma_2^{-2})$, and $\boldsymbol{\theta}_5 = (\phi_1, ..., \phi_k)$. To obtain a sample from the posterior distribution of interest, $p(\boldsymbol{\theta}|\mathcal{Y}_T)$, we employ the following iterative procedure:

(S1) Specify arbitrary initial values $\boldsymbol{\theta}_2^{(m)}$, $\theta_3^{(m)}$, $\boldsymbol{\theta}_4^{(m)}$, and $\boldsymbol{\theta}_5^{(m)}$ for $m = 0$.

(S2) Cycle through the full set of conditionals, drawing

(i) $\boldsymbol{\theta}_1^{(m+1)}$ from $p(\boldsymbol{\theta}_1|\boldsymbol{\theta}_2^{(m)}, \boldsymbol{\theta}_3^{(m)}, \boldsymbol{\theta}_4^{(m)}, \boldsymbol{\theta}_5^{(m)}, \mathcal{Y}_T)$

(ii) $\boldsymbol{\theta}_2^{(m+1)}$ from $p(\boldsymbol{\theta}_2|\boldsymbol{\theta}_1^{(m+1)}, \boldsymbol{\theta}_3^{(m)}, \boldsymbol{\theta}_4^{(m)}, \boldsymbol{\theta}_5^{(m)}, \mathcal{Y}_T)$

(iii) $\boldsymbol{\theta}_3^{(m+1)}$ from $p(\boldsymbol{\theta}_3|\boldsymbol{\theta}_1^{(m+1)}, \boldsymbol{\theta}_2^{(m+1)}, \boldsymbol{\theta}_4^{(m)}, \boldsymbol{\theta}_5^{(m)}, \mathcal{Y}_T)$

(iv) $\boldsymbol{\theta}_4^{(m+1)}$ from $p(\boldsymbol{\theta}_4|\boldsymbol{\theta}_1^{(m+1)}, \boldsymbol{\theta}_2^{(m+1)}, \boldsymbol{\theta}_3^{(m+1)}, \boldsymbol{\theta}_5^{(m)}, \mathcal{Y}_T)$

(v) $\boldsymbol{\theta}_5^{(m+1)}$ from $p(\boldsymbol{\theta}_5|\boldsymbol{\theta}_1^{(m+1)}, \boldsymbol{\theta}_2^{(m+1)}, \boldsymbol{\theta}_3^{(m+1)}, \boldsymbol{\theta}_4^{(m+1)}, \mathcal{Y}_T)$.

(S3) Repeat (S2) for $m = 1, 2, \ldots, M$.

Geman and Geman (1984) have shown that the sequence $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots$ converges in distribution to a random quantity that has posterior distribution $p(\boldsymbol{\theta}|\mathcal{Y}_T)$. Discarding the first, say, b iterates to attain the desired convergence in distribution (the "burn-in" period), we approximate the joint and marginal distributions of $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3, \boldsymbol{\theta}_4, \boldsymbol{\theta}_5)$ by the empirical distributions of $(\boldsymbol{\theta}_1^{(m)}, \boldsymbol{\theta}_2^{(m)}, \boldsymbol{\theta}_3^{(m)}, \boldsymbol{\theta}_4^{(m)}, \boldsymbol{\theta}_5^{(m)})$, where $m = b+1, \ldots, M$.

The set of conditionals used in (S2) are derived in Sections A.2 to A.6.

A.2 Generating $\boldsymbol{\theta}_1 = (S_1, S_2, \ldots, S_T)$ from $p(\boldsymbol{\theta}_1|\boldsymbol{\theta}_2, \boldsymbol{\theta}_3, \boldsymbol{\theta}_4, \boldsymbol{\theta}_5, \mathcal{Y}_T)$

Suppressing the conditioning on $\boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_5$, the distribution of $\boldsymbol{\theta}_1$ conditional on the data may be written as

$$p(\boldsymbol{\theta}_1|\mathcal{Y}_T) = p(S_T|\mathcal{Y}_T) \prod_{t=1}^{T-1} p(S_t|S_{t+1}, \ldots, S_T, \mathcal{Y}_T).$$

A full draw of the vector $\boldsymbol{\theta}_1$ may thus be generated iteratively by drawing S_T from $p(S_T|\mathcal{Y}_T)$, and then recursively drawing each S_t, $t = T-1, \ldots, 1$, from $p(S_t|S_{t+1}, \ldots, S_T, \mathcal{Y}_T)$, its distribution conditional on having drawn $S_{t+1}, \ldots, S_T$ and on the data $\mathcal{Y}_T$.

The probabilities $p(S_t|S_{t+1}, \ldots, S_T, \mathcal{Y}_T)$ required in each such step may in turn be based on the filtered probabilities $p(S_t|\mathcal{Y}_t)$ produced by the well-known Hamilton filter (Hamilton, 1989). As we show at the end of this section,

$$p(S_t|S_{t+1}, \ldots, S_T, \mathcal{Y}_T) = p(S_t|S_{t+1}, \mathcal{Y}_t), \tag{A.1}$$

so that the realizations of $S_{t+2}, \ldots, S_T$ and $y_{t+1}, \ldots, y_T$ are irrelevant when drawing S_t in the above procedure. Furthermore, by simple application of Bayes' law (and using the Markov property of S_t),

$$p(S_t|S_{t+1}, \mathcal{Y}_t) = \frac{p(S_{t+1}|S_t)p(S_t|\mathcal{Y}_t)}{p(S_{t+1}|\mathcal{Y}_t)} = \frac{p(S_{t+1}|S_t)p(S_t|\mathcal{Y}_t)}{\sum_{\tilde{S}_t \in \{1,2\}} p(S_{t+1}|\tilde{S}_t)p(\tilde{S}_t|\mathcal{Y}_t)}. \tag{A.2}$$

Noting that the transition probabilities $p(S_{t+1}|S_t)$ appearing in (A.2) are determined by $\boldsymbol{\theta}_2$ (which is fixed) and that $p(S_t|\mathcal{Y}_t)$, $t = 1, \ldots, T$ can be determined by the Hamilton filter given $\boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_5$ provides us with the following scheme for generating $\boldsymbol{\theta}_1$:

(i) Employ Hamilton's filter to obtain the filtered probabilities $p(S_t|\mathcal{Y}_t)$ for $t = 1, 2, ..., T$ (given the fixed parameters $\boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_5$) and save them. The last iteration of the filter provides $p(S_T|\mathcal{Y}_T)$, from which S_T is generated.

(ii) Iteratively generate S_t conditional on $\mathcal{Y}_t$ and S_{t+1} , for $t = T-1, T-2, ..., 1$, making use of (A.2) and the filtered probabilities derived above.

To make this procedure more concrete, let

$$\hat{\boldsymbol{\xi}}_{t|t} \equiv \begin{bmatrix} p(S_t = 1|\mathcal{Y}_t) \\ p(S_t = 2|\mathcal{Y}_t) \end{bmatrix}$$

contain the filtered probabilities for each state $i = 1, 2$ at date t, let

$$\boldsymbol{\eta}_t \equiv \begin{bmatrix} p(y_t|S_t = 1, \mathcal{Y}_{t-1}) \\ p(y_t|S_t = 2, \mathcal{Y}_{t-1}) \end{bmatrix}$$

collect the conditional densities of y_t for each state $i = 1, 2$ (given the fixed state-contingent parameters), and let $\mathbf{P}$ represent the transition matrix for the state vector (p_{ij} is the row i, column j element thereof). The Hamilton filter provides the relation

$$\hat{\boldsymbol{\xi}}_{t|t} = \frac{(\boldsymbol{\eta}_t \odot \mathbf{P}\hat{\boldsymbol{\xi}}_{t-1|t-1})}{\mathbf{1}'(\boldsymbol{\eta}_t \odot \mathbf{P}\hat{\boldsymbol{\xi}}_{t-1|t-1})},$$

where '$\odot$' denotes element-by-element multiplication, and $\mathbf{1}$ denotes a (2×1) vector of 1s. Herewith, employing the 'uninformative' initial value $\hat{\boldsymbol{\xi}}_{0|0} = (0.5, 0.5)'$, we generate all subsequent $\hat{\boldsymbol{\xi}}_{t|t}$. Moreover, the distribution of S_t conditional on $S_{t+1}, \mathcal{Y}_t$ required in (A.2) may be expressed explicitly as

$$p(S_t = i|S_{t+1} = j, \mathcal{Y}_t) = \frac{p_{ij}\mathbf{e}_i'\hat{\boldsymbol{\xi}}_{t|t}}{\mathbf{e}_j'\mathbf{P}\hat{\boldsymbol{\xi}}_{t|t}}. \tag{A.3}$$

Finally, we establish equation (A.1), which says that $S_{t+2}, \ldots, S_T$ and $y_{t+1}, \ldots, y_T$ contain no information on S_t beyond that already contained in S_{t+1} and $y_1, \ldots, y_t$. Letting $\mathcal{Y}_{t+1,T} \equiv (y_{t+1}, y_{t+2}, \ldots, y_T)$, $T > t$, denote the vector of observations from date $t + 1$ through T, we have

$$\begin{aligned} p(S_t|S_{t+1}, \ldots, S_T, \mathcal{Y}_T) &= p(S_t|S_{t+1}, \ldots, S_T, \mathcal{Y}_t, \mathcal{Y}_{t+1,T}) \\ &= \frac{p(S_t, \mathcal{Y}_{t+1,T}|S_{t+1}, \ldots, S_T, \mathcal{Y}_t)}{p(\mathcal{Y}_{t+1,T}|S_{t+1}, \ldots, S_T, \mathcal{Y}_t)} \\ &= \frac{p(\mathcal{Y}_{t+1,T}|S_t, S_{t+1}, \ldots, S_T, \mathcal{Y}_t)}{p(\mathcal{Y}_{t+1,T}|S_{t+1}, \ldots, S_T, \mathcal{Y}_t)} \cdot p(S_t|S_{t+1}, \ldots, S_T, \mathcal{Y}_t). \end{aligned} \tag{A.4}$$

Now

$$p(\mathcal{Y}_{t+1,T}|S_t, S_{t+1}, \ldots, S_T, \mathcal{Y}_t) = \prod_{j=1}^{T-t} p(y_{t+j}|S_t, S_{t+1}, \ldots, S_T, \mathcal{Y}_{t+j-1}),$$

where, by definition of our model,

$$y_{t+j}|S_t, S_{t+1}, \ldots, S_T, \mathcal{Y}_{t+j-1} \quad \sim \quad \mathrm{N}\left(\mu_{S_{t+j}} - \textstyle\sum_{i=1}^{k} \phi_i(y_{t+j-i} - \mu_{S_{t+j}}), \sigma^2_{S_{t+j}}\right)$$

for $j = 1, ..., T-t$. Hence, $p(\mathcal{Y}_{t+1,T}|S_t, S_{t+1}, \ldots, S_T, \mathcal{Y}_t)$ is independent of S_t, which implies $p(S_t|S_{t+1}, \ldots, S_T, \mathcal{Y}_T) = p(S_t|S_{t+1}, \ldots, S_T, \mathcal{Y}_t)$ from (A.4). Next, by a similar transformation as above,

$$p(S_t|S_{t+1}, \ldots, S_T, \mathcal{Y}_t) = \frac{p(S_{t+2}, \ldots, S_T|S_{t+1}, S_t, \mathcal{Y}_t)}{p(S_{t+2}, \ldots, S_T|S_{t+1}, \mathcal{Y}_t)} \cdot p(S_t|S_{t+1}, \mathcal{Y}_t).$$

Given S_{t+1} and $\mathcal{Y}_t$, $S_{t+2}, \ldots, S_T$ will be independent of S_t by the Markov assumption on the state sequence, from which equation (A.1) follows.

A.3 Generating $\boldsymbol{\theta}_2 = (p_{11}, p_{22})$ from $p(\boldsymbol{\theta}_2|\boldsymbol{\theta}_1, \boldsymbol{\theta}_3, \boldsymbol{\theta}_4, \boldsymbol{\theta}_5, \mathcal{Y}_T)$

Since given the sequence of states $\boldsymbol{\theta}_1$, p_{11} and p_{22} are independent of the data set and the model's other parameters, $\boldsymbol{\theta}_2$ may be drawn conditional only on $\boldsymbol{\theta}_1$. Based on standard Bayesian results on Markov chains, Albert and Chib (1993) derive the following likelihood function:

$$L(\boldsymbol{\theta}_2|\boldsymbol{\theta}_1) = p_{11}^{n_{11}}(1-p_{11})^{n_{12}} p_{22}^{n_{22}}(1-p_{22})^{n_{21}},$$

where n_{ij} refers to the number of transitions from state i to state j contained in the sequence $\boldsymbol{\theta}_1$. Considering the form of the likelihood function, we assume two independent beta distributions for the priors of p_{11} and p_{22}:

$$\begin{aligned} p_{11} &\sim \mathrm{Beta}(\alpha_1, \gamma_1) \\ p_{22} &\sim \mathrm{Beta}(\alpha_2, \gamma_2), \end{aligned}$$

where α_1 and γ_1 (α_2 and γ_2) are known parameters. By the independence assumption, the prior distribution is given by

$$g(\boldsymbol{\theta}_2) \propto p_{11}^{\alpha_1-1}(1-p_{11})^{\gamma_1-1} p_{22}^{\alpha_2-1}(1-p_{22})^{\gamma_2-1}.$$

Combining the priors and the likelihood function, we obtain

$$p(\boldsymbol{\theta}_2|\boldsymbol{\theta}_1) \propto p_{11}^{\alpha_1+n_{11}-1}(1-p_{11})^{\gamma_1+n_{12}-1} p_{22}^{\alpha_2+n_{22}-1}(1-p_{22})^{\gamma_2+n_{21}-1}.$$

Thus, p_{11} and p_{22} may be sampled from the (independent) posterior distributions

$$\begin{aligned} p_{11}|\boldsymbol{\theta}_1, \boldsymbol{\theta}_3, \boldsymbol{\theta}_4, \boldsymbol{\theta}_5, \mathcal{Y}_T &\sim \mathrm{Beta}(\alpha_1^*, \gamma_1^*), \\ p_{22}|\boldsymbol{\theta}_1, \boldsymbol{\theta}_3, \boldsymbol{\theta}_4, \boldsymbol{\theta}_5, \mathcal{Y}_T &\sim \mathrm{Beta}(\alpha_2^*, \gamma_2^*), \end{aligned}$$

where

$$\alpha_1^* = \alpha_1 + n_{11}, \qquad \gamma_1^* = \gamma_1 + n_{12}$$

and

$$\alpha_2^* = \alpha_2 + n_{22}, \qquad \gamma_2^* = \gamma_2 + n_{21}.$$

A.4 Generating $\boldsymbol{\theta}_3 = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2)$ from $p(\boldsymbol{\theta}_3|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_4, \boldsymbol{\theta}_5, \mathcal{Y}_T)$

We may conveniently rewrite our model for $t = 1, 2, ..., T$ as

$$y_t - \sum_{i=1}^{k} \phi_i y_{t-i} = \mu_{S_t} \left[1 - \sum_{i=1}^{k} \phi_i \right] + \varepsilon_t,$$

and, letting $\mathbf{1}_{[\cdot]}$ denote the indicator function and defining

$$\mathbf{x}_t \equiv \left[1 - \sum_{i=1}^{k} \phi_i \right] \begin{bmatrix} \mathbf{1}_{[S_t=1]} \\ \mathbf{1}_{[S_t=2]} \end{bmatrix},$$

as

$$y_t - \sum_{i=1}^{4} \phi_i y_{t-i} = \mathbf{x}_t' \boldsymbol{\theta}_3 + \varepsilon_t.$$

Dividing both sides of the preceeding equation by the known standard deviation, σ_{S_t}, leads to

$$y_t^* = \mathbf{x}_t^{*\prime} \boldsymbol{\theta}_3 + \varepsilon_t^* \qquad t = 1, 2, ..., T,$$

where

$$y_t^* \equiv \frac{1}{\sqrt{\sigma_{S_t}^2}} \left(y_t - \sum_{i=1}^{4} \phi_i y_{t-i} \right)$$

and

$$\mathbf{x}_t^* \equiv \frac{1}{\sqrt{\sigma_{S_t}^2}} \mathbf{x}_t.$$

The model can be expressed more compactly in matrix notation as

$$\mathbf{Y}^* = \mathbf{X}^* \boldsymbol{\theta}_3 + \boldsymbol{\varepsilon},$$

where $\mathbf{Y}^* \equiv (y_1^*, y_2^*, \ldots, y_T^*)'$ and $\mathbf{X}^* \equiv (\mathbf{x}_1^*, \mathbf{x}_2^* \ldots, \mathbf{x}_T^*)'$. Note that this model satisfies the assumptions of the classical regression model, since $\boldsymbol{\varepsilon} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{I}_T)$.

Following standard practice, we assume a bivariate normal distribution for the priors of μ_1 and μ_2:

$$\boldsymbol{\theta}_3 \sim \mathrm{N}(\boldsymbol{\mu}_{\boldsymbol{\theta}_3}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}_3}),$$

where $\boldsymbol{\mu}_{\boldsymbol{\theta}_3}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\theta}_3}$ are known parameters.

Then, from standard Bayesian statistics, the posterior distribution from which the means are sampled is given by

$$\boldsymbol{\theta}_3|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_4, \boldsymbol{\theta}_5, \mathcal{Y}_T \sim \mathrm{N}(\boldsymbol{\mu}^*, \boldsymbol{\Sigma}^*),$$

where

$$\boldsymbol{\Sigma}^* = \left(\boldsymbol{\Sigma}_{\boldsymbol{\theta}_3}^{-1} + \sum_{t=1}^{T} \mathbf{x}^* \mathbf{x}_t^{*\prime} \right)^{-1},$$

$$\boldsymbol{\mu}^* = \boldsymbol{\Sigma}^* \left(\boldsymbol{\Sigma}_{\boldsymbol{\theta}_3}^{-1} \boldsymbol{\mu}_{\theta_3} + \sum_{t=1}^{T} \mathbf{x}_t^* y_t^* \right).$$

We then draw (μ_1, μ_2) from the above bivariate normal distribution. To meet the normalization constraint $\mu_1 < \mu_2$, we employ rejection sampling: We accept the draw if the constraint is met, otherwise it is discarded and resampled.

A.5 Generating $\boldsymbol{\theta}_4 = (\boldsymbol{\sigma}_1^{-2}, \boldsymbol{\sigma}_2^{-2})$ from $p(\boldsymbol{\theta}_3|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3, \boldsymbol{\theta}_5, \mathcal{Y}_T)$

We shall assume that the prior distributions for the inverse of the variances (σ_1^{-2} and σ_2^{-2}, respectively, also called the 'precisions') are independently Gamma:

$$\begin{aligned} \sigma_1^{-2} &\sim \Gamma(N_1, \lambda_1) \\ \sigma_2^{-2} &\sim \Gamma(N_2, \lambda_2), \end{aligned}$$

where N_1 and λ_1 (N_2 and λ_2) are known parameters.

Observe that conditioning on $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3, \boldsymbol{\theta}_5$, and $\mathcal{Y}_T$ is equivalent to observing

$$\varepsilon_t = (y_t - \mu_{S_t}) - \sum_{i=1}^{k} \phi_i (y_{t-i} - \mu_{S_t}) \qquad t = 1, 2, ..., T.$$

Let

$$\begin{aligned} \text{SSR}_1 &\equiv \sum_{t=1}^{T} \varepsilon_t^2 \mathbf{1}_{[S_t=1]}, \qquad n_1 \equiv \sum_{t=1}^{T} \mathbf{1}_{[S_t=1]}, \\ \text{SSR}_2 &\equiv \sum_{t=1}^{T} \varepsilon_t^2 \mathbf{1}_{[S_t=2]}, \qquad n_2 \equiv \sum_{t=1}^{T} \mathbf{1}_{[S_t=2]}, \end{aligned}$$

such that, for every state $i = 1, 2$, SSR_i denotes the sum of squared residuals and n_i the number of times that the state sequence has been seen in that particular state.

Then, from standard Bayesian statistics, the (independent) posterior distributions from which the 'precisions' are sampled are given by

$$\begin{aligned} \sigma_1^{-2}|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3, \boldsymbol{\theta}_5, \mathcal{Y}_T &\sim \Gamma(N_1 + n_1, \lambda_1 + \text{SSR}_1), \\ \sigma_2^{-2}|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3, \boldsymbol{\theta}_5, \mathcal{Y}_T &\sim \Gamma(N_2 + n_2, \lambda_2 + \text{SSR}_2). \end{aligned}$$

A.6 Generating $\boldsymbol{\theta}_5 = (\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_k)$ from $p(\boldsymbol{\theta}_5|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3, \boldsymbol{\theta}_4, \mathcal{Y}_T)$

Proceeding exactly as above, we may rewrite our model as

$$y_t^* = \mathbf{x}_t^{*\prime} \boldsymbol{\theta}_5 + \varepsilon_t^*, \qquad t = 1, 2, \ldots, T$$

where

$$y_t^* \equiv \frac{1}{\sqrt{\sigma_{S_t}^2}} (y - \mu_{S_t})$$

and

$$\mathbf{x}_t^* \equiv \frac{1}{\sqrt{\sigma_{S_t}^2}} \begin{bmatrix} y_{t-1} - \mu_{S_t} \\ y_{t-2} - \mu_{S_t} \\ y_{t-3} - \mu_{S_t} \\ y_{t-4} - \mu_{S_t} \end{bmatrix}.$$

Again, the model can be expressed more compactly in matrix notation as

$$\mathbf{Y}^* = \tilde{\boldsymbol{X}}^* \boldsymbol{\theta}_5 + \boldsymbol{\varepsilon},$$

where $\mathbf{Y}^* \equiv (y_1^*, y_2^*, \ldots, y_T^*)'$ and $\tilde{\boldsymbol{X}}^* \equiv (\mathbf{x}_1^*, \mathbf{x}_2^* \ldots, \mathbf{x}_T^*)'$. As before, this model satisfies the assumptions of the classical regression model, since $\boldsymbol{\varepsilon} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{I}_T)$.

Following standard practice, we assume a multivariate normal distribution for the priors of $\phi_1, \ldots, \phi_k$:

$$\boldsymbol{\theta}_5 \sim \mathrm{N}(\boldsymbol{\mu}_{\boldsymbol{\theta}_5}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}_5}),$$

where $\boldsymbol{\mu}_{\boldsymbol{\theta}_5}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\theta}_5}$ are known parameters.

Then, from standard Bayesian statistics, the posterior distribution from which the autoregressive coefficients are sampled is given by

$$\boldsymbol{\theta}_5 | \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3, \boldsymbol{\theta}_4, \mathcal{Y}_T \sim \mathrm{N}(\boldsymbol{\mu}^*, \boldsymbol{\Sigma}^*),$$

where

$$\boldsymbol{\Sigma}^* = \left(\boldsymbol{\Sigma}_{\boldsymbol{\theta}_5}^{-1} + \sum_{t=1}^{T} \mathbf{x}_t^* \mathbf{x}_t^{*\prime} \right)^{-1},$$

$$\boldsymbol{\mu}^* = \boldsymbol{\Sigma}^* \left(\boldsymbol{\Sigma}_{\boldsymbol{\theta}_5}^{-1} \boldsymbol{\mu}_{\boldsymbol{\theta}_5} + \sum_{t=1}^{T} \mathbf{x}_t^* y_t^* \right).$$

We then draw $\phi_1, \ldots, \phi_k$ from the above multivariate normal distribution. To meet the restriction that the autoregressive coefficients lie within the stationary region, we employ rejection sampling: We accept the draw if the roots of the polynomial in the lag operator lie outside the unit circle. Whenever the sampled values do not lie in stationary region, the draw is discarded and resampled.

References

Albert, James H., and Siddhartha Chib (1993) 'Bayes inference via Gibbs sampling of autoregressive time series subject to Markov mean and variance shifts.' *Journal of Business and Economic Statistics* 11(1), 1–15

Andrade, Gregor, Mark Mitchell, and Erik Stafford (2001) 'New evidence and perspectives on mergers.' *Journal of Economic Perspectives* 15(2), 103–120

Andrews, Donald W. K. (1993) 'Tests for parameter instability and structural change with unknown change point.' *Econometrica* 61(4), 821–856

Barkoulas, John T., Christopher F. Baum, and Atreya Chakraborty (2001) 'Waves and persistence in merger and acquisition activity.' *Economics Letters* 70(2), 237–243

Chow, Gregory C. (1960) 'Tests of equality between sets of coefficients in two linear regressions.' *Econometrica* 28(3), 821–856

— (1984) 'Random and changing coefficient models.' In *Handbook of Econometrics,* ed. Zvi Griliches and Michael D. Intrilligator, vol. 2 (Amsterdam: North-Holland) pp. 1213–1245

Chowdhury, Abdur R. (1993) 'Univariate time-series behaviour of merger activity and its various components in the United States.' *Applied Financial Economics* 3(1), 61–66

Clark, John J., Alok K. Chakrabarti, and Thomas Chiang (1988) 'Trend and stochastic movements in US merger activity.' *Quarterly Review of Economics and Business* 28(2), 6–19

Fauli-Oller, Ramon (2000) 'Takeover waves.' *Journal of Economics and Management Strategy* 9(2), 189–210

Gelfand, Alan E., and Adrian F. M. Smith (1990) 'Sampling-based approaches to calculating marginal densities.' *The American Statistician* 85(410), 398–409

Geman, Stuart, and Donald Geman (1984) 'Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images.' *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6(6), 721–741

Golbe, Devra L., and Lawrence J. White (1988) 'A time series analysis of mergers and acquisitions in the U.S. economy.' In *Corporate Takovers: Causes and Consequences,* ed. Alan J. Auerbach (Chicago: University of Chicago Press) pp. 265–305

— (1993) 'Catch a wave: The time series behavior of mergers.' *Review of Economic Studies* 75(3), 493–499

Gugler, Klaus, Dennis C. Mueller, and B. Burcin Yurtoglu (2005) 'The determinants of merger waves.' Working paper, University of Vienna

Hamilton, James D. (1989) 'A new approach to the economic analysis of nonstationary time series and the business cycle.' *Econometrica* 57(2), 357–384

— (1993) 'Estimation, inference and forecasting of time series subject to changes in regime.' In *Handbook of Statistics,* ed. G. S. Maddala, C. R. Rao, and H. D. Vinod, vol. 11 (Amsterdam: Elsevier) chapter 9, pp. 231–260

Harford, Jarrad (2005) 'What drives merger waves?' *Journal of Financial Economics* 77(3), 529–560

Holmstrom, Bengt, and Steven N. Kaplan (2001) 'Corporate governance and merger activity in the United States: Making sense of the 1980s and 1990s.' *Journal of Economic Perspectives* 15(2), 121–144

Hughes, Alan (1993) 'Mergers and economic performance in the UK: A survey of empirical evidence 1950–1990.' In *European Mergers and Merger Policy,* ed. Matthew Bishop and John Kay (Oxford: Oxford University Press) pp. 9–95

Jovanovic, Boyan, and Peter L. Rousseau (2002) 'The Q-theory of mergers.' *American Economic Review, Papers and Proceedings* 92(2), 198–204

Kim, Chang-Jin, and Charles R. Nelson (1999) *State-Space Models with Regime Switching* (Cambridge, Mass.: MIT Press)

Linn, Scott C., and Zhen Zhu (1997) 'Aggregate merger activity: New evidence on the wave hypothesis.' *Southern Economic Journal* 64(1), 130–146

Mitchell, Mark L., and J. Harold Mulherin (1996) 'The impact of industry shocks on takeover and restructuring activity.' *Journal of Financial Economics* 41(2), 193–229

Mueller, Dennis C. (1997) 'Merger policy in the United States: A reconsideration.' *Review of Industrial Organization* 12(5–6), 655–685

Mulherin, J. Harold, and Audra L. Boone (2000) 'Comparing acquisitions and divestitures.' *Journal of Corporate Finance* 6(2), 117–139

Nelson, Charles R., Jeremy Piger, and Eric Zivot (2001) 'Markov regime switching and unit root tests.' *Journal of Business and Economic Statistics* 19(4), 404–415

Nelson, Ralph L. (1959) *Merger Movements in the American Industry, 1895–1956* (Princeton: Princeton University Press)

Owen, Sian (2004) 'A Markov switching model for UK aquisition levels.' Working Paper, University of New South Wales, Sydney

Ravenscraft, David J. (1987) 'The 1980's merger wave: An industrial organization perspective.' In *The Merger Boom,* ed. L. E. Brown and E. S. Ronsengren Federal Reserve Bank of Boston Conference Series, No. 31, pp. 17–37

Resende, Marcelo (1999) 'Wave behaviour of mergers and acquisitions in the UK: A sectoral study.' *Oxford Bulletin of Economics and Statistics* 61(1), 85–94

Rhodes-Kropf, Matthew, and S. Viswanathan (2004) 'Market valuation and merger waves.' *Journal of Finance* 59(6), 2685–2718

Scherer, Frederic M., and David Ross (1990) *Industrial Market Structure and Economic Performance* (Boston: Houghton Mifflin Company)

Shleifer, Andrei, and Robert W. Vishny (1990) 'The takeover wave of the 1980s.' *Science* 249(4971), 745–749

Shughart, II, William F., and Robert D. Tollison (1984) 'The random character of merger activity.' *RAND Journal of Economics* 15(4), 500–509

Town, Robert J. (1992) 'Merger waves and the structure of merger and acquisition activities.' *Journal of Applied Econometrics* 7, Supplement: Special Issue on Nonlinear Dynamics and Econometrics, S83–S100

Toxvaerd, Flavio (2004) 'Strategic merger waves: A theory of musical chairs.' Working paper, Hebrew University of Jerusalem

Lebenslauf

Geboren am 16. Oktober 1976 in Konstanz

1987 – 1989	Primarschule in Kaiseraugst und Bernhardzell
1989 – 1991	Sekundarschule in Wittenbach
1991 – 1996	Kantonsschule St. Gallen, mathematisch-naturwissenschaftliche Abteilung
1997 – 2001	Diplomstudiengang Volkswirtschaftslehre am Fachbereich Wirtschaftswissenschaften der Universität Konstanz
2000	Gast am Sonderforschungsbereich 373 an der wirtschaftswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin
2002 – 2006	Doktorandenstudium Wirtschaftswissenschaften, Universität Zürich
2002 – 2003	„Swiss Program for Beginning Doctoral Students“, Gerzensee (Makroökonomik, Mikroökonomik und Ökonometrie)
2003	Doktorandenkurs „Time Series Econometrics“, Gerzensee
2005	Forschungsaufenthalt an der Universität Toulouse, Frankreich, als Stipendiat des Schweizerischen Nationalfonds