The Organization of Education

Dissertation

zur Erlangung des Grades
Doktor der Wirtschaftswissenschaften (Dr. rer. pol.)
am Fachbereich Wirtschaftswissenschaften
der Universität Konstanz

Gerald Eisenkopf

Griesseggstr. 21

78462 Konstanz

Konstanz, June 11, 2007

Vorwort

Das Urteil über diese Arbeit steht noch aus, aber ich bin froh und dankbar, dass ich es überhaupt so weit gebracht habe. Als ich 2002 nach Konstanz kam, war ökonomische Theorie in mathematischer Form nicht die nächstliegendste Methode für meine anstehende Forschung. In den viereinhalb vergangenen Jahren habe ich in Konstanz eine Einweisung in das ökonomische Denken erhalten, die mir einen neuen Zugang zu den mich interessierenden Problemstellungen geöffnet haben. Vielen Dank an alle, die mir dabei geholfen haben.

Diesen Fortschritt verdanke ich insbesondere meinem Doktorvater, Prof. Dr. Oliver Fabel. Seine konstruktive Kritik an meiner Arbeit, seine zahlreichen Denkimpulse und nicht zuletzt die freundschaftliche und vertrauensvolle Umgangsform, die er praktiziert hat, waren besonders förderlich. Thank you very much also to Claude Fluet who immediately agreed to evaluate this thesis.

Meine Kollegen am Lehrstuhl und am Fachbereich haben mir bei den zahlreichen Bergen auf dem Weg tatkräftig beigestanden. Benjamin Weigert half mir bei den ersten Rechenübungen, Christoph Safferling und Christian Lukas gaben stets ihr Bestes bei den zahlreichen großen und kleinen Fragezeichen.

Das Doktorandenprogramm des Fachbereichs erwies sich für mich als sehr hilfreich. Ich lernte schnell einen Einblick in verschiedenste Methoden und Fächer der Wirtschaftswissenschaften kennen und konnte so einige Rückstände aufholen. Nicht zuletzt gab es durch das Programm die Kommilitonen, die vor den gleichen Problemen standen und von deren Kooperation ich sehr profitieren konnte.

In den letzten Jahren hat mich meine liebe Frau Doris bei der Dissertation begleitet und moralisch unterstützt. Dafür und für die Geburt unserer Tochter Greta hat sie ihre eigene Berufsplanung zurück gestellt. Ihr gilt mein besonderer Dank. Meinen Eltern sei die Arbeit als Dank für ihre treue Unterstützung bei meiner langjährigen Ausbildung gewidmet.

Konstanz, 6. Februar 2007

Contents

List of Figures

Executive Summary

This dissertation is a collection of three stand-alone research papers. My research since 2002 has focused on the economics of education and organizational economics. The dissertation covers papers which link these two research areas. The introduction puts the papers in a more general context.

Chapter two introduces and discusses a mechanism which minimizes gaming or manipulation activities, if payments are linked to results from manipulable methods. The idea is to add non-manipulable information to manipulable information to improve the evaluation of a given output. A score declining in increasing evaluation quality indicates gaming. A simple agency model explains the advantages and disadvantages of the approach. The introduced mechanism dominates a single evaluation method if risk aversion is positive but not too great. The principal should let each agent decide about which evaluation method he prefers. The mechanism is applied to performance measurement in research organizations and hospitals. The problem of school accountability is discussed in the introduction.

Chapter three analyzes the impact of deregulation policies in higher education on the requirements for student input. Requirements decline if universities can choose the level of tuition fees (autonomous fees). If regulations keep tuition fees artificially low (regulated fees) or allow low ability students into higher education, universities increase requirements to deter undesired students. In a duopoly with regulated fees two ex-ante identical universities have identical requirements. Autonomous fee setting induces product differentiation. One university chooses high requirements and low tuition fees, the competitor low requirements and high fees. The chapter provides explanations for price-cost ratios in American universities, the differences in the industrial organization of higher education in the US and Europe, and the existence of profitable private universities with relatively low academic standards.

Chapter four discusses the impact of ability grouping in secondary education on student incentives. Education provides a signal on unobservable ability for employers and improves productivity after education. Selection sets better incentives in primary education and allows for improved peer group effects in secondary education. In comprehensive schools qualification has a greater impact on the employers' beliefs. Hence, students have a higher incentive to invest in qualification. The chapter provides an explanation why selective systems do not outperform comprehensive systems in comparative studies, even if standard peer effect assumptions hold. It also undermines the assumption that better scores with given observable inputs mean a higher efficiency level. Finally it shows why the measurement of value added - or differences-in-differences - in secondary education underestimates the impact of selection on performance.

Zusammenfassung

Diese Dissertation beinhaltet eine Zusammenstellung dreier eigenständiger Forschungspapiere. Meine Forschung seit dem Jahr 2002 hat sich auf Bildungsökonomik und Organisationsökonomik fokusiert. Die Papiere in der Dissertation verbinden beide Forschungsgebiete. In der Einleitung werden sie dabei in einen größeren Kontext eingebunden.

Kapitel zwei stellt einen Mechanismus vor, der Manipulationsaktivitäten minimiert, wenn Auszahlungen von den Ergebnissen manipulierbarer Bewertungsverfahren abhängen. Der Mechanismus beruht auf der Addition nicht manipulierbarer Information zu manipulationsanfälligen Informationen, um ein Bewertungsverfahren zu verbessern. Wenn das Ergebnis mit steigender Testgüte fällt, so weist das auf Manipulation hin. Im Rahmen eines einfachen Agency-Modells werden die Vor- und Nachteile des Ansatzes erklärt. Der vorgestellte Mechanismus dominiert jede individuelle Methode dominieren, so lange der Agent nur in geringem Maße risiko-avers ist. Der Prinzipal sollte jedoch jeden Agenten für sich über die Wahl des Evaluationsmethode entscheiden lassen. Als Anwendungsbeispiele werden Leistungsmessung in Forschungseinrichtungen und Hospitälern diskutiert.

Kapitel drei untersucht die Auswirkung von Deregulierungsmaßnahmen im Hochschulbereich auf die Anforderungen an studentische Beiträge (inputs). Diese

Anforderungen sinken, wenn die Universitäten die Höhe Studiengebühren selbst festlegen können (Gebührenautonomie). Werden die Studiengebühren durch Regulierungsmaßnahmen auf niedrigem Niveau gehalten werden (regulierte Gebühren) oder wenn Studenten mit niedrigen Fähigkeiten ein Recht auf einen Studienplatz haben, so werden die Anforderungen als Abschreckungsmaßnahme steigen. In einem Duopol mit identischen Gebühren werden ex ante identische Universitäten das selbe Anforderungsniveau wählen. Gebührenautonomie wird zur Produktdifferenzierung führen. Eine Universität wird hohe Anforderungen mit niedrigen Gebühren wählen, die andere niedrige Anforderungen mit hohen Gebühren. Das Papier liefert Erklärungen für Preis-Kosten-Verhältnisse an amerikanischen Universitäten, die Unterschiede in der industriellen Organisation des Hochschulsektors in Europa und den Vereinigten Staaten und die Existen profitabler Privatuniversitäten mit relativ niedrigen akademischen Standards.

In Kapitel vier werden die Auswirkung einer Selektion nach Fähigkeiten in der sekundären Bildungsstufe auf die Anreize für Schüler diskutiert. Bildung bietet sowohl ein Signal für Arbeitgeber bezüglich der unbeobachtbaren Fähigkeit eines Schülers als auch eine direkte Erhöhung der Produktivität nach der Schulzeit. Durch Selektion werden die Anreize in der Grundschule erhöht und gleichzeitig auch die Peer-Effekte in der Sekundarstufe verbessert. In nichtselektiven Bildungssystemen (Gesamtschulen) hat die erworbene Qualifikation aber eine höhere Signalwirkung. Deswegen haben die Schüler auch einen höheren Anreiz, sich anzustrengen. Der Abschnitt erklärt, warum Gesamtschulsysteme nicht schlechter abschneiden als selektive Systeme, selbst wenn man herkömmliche Annahmen zu Peer-Effekten berückschtigt. Außerdem zeigt der Abschnitt, dass ein besserer Leistungstest bei gegebenen finanziellen Inputs nicht unbedingt auf ein effizienteres Schulsystem schließen lässt. Schließlich zeigt das Papier, dass die Messung des Wertzuwachses in der Sekundarstufe allein den Einfluss der Selektion auf die Leistung unterschätzt.

Chapter 1

Introduction

Education is one of the most important determinants for individual careers and the development of nations. According to a meta-analysis by Groot and van den Brink (2000), the returns to required education are roughly 8 percent in the USA and Europe. The health of children also improves (Currie and Moretti,2003). The educational output may rise with increased expenditures and other quantitative improvements but the quality of education appears to be even more important. Hanushek and Wößmann (2007, p. 1) argue in their comprehensive review of the empirical literature that

> educational quality – particularly in assessing policies related to developing countries - is THE key issue [in promoting economic well-being]

The quality of education depends on the organization of education. This organization reflects underlying institutions and the educational production function. Examples for such institutions are school autonomy, tuition fees, school accountability, school choice, selection, and competition. Parameters in the educational production function are, among others, class and school size, family and social background, characteristics of teachers and fellow-students (the peer group), student effort and more.

This dissertation is a collection of three essays on the organization of education. Each essay focuses on some institutions and properties of the educational production function. This introductory chapter will put them in a more general context. Its first section explains the analyzed educational institutions and educational production function. All three papers are theoretical papers, but with reference to empirical problems. This link between the theoretical approach and the empirical literature is explained in greater detail in section two. Section three sketches normative implications of the positive analysis in the different chapters.

1.1 Institutions and Educational Production

The different papers analyze different institutions. Chapter two focuses on accountability, chapter three on autonomy and competition, while chapter four considers ability grouping. Chapters three and four make assumptions about educational production functions. More specifically, they refer to student ability, student inputs and externalities generated by students. Chapter four is more specific. Students provide effort as an input and the externality is a peer effect. Both production parameters share the characteristic that they are difficult to observe. Hence, it is also difficult to estimate their impact on educational production (see next section). The chapters show that educational production is shaped by institutions. More specifically, they show that externalities and student input respond to the institutional framework. These papers consider students as utility maximizing agents who respond rationally to institutions and subsequent incentives. This assumption may not be entirely surprising but has often been ignored in the literature (see also De Fraja and Landeras, 2006).

Chapter two focuses on incentive compatible performance measurement in areas where measurement is difficult and subject to manipulation. It provides a contribution to the literature on accountability in areas like health care, research or education. The paper in this chapter introduces and discusses a mechanism which minimizes gaming or manipulation activities, if payments are linked to results from manipulable methods. The idea is to add non-manipulable information to manipulable information to improve the evaluation of a given output. A score declining in increasing evaluation quality indicates gaming. A simple agency model explains the advantages and disadvantages of the approach. The introduced mechanism dominates a single evaluation method if risk aversion is positive but not too great. The principal should let each agent decide about which evaluation method he prefers. The mechanism is applied to performance measurement in research organizations and hospitals.

Schools face similar problems of performance measurement. Jacob (2005) shows that math and reading achievement increased sharply following the introduction of an accountability policy in the Chicago Public Schools in 1996-1997. His explanation for observed achievement gains refer to student effort and increases in test-specific skills. Therefore, accountability appears to be a fine idea but induces teaching to the test (see also Lazear, 2006). Schools, like research organizations or hospitals, face a measurement problem. Long term benefits and externalities are almost impossible to quantify. Recent studies identified evidence of externalities of education, e.g. reduced crime (Lochner and Moretti (2004)) or improved civic participation (Dee (2004); Milligan, Moretti, and Oreopoulos (2004)). Educational performance also depends on other agents than schools. For example, the socio-economic background and the genetic background also have an impact (see Plug and Vijverberg, 2003). Simplistic measures also induce gaming behavior. The mechanism in chapter two considers both the measurement and the gaming problem.

Chapter three analyzes the impact of deregulation policies in higher education. Being a key parameters in the educational production function, student input requirements depend on the institutional framework. Requirements decline if universities can choose the level of tuition fees (autonomous fees). If regulations keep tuition fees artificially low (regulated fees) or allow low ability students into higher education, universities increase requirements to deter undesired students. In a duopoly with regulated fees two ex-ante identical universities have identical requirements. Autonomous fee setting induces product differentiation. One university chooses high requirements and low tuition fees, the competitor low requirements and high fees. The chapter does not consider the impact of an explicit peer effect. However externalities generated by students drive the behavior of the university. In a way one can call the result an indirect peer effect. Students influence their fellow students via their impact on university policy.

The fourth chapter discusses the impact of ability grouping in secondary education on student incentives. Education provides a signal on unobservable ability for employers and improves productivity after education. Selection sets better incentives in primary education and allows for improved peer group effects in secondary education. In comprehensive schools qualification has a greater impact on the employers' beliefs. Hence, students have a higher incentive to invest effort in qualification. The paper represents one of the first theoretical contributions to the literature which include effort incentives for students into the analysis. It shows that selection or ability grouping has a non-trivial impact on the incentives of students. Selection increases the marginal productivity of the best students but predetermines labor market success. Acknowledging for effort provides a new interpretation of empirical results on educational performance. A higher performance is not necessarily more efficient. It might be caused by an inefficiently high level of student effort.

1.2 Theory and Evidence

All papers collected in this dissertation are theoretical papers, but with a strong reference to the empirical literature. This holds in particular for chapters four. Chapter two is a theoretical paper precisely because it addresses a measurement problem. If any principal, and subsequently any econometrician, could measure performance easily and without active distortion, the entire problem discussed in the chapter would be obsolete. Chapter three provides an explanation for price-cost ratios in US universities as documented in Winston (2004, see also Figure 3.1). Furthermore, it explains the differences in the industrial organization of higher education in the US and Europe, and the existence of profitable private universities with relatively low academic standards.

Chapter four includes a theoretical contribution to a largely empirical literature. Peer effects and student effort are hard to observe and estimate. In the case of peer effects, this problem has induced a large literature. Econometricians face three key problems when estimating peer effects, the most important methodological discussion of them is provided by Manski (1993). Nevertheless a large literature has been devoted to getting around them (e.g. Hoxby, 2000, McEwan, 2003, Hanushek et al., 2003, Cullen, Jacob and Levitt, 2003, these and others are summarized in Ammermüller and Pischke, 2006). As a first problem, most measures are endogenous. Take average performance of fellow class members as an example. If positive peer effects exist, the performance of any student should improve with the performance of his class members. However, the dependent variable in econometric models has an impact on the peer effect measure, as any observed student also has an impact on his fellow students. This problem could be ignored if an independent ability measure for any student was available which lacks in many data-sets. Secondly, students are not randomly assigned to their peer groups. Parents, schools or any other party decide where students enrol. Such a selection process precludes the identification of a counterfactual. How would a student have performed in a different peer group or as a single learner? Thirdly, teacher behavior and other environmental characteristics can change with the peer group composition. The case of changing input requirements in chapter three illustrates just this. A university will increase requirements if too many undesired students want to enrol. Meier (2004) makes a similar argument. The same teacher may teach the same topic in a different way, if the average ability or the ability distribution changes in a class. Chapter three in this dissertation how external effects generated by students influence the behavior of universities.

Student effort, on the other hand received little attention in the econometric literature. A lack of reliable effort data is probably the cause for this low attention. However, little attention does not mean that effort is meaningless. Notable exceptions are Jacob (2005) and de Fraja et al. (2006) who underline the importance of effort provision for educational performance.

In the theoretical economic literature, students with a higher marginal productivity are typically assumed to provide more beneficial external effects than low ability students (e.g. Lazear, 2001). A popular measure for such an assumption is average ability in a class (e.g. in Epple and Romano, 2003). The measure implies that high ability students can also benefit more strongly from better peer effects because their marginal productivity is greater. Such a pattern suggests that homogeneous learning groups are efficient. Students should be sorted according to ability. Now chapter four provides an explanation why selective systems do not outperform comprehensive systems in comparative studies (e.g. Hanushek and Wößmann (2006)), even if those peer effect assumptions hold.

Two further theoretical contribution to the empirical literature are discussed in greater detail within chapter four. The chapter undermines the assumption that better scores with given observable inputs mean a higher efficiency level. Unobservable effort inputs have to be taken into account as well. Finally it provides a critique to recent econometric publications. A differences-in-differences approache to the impact of selection in education like in Hanushek and Wößmann (2006) underestimates the impact of selection on performance. The subtrahend in these estimations is endogenous, as students in a selective system participate in a tournament to get into the better school.

1.3 Policy Implications

The papers in this dissertation provide a positive analysis of different aspects of the organization of education. But they also make statements about efficient allocations.

In chapter two the implications of the argument are straight-forward. With several evaluation scores attached to a given output, the quality and manipulability of each evaluation method has to be considered, too. A declining score at increasing evaluation quality indicates gaming. The proposed mechanism sets better incentives than the application of any single evaluation mechanism if agents are not too risk-averse. With risk-averse agents the principal can improve his situation by allowing each agent to choose between an individual evaluation method and the mechanism.

In Chapter three, competition with deregulated tuition fees leads to the most efficient outcome. In this case universities take the heterogenous preferences of prospective students into account and provide differentiated products. In chapter four ability grouping is the dominant solution, in spite a lack of evidence from achievement test (see above).

In those two last chapters however, the focus is on a parsimonious positive analysis. Important aspects of for a normative analysis are missing. In particular, each paper does not take external effects of education on agents outside the education sector into account. If external effects are positive and sufficiently large, the efficient allocation can imply a comprehensive educational system in chapter four. Only with such an externality, the benefit from greater performance would justify the extra costs of effort. If the externalities from education derive particularly from low ability students (chapter three), they would balance the heterogeneity in requirement preferences across students. In this case, uniform requirements across universities would be the efficient solution.

References

Ammermüller, Andreas and Jörn-Steffen Pischke. 2006. "Peer Effects in European Primary Schools: Evidence from PIRLS." ZEW Discussion Paper: Mannheim.

Cullen, Julie, Brian Jacob, and Steven Levitt. 2003. "The effect of school choice on student outcomes: Evidence from randomized lotteries." NBER Working Paper.

Currie, Janet and Enrico Moretti. 2003. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings." Quarterly Journal of Economics, 118:4, pp. 1495-532.

De Fraja, Gianni and Pedro Landeras. 2006. "Could do better: the effectiveness of incentives and competition in schools." Journal of Public Economics, 90, pp. 189–213.

De Fraja, Gianni, Tania Oliveira, and Luisa Zanchi. 2005. "Must try harder. Evaluating the role of effort on examination results." CEPR.

Dee, Thomas S. 2004. "Are there civic returns to education?"" Journal of Public Economics, 88:9-10, pp. 1697-720.

Epple, Dennis, Richard E. Romano, and Holger Sieg. 2003. "Peer effects, financial aid, and selection of students into colleges and universities: An empirical analysis." Journal of Applied Econometrics, 18:5, pp. 501-25.

Groot, W. and H. M. van den Brink. 2000. "Overeducation in the labor market: a meta-analysis." Economics of Education Review, 19:2, pp. 149-58.

Hanushek, Eric A., John F. Kain, Jacob M. Markman, and Steven Rivkin. 2003. "Does Peer Ability Affect Student Achievement?" Journal of Applied Econometrics, 18:5, pp. 18 (5), 527-44.

Hanushek, Eric A. and Ludger Wößmann. 2006. "Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence across Countries." Economic Journal, 116, pp. C63–C76.

Hanushek, Eric A. and Ludger Wößmann. 2007. "The role of school improvement in economic development." NBER.

Hoxby, Caroline M. 2000. "Does competition among public schools benefit students and taxpayers?" American Economic Review, 90:5, pp. 1209-38.

Jacob, Brian A. 2005. "Accountability, Incentives and Behavior: The Impact of High-Stakes Testing in the Chicago Public Schools." Journal of Public Economics 89:5-6, pp. 761-96.

Lazear, Edward P. 2001. "Educational Production." The Quarterly Journal of Economics, 116:3, pp. 777-803.

Lazear, Edward P. 2006. "Speeding, Terrorism, and Teaching to the Test." Quarterly Journal of Economics, 121:3, pp. 1029-61.

Lochner, Lance and Enrico Moretti. 2004. "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports." American Economic Review, 94:1, pp. 155-89.

Manski, Charles. 1993. "Identification of endogenous social effects: The reflection problem." Review of Economic Studies, 60:3, pp. 531-42.

McEwan, Patrick. 2003. ”Peer effects on student achievement: Evidence from Chile.” Economics of Education Review, 22:2, pp. 131-41.

Meier, Volker. 2004. ”Choosing between School Systems: The Risk of Failure.” Finanzarchiv, 60:1, pp. 83-93.

Milligan, Kevin, Enrico Moretti, and Philip Oreopoulos. 2004. ”Does education improve citizenship? Evidence from the United States and the United Kingdom.” Journal of Public Economics, 88:9-10, pp. 1667-95.

Plug, Erik and Wim Vijverberg. 2003. ”Schooling, Family Background, and Adoption: Is It Nature or Is It Nurture?” Journal of Political Economy, 111:3, pp. 611-41.

CHAPTER 2

Deterring Gaming with Imperfect Evaluation Methods

2.1 Introduction

Instruments of performance measurement rely on observable and often superficial indicators. From these indicators an evaluator hopes to conclude on the real performance. A link between payment and these measures may incidentally improve the desired output but channel inputs in wasteful, window-dressing activities. Agent will focus on improving the observable characteristics of the performance. Any principal gets what he pays for, which is not necessarily what he wants (Kerr, 1975). Considerable research has been carried out on the development of incentive compatible performance measures. The literature focuses on improving existent measures, combining and weighting them or introducing some measure of subjectivity into the measurement to align the incentives. This paper explains and discusses a different approach to combining performance measures. The assignment of negative weights to rather poor measures makes the difference.

The quality of performance measures is crucial for the delegation of problems from a principal to his agents. Moers (2006) shows empirically that, if financial performance measures are sensitive, precise and verifiable, then using these measures for incentive purposes increases delegation. Yet, in many cases, available measures fall short of these criteria. It appears obvious to improve and tighten existing measures as much as possible to restrict earnings management. Improved measures often imply a balancing of several measures into one score (Datar et al. 2001). Examples are the combination of accounting data with external market information (Dutta and Reichelstein, 2005) or the consideration of non-financial variables (e.g. Kaplan and Norton, 1992 and 1993, Dikolli, 2001, Sliwka, 2002, or Dutta and Reichelstein, 2003). However, tightened measures can lead to unintended substitution effects from accounting earnings management to real earnings management (Schipper, 2003, Ewert and Wagenhofer, 2005). Courty and Marschke (2003) recommend to take the dynamic characteristics of earnings management into account, as agents learn the mechanisms of the incentive contract much more precisely than the principal can anticipate. Hence, gaming increases over time and requires a change in the performance measurement.

Monitoring intensity can have behavioral consequences although experimental results do not suggest to reject conventional economic approaches. Dittrich and Kocher (2006) find experimental evidence for reciprocal behavior with low monitoring intensity, but employers do not benefit from relying on reciprocity. Nagin et al. (2002) investigate in a field how employees of a call center company respond to different monitoring rates. Some employees respond to a reduction in the perceived cost of opportunistic behavior by increased shirking but others do not respond at all to variations in the monitoring intensity. An empirical comparison of different theories about incentive compatible performance measurement with "real-life" data from companies implies serious econometric problems (see e.g. Prendergast, 1999).

Many contributions propose, as an alternative, a subjective element or some discretion in the measurement process (e.g. Baker et al. 1994, Murphy and Oyer, 2003, Liang, 2004, Rajan and Reichelstein, 2006). Of course, subjectivity and improved measurement are not mutually exclusive tools.

Gibbs et al. (2004) find evidence that subjective determination of bonuses increases with the complexity of tasks and the manager's tenure. The latter aspect is an indicator for trust between the principal and agent. Yet, subjectivity has its drawbacks. Krishnan et al. (2005) show that subjective performance measures can lead to suboptimal weights because of cognitive difficulties. Furthermore, discretion can induce favoritism and incalculable income risks (Ittner et al. 2003, see also Bergstresser and Phillipon, 2006).

This paper adds a different perspective to the combination of different measures. Developers of balanced performance measures typically adjust the weight of each individual measure according to its relevance or informativeness (e.g. Banker and Datar, 1989, Feltman and Xie, 1994, Datar et al. 2001). Rather poor measures receive a weak, if any weight, good and important measures a strong one. This paper explores the idea of assigning negative weights to poor measures. Hence, a negative weight in a poorly designed evaluation reduces the performance of the balanced or combined performance measure. Smith (2002) provides an earnings management model with non-financial performance measures which can assign a negative weight to some measure. In his case, consumer satisfaction may be rated negatively if long term objectives are not contractible and the agent boosts short term profits at the expense of long term profits.

The main reason for a negative weight is that poor performance measures are much more subject to manipulation than good ones. Hence, the poor measure may actually serve as a good measure for manipulation. Of course, this argument does not hold under all conditions. The paper discusses the limits of the approach in the context of a simple agency problem.

The analytic argument can be summarized as follow. Agents can game or manipulate the results of a poor performance measure. The principal can increase evaluation quality by considering more relevant information for the performance measurement. Hence, an evaluation instrument with high quality is an instrument which includes the most information into its analysis. It is costly to acquire information, of course. The principal looks at the relationship between quality and evaluation score and uses this relationship to calculate a score, which reveals the impact of undesired activities. Therefore, he uses the worse tests as information for manipulative behavior. To deter gaming completely, it is crucial that some information cannot be manipulated. However, the proposed process allows better determent of manipulation than any single evaluation method, even if specific information is subject to specific manipulation. In this case, a measure of subjectivity might bring further help, but the issue goes beyond the focus of this paper.

The proposed mechanism and the underlying principal-agent model are in many aspects similar to the proposed weighting of multiple performance measures in Datar et al. (2001). Indeed, this paper provides a simplification of Datar et al. (2001) to take account of notable differences. Datar et al. (2001) and many others in the literature (e.g. Baker 2002) claim that incentive-compatible performance measurement implies a trade-off between risk and accuracy. However, the risk declines if sufficient information is available at little cost. Adding information to a performance measure allows an identification of the statistical relationship between information availability and measured performance. Think of added information as the next hour or day, and the tools from time-series econometrics can be applied to forecast the performance measure if another unit of information is added. This stepwise increase in the considered information also implies that more sophisticated evaluation methods are less risky, as more available information typically means a more precise estimation. As an additional difference to Datar et al. (2001), the model considers the relationship of a principal with many agents with heterogeneous and unobservable abilities, which allows for self-selection processes among agents.

Two examples may explain the introduced mechanism. Research evaluators in Germany often use the number of publications as a proxy for quality, which is a cheap assessment method but dubious in terms of incentive compatibility (see Ursprung, 2003). Scientists are induced to publish countless papers with little if any new results if payments depend on such evaluations. Yet more papers are ceteris paribus better than less. Further measures such as citation impact, or the prestige of the journal could also be taken into account, for the price of higher evaluation costs. Manipulating these measures is difficult for most scientists. The proposed evaluation record would reveal the excessive output of meaningless papers. The number of published papers would get a negative weight but the number of papers in a peer-reviewed quality journal a positive and relatively large one.

In health care, policy makers ideally want hospitals and doctors to heal as many people at minimum cost. The output measurement is far from easy, as treatments often take a long time and interact with various factors which are beyond the control of any doctor. Any single, naive performance measure (e.g. the number of medical treatments, the number of successful treatments or the cost-per-treatment ratio) is subject to manipulation. Doctors may account minor consultations as full medical treatments, subscribe too many or too little medicine, or may be inclined to refer difficult cases to other hospitals. Reviewing simple measures in the context of more and better information will reveal how reliable the simple numbers are and indicate manipulation. The number of successful treatments for example has to be discounted, if the hospital or the physician deals with minor health problems only or if quick fixes have negative long term implications. In both examples the mechanism seems to be more suitable for organizations, e.g. entire research units or hospitals. Its application in individual performance based wage contracts would transfer relatively high risks to an agent.

The paper is structured as follows. The evaluation process and the proposed evaluation mechanism are introduced in the following section. Section 3 analyzes the application to an agency problem with risk averse agents. Section 4 tests the sensitivity of the results with respect to a change in the underlying assumptions.

2.2 The evaluation process

Evaluation constitutes an examination of the output of many agents by a single principal in a specified period. The mass of agents is standardized to one. Consider two methods, a good test T_g and a bad test T_b. The score of the good method reacts less sensitively to manipulation. Both methods are verifiable and use the same score scale.

The time structure is as follows. The principal chooses the evaluation mechanism and a wage contract. Then, the agents produce some output for one period. After production, the output is evaluated and payments are made. The principal evaluates honestly. If the principal would not evaluate agents would receive either a lump sum payment or average output. Given an infinite number of agents, both alternatives do not generate any incentive.

The principal can choose either method T_g, or method T_b, or a combination of both methods (T_c, with costs $K_c \geq K_g$). For simplicity, the evaluation score of agent i evaluated with method T is the sum of two linear functions and an error term.

$$Q_{i,T} = P_i + S_i(T) + \varepsilon_{i,T} \tag{2.1}$$

The two included functions are a desired-effort or "true output production" function

$$P_i = \theta_i e_i \tag{2.2}$$

and an undesired-effort, manipulation or "superficial output" function

$$S_i(T) = \gamma_T \mu_i s_i \tag{2.3}$$

The evaluation depends on characteristics of the agent and the properties of the evaluation method. The agent's ability in producing the desired output is denoted θ_i, and $e_i \geq 0$ is the desired effort. Agents have also an ability to manipulate, μ_i, and can choose some undesired manipulation effort $s_i \geq 0$. Manipulation thus constitutes an investment in superficial output, i.e. window-dressing activities. The agents are not homogeneous. Both abilities θ and μ are distributed across the agents according to identical independent uniform distributions, with $cov(\theta, \mu) = 0$ and

$$0 < \underline{\theta} = \underline{\mu} < \theta, \mu < \overline{\theta} = \overline{\mu}$$

The individual abilities are private knowledge, their distributions are common knowledge. To get explicit solutions, the cost functions of the effort inputs are specified as

$$C(e_i) = \frac{e_i^2}{2} \tag{2.4}$$

$$C(s_i) = \frac{s_i^2}{2}. \tag{2.5}$$

In this setting the costs of manipulation do not depend on the costs of actual production, and vice versa. This assumption allows for a simplification of the argument. The main results hold, even if, say, window-dressing had a positive impact on the actual output.

The measurement error is subject to some evaluation-specific and agent-specific observation error $\varepsilon_{i,T}$. Furthermore, the realisations of $\varepsilon_{i,T}$ are identically and independently distributed across agents as well as across evaluation methods.

$$\varepsilon_T \sim N(0, \sigma_T^2) \text{ with cdf } F_T(\varepsilon_{iT})$$

The respective cumulative distribution function is denoted with $F(\varepsilon_{i,T})$ and the density function with $f(\varepsilon_{i,T})$.

The chosen approach to information manipulation rests on the assumption that information acquisition is more expensive for information which is less prone to manipulation. If unmanipulated information was available on the cheap the problem of incentive compatible performance measurement would not exist.

Therefore, the properties of γ_T can be explained in the following way. Perfect evaluation requires a certain set of information units, but the costs for a perfect evaluation are prohibitively high. For simplicity, each unit of information receives a weight of one in the evaluation. The bad evaluation method T_b includes some of the necessary information. The information in this subset has two characteristics: It is very cheap to gather and it is easily manipulated. The variable γ_b denotes the degree of manipulability. The good evaluation method T_g represents an improvement on T_b. It includes the information units from T_b and a further subset of necessary information units. The units in this second subset are less prone to manipulation and more costly. Furthermore, the manipulability of all information in T_g which is not in T_b, is assumed to be zero. Then γ_g is equal to the product of γ_b and the share of information units from T_b among all information units in T_g. The costs for this method are higher as well ($K_g > K_b$).

2.2.1 *The proposed mechanism (combined evaluation)*

The principal and the agents know $0 < \gamma_g < \gamma_b$, i.e. to what extent each test is prone to manipulation. The principal calculates the combined evaluation score with the help of the function $G_c\left(Q_{i,1}, Q_{i,2}\right)$:

$$G_c\left(Q_{i,b}, Q_{i,g}\right) = Q_{i,c} = Q_{i,b} - \left(\frac{\gamma_b}{\gamma_b - \gamma_g}\right)\left(Q_{i,b} - Q_{i,g}\right) \tag{2.6}$$
$$= \left(\frac{\gamma_b}{\gamma_b - \gamma_g}\right) Q_{i,g} + \left(1 - \left(\frac{\gamma_b}{\gamma_b - \gamma_g}\right)\right) Q_{i,b}$$

The principal uses this function because it provides a consistent estimation of the "true" or desired output.

Lemma 2.1. *The desired output (P_i) is equal to the expected value of the combined evaluation.*

Proof. Notice that

$$E\left\{G_c\left(.\right)\right\} = \left(P_i + S_i(T_b) + \varepsilon_{ib}\right) - \left(\frac{\gamma_b}{\gamma_b - \gamma_g}\right)\left(P_i + S_i(T_b) + \varepsilon_{ib} - P_i - S_i(T_g) - \varepsilon_{ig}\right) \tag{2.7}$$

This equation can be transformed into

$$E\left\{G_c\left(.\right)\right\} = \left(P_i + \gamma_b \mu_i s_i\right) - \left(\frac{\gamma_b}{\gamma_b - \gamma_g}\right)\left(\gamma_b - \gamma_g\right)\mu_i s_i \tag{2.8}$$

which implies

$$E\left\{G_c\left(.\right)\right\} = P_i \tag{2.9}$$

□

The mechanism deters better than any single method even if information units in $T_g \ni T_b$ are subject to manipulation. At least all the manipulation of information in T_b is identified.

Lemma 2.1 requires that two evaluations have an intersect of information used. In particular, agents can invest in one type of unspecific manipulation only ($S_i(T_c) = \gamma_c \mu_i s_i$). This manipulation affects all information units in the bad test and some units in the good test. The qualification is important because if agents could manipulate a single evaluation score specifically while keeping all others constant, the results from Lemma 1 could not be maintained.

The expected value of the combined error term ε_c is given by

$$\varepsilon_{ic} = \left(\frac{\gamma_b}{\gamma_b - \gamma_g}\right) E\left(\varepsilon_{ig}\right) + \left(1 - \left(\frac{\gamma_b}{\gamma_b - \gamma_g}\right)\right) E\left(\varepsilon_{ib}\right) = 0 \tag{2.10}$$

Hence, the properties of ε_{ic} are as follows:

$$\varepsilon_{ic} \sim N(0, V(\varepsilon_{ic})) \text{ with cdf } f_c(\varepsilon_{ic})$$

with

$$V(\varepsilon_{ic}) = \left(\frac{\gamma_b}{\gamma_b - \gamma_g}\right)^2 \sigma_g^2 + \left(1 - \left(\frac{\gamma_b}{\gamma_b - \gamma_g}\right)\right)^2 \sigma_b^2 \tag{2.11}$$

Lemma 2.2. *(i) The variance of the combined error term is always larger than the variance of the best single evaluation method*

$$V(\varepsilon_c) > V(\varepsilon_g) > 0$$

(ii) The variance of the combined evaluation converges to infinity if the available evaluation methods are almost equal with respect to quality

$$\lim_{\gamma_g \to \gamma_b} V(\varepsilon_{ic}) = \infty \tag{2.12}$$

Proof. The variance of method T_g is σ_g^2. The difference in evaluation quality $\gamma_b > \gamma_g > 0$ implies $\left(\frac{\gamma_b}{\gamma_b - \gamma_g}\right) > 1$. Therefore, the first term in (2.11) is larger than σ_g^2 (i). Statement (ii) holds because of

$$\lim_{\gamma_g \to \gamma_b} \frac{\gamma_b}{\gamma_b - \gamma_g} = \infty \tag{2.13}$$

□

Lemma 2.1 states that, on average, the combined evaluation reveals manipulation. This revelation is caused by the negative weight for the less sophisticated test score. Lemma 2.2 shows that any principal applying the mechanism may face a trade-off between evaluation quality and risk if the possibility to improve a test is small (i.e. $\gamma_g \tilde{=} \gamma_b$). However, the risk of the combined evaluation can be lower than the risk of method T_b if an evaluation method can be substantially improved and $\sigma_b^2 > \sigma_g^2$ holds. Taking additional information into account allows to hedge against shocks, which do not affect all information symmetrically. Then, the actual trade-off is between measurement quality and measurement cost. Lemma 2.2 rests on iid assumptions for each method's error term. It may not hold if $cov\left(\varepsilon_g, \varepsilon_b\right)$ is positive and sufficiently large.

2.3 The Principal-Agent-Problem

This section discusses the incentive effects of the combined evaluation on risk-averse agents. The analysis focuses on a linear wage with a fixed payment α and a variable payment which is contingent on the evaluation score of method T.:

$$w = \alpha + \beta Q\left(\theta_i, \gamma_T, \mu_i\right) \tag{2.14}$$

A similar analysis could be done with other wage formulae but this would yield no meaningful differences with respect to the application of the combined evaluation mechanism.

The analysis of the incentive effects of the mechanism is structured as follows. At first, I identify the general trade-off for the principal which is associated with application of the combined evaluation. In this setting all agents are subject to the same linear wage. The second step shows that the principal can improve by letting the agents decide which evaluation method they want. Finally it is shown that T_c, but not the single measure T_g, is a screening instrument for θ.

The utility function of each agent is affected by this wage w, his effort supply and a risk premium $r\beta^2 V(\varepsilon_T)$. The degree of risk-aversion is given by r. The risk premium is a function of the size of the risky component of the payment, indicated by β. The component is risky because it depends on the evaluation score, which is associated with some uncertainty ($\varepsilon_{i,T}$).

$$\max_{e,s,T} U_i = w - C(e_i) - C(s_i) - r\beta^2 V(\varepsilon_{i,T}) \tag{2.15}$$

Other payment schemes are possible, of course. A discussion in Section D shows that the results on the evaluation mechanism are robust for bonus contracts and tournaments, too. Since this paper focuses on incentive compatible performance measurement rather than on optimal wage contracts, the formal analysis is restricted to linear wages.

Each agent invests in effort and manipulation to maximize his utility.

$$\max_{e,s} U_i = w - C(e_i) - C(s_i) - r\beta^2 V(\varepsilon_T) \tag{2.16}$$

with the respective first-order conditions:

$$\frac{\partial U}{\partial e} = \beta\theta_i - e_i^{**} = 0 \tag{2.17}$$

$$\frac{\partial U}{\partial s} = \beta\gamma_T\mu_i - s_i^{**} = 0 \tag{2.18}$$

The first-order-conditions (2.17) and (2.18) allow to calculate the incentive compatibility constraints in the principal-agent-model, with e_i^* and s_i^* denoting the optimal effort and manipulation levels of the agent:

$$\beta\theta_i = e_i^* \tag{2.19}$$

$$\beta\gamma_T\mu_i = s_i^* \tag{2.20}$$

The amount of each input is independent of the amount of the other input. The ratio of both inputs depends on the respective abilities of the agent and the evaluation quality.

$$\frac{e_i^*}{s_i^*} = \frac{\theta_i}{\gamma_T \mu_i} \tag{2.21}$$

A change in the incentive factor β causes proportional changes of both inputs. Only a change in the evaluation quality changes the relative effort supply via γ_T. Furthermore, equation (2.18) reveals that perfect evaluation makes manipulation efforts prohibitively costly for the agent. Of course, if the principal cannot not credibly commit himself to honest evaluation and payments the agents will not provide effort.

The subsequent analysis is built on the following assumptions. The outside option for each agent is equal to zero

$$\alpha - \beta Q_{i,T}(\theta, \gamma_T, \mu) - C(e_i) - C(s_i) - r\beta^2 V(\varepsilon_T)) \geq 0 \mid \forall (\theta, \mu) \tag{2.22}$$

Furthermore, the employer has to employ every agent. The latter condition ensures that wage contracts do not provide a screening instrument for the employer. It serves as a simplification of the analysis. The condition will be relaxed later in the paper.

Proposition 2.1. *Given a sufficiently low degree of risk-aversion among agents and a uniform linear wage for all agents, combined evaluation dominates the best single method (T_g).*

Proof. see Appendix □

The variable payment factor β decreases in the manipulability of the chosen evaluation method and in the associated risk. The combined mechanism is not subject to manipulation but implies a greater risk than method T_g. Therefore, if risk aversion is very high it is more suitable to apply a single method. Actually, the application of the combined mechanism is restricted on two sides. If risk aversion is low and the costs of testing exceed the risk premium, each agent becomes an individual entrepreneur and the incentive for manipulation vanishes.

Proposition 2.1 was based on a uniform linear contract for all agents. However, the principal can improve the situation by allowing for self-selection.

Proposition 2.2. *Let each agent choose between the combined evaluation method T_c or the best single evaluation method T_g. Assuming the principal sets optimal incentive parameters for each method, then*

1. *an agent with a high θ and a low μ will choose T_c;*
2. *an agent with a low θ and a high μ will choose T_g;*
3. *an agent will choose T_c if he was indifferent between T_g and T_c when evaluation methods were assigned by the principal;*
4. *the choice of evaluation methods by the agents dominates the assignment by the principal.*

Proof. First recall that the risk of an evaluation is constant for all agents regardless of their specific abilities. Then, the first two statements derive from (2.16) and the resulting relationship between expected income and associated risk. For the third statement, note that the agent was indifferent when all agents where evaluated in the same way. Due to the self-selection, the rewards of the best single method T_g

decline, as all agents with a high productivity and little manipulation ability choose against it. Self-selection dominates as incentives increase with productivity. □

The principal can take into account that people with different characteristics have a different perspective on a given risk. The good agents have less to fear from the combined evaluation and will choose it eventually. The "average" agents will join them, because otherwise they will be lumped together with the not so good and honest agents. Any agent with a relatively high productive ability will cross subsidize someone with a relatively high manipulative ability.

Lemma 2.3. *Assume the principle applies the best single method T_g. In this case a menu of linear wage contracts is not a screening instrument for the gaming ability μ of each agent.*

Proof. Linear wage contracts differ in α and β. Differences in α are irrelevant for efficiency considerations. The variable β denotes the share of the *observed* output which is paid to the agent. In the formula for observed output (2.1), true and manipulative output (P_i and S_i) are substitutes if $\gamma_T > 0$. Since $\gamma_g > 0$ and constant absolute risk aversion across agents the variable payment factor β chosen by an agent increases in $(\theta + \gamma_g \mu)$. Hence, productive and manipulative agents are indistinguishable. □

However, the application of the combined evaluation mechanism allows the principal to offer a menu of contracts. Each agent can choose the contract which is optimal for him because no agent has an incentive to manipulate. The variable payment factor β chosen by an agent increases only in (θ).

2.4 Variations of the model

In this section some of the underlying assumptions are relaxed to investigate how the results change with a modification of these assumptions.

2.4.1 Individual contracts

Evaluating the agents over time allows the principal to identify the abilities of the agents. Hence, he is able to design specific contracts for each agent. This knowledge affects the value of applying the most suitable single evaluation method.

Lemma 2.4. *The principal proposes individual contracts for each agent if he knows the specific abilities. With individual contracts, the combined evaluation becomes less attractive for agents with a high θ.*

Proof. In a wage contract with T_g as evaluation method, the more able and less manipulating agents subsidized their counterparts (see Lemma 2.3). This transfer payment is eliminated with individual contracts, which in turn makes the best single method more attractive for the more able and less manipulating agents.

The individual wage contract is provided in the appendix. □

As a consequence of this result, the combined evaluation mechanism is more attractive if the characteristics of the agents are unknown, because more productive agents cannot identify themselves while choosing T_g.

2.4.2 *Interdependence between the cost functions*

Previously, the costs of the different inputs were independent. However, marginal costs for one input may depend on the provision of the other input. Let $C_1(e_i, s_i)$ substitute $C(e_i)$ and replace $C(s_i)$ with $C_2(e_i, s_i)$ such that

$$\frac{\partial C_1(e_i, s_i)}{\partial e_i} = C_1' > 0; \frac{\partial^2 C_1(e_i, s_i)}{\partial (e_i)^2} = C_1'' > 0; \frac{\partial C_2(e_i, s_i)}{\partial s_i} = C_2' > 0; \frac{\partial^2 C_2(e_i, s_i)}{\partial (s_i)^2} = C_2'' > 0$$

hold.

Four relationships between the cost functions are possible:

$$\tfrac{\partial C_1'}{\partial s_i} < 0; \tfrac{\partial C_1'}{\partial s_i} > 0; \tfrac{\partial C_2'}{\partial e_i} < 0; \tfrac{\partial C_2'}{\partial e_i} > 0$$

Only the first of these relationships changes the case for the application of the combined evaluation mechanism qualitatively. If manipulation would increase costs for the desired input ($\frac{\partial C_1'}{\partial s_i} > 0$), it would strengthen the case for the combined mechanism. The third case ($\frac{\partial C_2'}{\partial e_i} < 0$) also increases the necessity to go again window-dressing. If the last case would hold, the entire discussion about gaming and manipulation would be obsolete.

Therefore, the focus is on the impact of $\frac{\partial C_1'}{\partial s_i} < 0$. Such a property changes the characteristics of what has been called manipulation. If more provision of this input decreases the costs for the desired input, then it is not entirely wasted. As an analogy, one may refer advertising, which is in itself unproductive but generates a positive externality on the desired objective. To stay in the research example, redundant publications by a researcher may spread his results wider than a single paper and induce more people to use the results for further research. Another label for such a relationship may be a trial-and-error approach. People work superficially on a lot of projects. Any of these projects may carry the big idea, so the pursuit of many projects may reduce the search costs for finding that good idea.

For any realistic consideration, $\frac{\partial^2 C'(e_i)}{\partial s_i \partial e_i} > 0$ and $\frac{\partial^2 C'(e_i)}{\partial s_i^2} > 0$ have to hold, otherwise the desired effort could be increased without limits.

Lemma 2.5. *Suppose increased manipulation decreases the marginal costs of effort (i.e. $\frac{\partial C_1'}{\partial s_i} < 0$ and $\frac{\partial C_2'}{\partial e_i} = 0$ hold). Then the combined evaluation mechanism does not deter agents from showing some manipulative effort.*

Proof. See appendix □

Agents will always game, even with the combined evaluation process being applied, because gaming facilitates the production of the desired output. Principals have a lower desire to deter agents from gaming, since it has some positive side effects. These external effects of manipulation has a further consequence. The evaluation quality from the best available single method is lower than in the case of $\frac{\partial C_1'}{\partial s_i} = 0$, i.e.

$$\left(\gamma_g \mid \frac{\partial C_1'}{\partial s_i} = 0\right) < \left(\gamma_g \mid \frac{\partial C_1'}{\partial s_i} < 0\right). \tag{2.23}$$

This statement holds because the term $\frac{\partial C_1'}{\partial s_i} < 0$ states that marginal costs decrease with manipulation. Increasing evaluation quality leads ceteris paribus to higher implicit costs. While the overall amount of gaming increases with positive external effects of manipulation, the difference between the combined mechanism and the best single method is not eliminated. Hence, the general trade-off for a principal remains as in the previous section.

2.4.3 *Change in the wage contract*

The mechanism was applied to an agency problem in which a principal paid linear wages. The proposed mechanism provides an efficient solution only if risk aversion is positive but not too high. Other wage formulas do not provide a different solution even if the set of possible payments is restricted. Since this paper does not primarily discuss the optimality of different wage formulae I do provide only a brief verbal description of the argument. In both bonus schemes or tournaments agents get a higher wage if they exceed a certain performance measure. In the case of bonus payments it is an absolute threshold, in tournaments with a finite number of agents it is a relative threshold.

In the most simple setting only two wages are paid, a high one for passing the threshold and a low one for failing to do so. This wage spread provides the incentives for effort supply. It also provides a certain insurance relative to a linear payment formula. On the other hand, increasing uncertainty in the production or measurement process reduces the incentives. Hence, the application of the proposed mechanism requires a greater wage spread than the best available single evaluation method to provide a similar incentive. Therefore, the same argument for the application of the mechanism applies, even if linear contracts do not provide the optimal payment scheme.

2.5 Conclusions

The paper has introduced an evaluation mechanism which eliminates incentives for manipulation even if individual methods cannot perfectly detect manipulation. The mechanism was based on the extrapolation of the perfect evaluation score from the scores of imperfect evaluations. Given full information about test quality and ability distributions, the "true" evaluation score can consistently be estimated.

The implications of the argument are straight-forward. If there are several evaluation scores attached to a given output, it is not the best idea to assign a positive weight to all relevant scores. The quality of the evaluation methods has to be considered, too. A declining score at increasing evaluation quality indicates window-dressing.

The proposed mechanism sets better incentives than the application of any single evaluation mechanism if agents are risk averse, but not too much. With risk-averse agents the selection of the evaluation method by each agent dominates uniform assignment by the principal. In this situation, the principal can improve his situation by allowing each agent to choose between an individual evaluation method and the combined mechanism. Finally, the application of the proposed mechanism is not restricted even if manipulative activities have a positive effect on the production process or linear wage contracts alone.

Appendix

Proof of Proposition 2.1.

The problem of the principal is given by

$$\max_{\alpha,\beta,T} \pi = \int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{\mu}}^{\overline{\mu}} \left(\theta e_i^* - K_T - \alpha - \beta Q_{i,T}\left(\theta, \gamma_T, \mu\right)\right) d\theta d\mu \tag{2.A.1}$$

The principal has to take the incentive compatibility constraints (2.19) and (2.20) into account. The principal sets the utility of the agent with the lowest expected income to zero. Such a worker is characterized by $\theta = \underline{\theta}$ and $\mu = \underline{\mu}$. The wage payments have to cover the costs and the risk premium for this agent. In our case of uniform contracts for all agents, this leads to the following necessary condition

$$\frac{\partial \pi}{\partial \beta} = \underline{\theta}^2 - \beta^* \left(\underline{\theta}^2 + \left(\gamma_{(T^*)} \underline{\mu} \right)^2 \right) - 2\beta^* r V(\varepsilon_T)) = 0 \tag{2.A.2}$$

Furthermore the following condition has to hold for the optimal method T^* relative to all other methods $T \neq T^*$

$$\int_{\underline{\mu}}^{\overline{\mu}} K_{T^*} - K_{(T \neq T^*)} + \left((\gamma_{T^*})^2 ((\beta^* \mid T = T^*)^2 - (\gamma^2_{(T \neq T^*)})(\beta^* \mid T \neq T^*)^2 \right) \mu_i^2$$
$$+ r \left(\beta^* \mid T = T^* \right)^2 V(\varepsilon_{T^*})) - r(\beta^* \mid T \neq T^*)^2 V(\varepsilon_{(T \neq T^*)}) d\mu < 0 \tag{2.A.3}$$

An optimal method T^* in (2.A.3) is characterized by the greatest spread between benefits, i.e. reduced payments for manipulative output: $-(\gamma_{T^*})^2 \left(\beta^* \mid T = T^* \right)^2 \mu_i^2$, and the respective costs $\left(K_{T^*} + r \left(\beta^* \right)^2 V(\varepsilon_{T^*}) \right)$ compared with all other available methods $T \neq T^*$. The optimal incentive weight for risk-averse agents is given by:

$$\beta^* = \frac{\underline{\theta}^2}{\underline{\theta}^2 + \left(\gamma_{T^*} \underline{\mu} \right)^2 + 2rV(\varepsilon_{i,T})} \tag{2.A.4}$$

Risk neutral agents would imply the elimination of risk premia and the full transfer of revenues to the agent.

If additional costs for the combined mechanism are low the combined evaluation is strictly dominant.

$$\lim_{K_c \to K_g} \int_{\underline{\mu}}^{\overline{\mu}} K_c - K_{(T \neq T_c)} - \left(\gamma_{(T \neq T_c)}\right)^2 \left((\beta^* \mid T \neq T_c)\right)^2 \mu_i^2 d\mu < 0 \tag{2.A.5}$$

A marginal increase in r does not affect this result. Therefore, a sufficiently low but positive level of risk aversion exists in which the combined mechanism dominates the best single mechanism. □

Individual contract in Lemma 2.4

The contract for an agent i solves the following problem

$$\max_{\alpha, \beta_i, T} \theta_i e_i - K_T - \alpha - \beta \left(Q_{i,T}\left(\theta, \mu\right)\right)$$

The incentive compatibility constraints (2.19) and (2.20) hold as well as the zero-profit condition for the agent. With risk averse agents, the incentive weight is

$$\beta_i^{**} = \frac{\theta_i^2}{\theta_i^2 + 2rV(\varepsilon_{T_c})}$$

for the combined evaluation and

$$\beta_i^{**} = \frac{\theta_i^2}{\theta_i^2 + \left(\gamma_T \mu_i\right)^2 + 2rV(\varepsilon_{T_g})}$$

for evaluation method T_g. The variable β_i^{**} denotes the optimal variable payment for an individual worker. Hence the incentive weight depends on the individual abilities. No cross subsidy from rather productive to rather manipulative agents takes place.

Proof of Lemma 2.5

For simplicity, I solve the case for an individual agent. The utility function for an agent i turns into

$$\max_{e,s} U_i = \alpha + \beta\left(Q_{i,T}\left(\theta\right)\right) - C_1(e_i, s_i) - C_2(e_i, s_i) - 2r\beta^2 V(\varepsilon_{T_g}) \tag{2.A.6}$$

with the incentive compatibility constraints

$$\frac{\partial U_i}{\partial e_i} = \beta\theta_i - C_1' - \frac{\partial C_2(e_i, s_i^{**})}{\partial e_i} = \beta\theta_i - C_1' = 0 \tag{2.A.7}$$

$$\frac{\partial U_i}{\partial s_i} = \beta\gamma_T\mu_i - C_2' - \frac{\partial C_1(e_i^{**}, s_i)}{\partial s_i} = 0 \tag{2.A.8}$$

Let s_i^{**} and e_i^{**} denote the incentive compatible inputs. The second of these conditions lead to $s_i^{**} > 0$ even for $\gamma_T = 0$. □

References

Baker, George. 2002. "Distortion and Risk in Optimal Incentive Contracts." Journal of Human Resources, 37:4, pp. 728-51.

Baker, G., R. Gibbons, and K. J. (1994) Murphy. 1994. "Subjective Performance Measures in Optimal Incentive Contracts." The Quarterly Journal of Economics, 109:4, pp. 1125-56.

Banker, R. D. and S. Datar. 1989. "Sensitivity, precision, and linear aggregation of signals for performance evaluation." Journal of Accounting Research, 27:1, pp. 21-39.

Bergstresser, D. and T. Philippon. 2006. "CEO Incentives and Earnings Management." Journal of Financial Economics, 80:3, pp. 511-29.

Courty, P. and G. Marschke. 2003. "Dynamics of Performance Measurement Systems." Oxford Review of Economic Policy, 19:2, pp. 268-84.

Datar, S., S. Cohen Kulp, and R. A. Lambert. 2001. "Balancing Performance Measures." Journal of Accounting Research, 39:1, pp. 75-92.

Dikolli, S. 2001. "Agent employment horizons and contracting demand for forward-looking performance measures." Journal of Accounting Research, 39:3, pp. 481-94.

Dittrich, D. and M. Kocher. 2006. "Monitoring and Pay: An Experiment on Employee Performance under Endogenous Supervision." Max Planck Institute of Economics.

Dutta, S. and S. Reichelstein. 2003. "Leading Indicators, performance measurement and long-term versus short-term contracts." Journal of Accounting Research, 41:5, pp. 837-66.

Dutta, S. and S. Reichenstein. 2005. "Stock Price, Earnings, and Book Value in Managerial Performance Measures." Accounting Review, 80:4, pp. 1069-100.

Ewert, R. and A. Wagenhofer. 2005. "Economic Effects of Tightening Accounting Standards to Restrict Earnings Management." Accounting Review, 80:4, pp. 1101-24.

Feltham, G. and J. Xie. 1994. "Performance measure congruity and diversity in multi-task principal-agent relations." Accounting Review, 69:3, pp. 429-53.

Gibbs, M., K. A. Merchant, W. A. Van der Stede, and M. E. Vargus. 2004. "Determinants and Effects of Subjectivity in Incentives " Accounting Review, 79:2, pp. 409-36.

Ittner, C. D., D F. Larcker, and M. W. Meyer. 2003. "Subjectivity and the Weighting of Performance Measures: Evidence from a Balanced Scorecard." Accounting Review, 78:3, pp. 725-58.

Kaplan, R. S and D. P. Norton. 1992. "The Balanced Scorecard - measures that drive performance." Harvard Business Review.

Kaplan, R. S and D. P. Norton. 1993. "Putting the Balanced Scorecard to work. ." Harvard Business Review, pp. 134-47.

Kerr, S. 1975. "On the Folly of Rewarding for a While Hoping for B." Academy of Management Journal, 18:4, pp. 769-83.

Krishnan, R., J. L. Luft, and M. D. Shields. 2005. "Effects of Accounting-Method Choices on Subjective Performance-Measure Weighting Decisions: Experimental Evidence on Precision and Error Covariance." Accounting Review, 80:4, pp. 1163-92.

Liang, P. J. 2004. "Equilibrium Earnings Management, incentive contracts, and accounting standards." Contemporary Accounting Research, 21, pp. 685-717.

Moers, F. 2006. "Performance Measure Properties and Delegation." Accounting Review, 81:4, pp. 897-924.

Murphy, K. J. and P. Oyer. 2003. "Discretion in Executive Incentive Contracts." University of Southern California.

Nagin, D. S., J. B. Rebitzer, S. Sanders, and L. J. Taylor. 2002. "Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment." American Economic Review, 92:4, pp. 850-73.

Prendergast, Canice. 1999. "The provision of incentives in firms." Journal of Economic Literature, 37, pp. 7-63.

Rajan, M. V. and S. Reichelstein. 2006. "Subjective Performance Indicators and Discretionary Bonus Pools " Journal of Accounting Research, 44:3, pp. 585-618.

Schipper, K. 2003. "Principle-based accounting standards." Accounting Horizons, pp. 61-72.

Sliwka, D. 2002. ”On the use of nonfinancial performance measures in management compensation.” Journal of Economics and Management Science, 11:3, pp. 487-511.

Smith, M. J. 2002. ”Gaming Nonfinancial Performance Measures.” Journal of Management Accounting Research, 14, pp. 119-33.

Ursprung, H. W. 2003. ”Schneewittchen im Land der Klapperschlangen: Evaluation eines Evaluators”.” Perspektiven der Wirtschaftspolitik:Band 2, pp. 177-89.

Chapter 3

Deregulation and Customer Input in Higher Education

3.1 Introduction

Education is a typical customer-input technology (Rothschild and White, 1995). The quality of the product depends on a student's inputs, e.g. his preparations for the courses and exams. A university can elicit these inputs, e.g. by setting academic standards or assigning workloads. More (or higher or tougher) requirements force students to study harder. More intensive studying is certainly beneficial for human capital output but comes at a cost. Hence, tougher requirements can also deter. This paper analyzes to what extent universities elicit student inputs in different regulatory settings.

Beside this 'customer-relationship management' the paper provides an explanation for different market structures. The differences between universities are arguably much smaller in Europe than in the United States where regulation is less prevalent. The hierarchical stratification in the US is well documented (e.g. Epple, D., Romano, R. E. and Sieg, H., 2003). This paper provides explanations for the emergence of elite universities in the US and for the existence of rather homogeneous universities in many European countries. Furthermore, it explains a rationale for the existence of profitable education businesses like the University of Phoenix. These schools deliberately position themselves on a lower academic level, responding to the demand of students for more "down-to-earth" courses.

The university is modeled as a surplus-maximizing risk-neutral firm in a regulated market. It benefits from the return on education of their students. A university certainly follows objectives beyond the profit motive. However, the surplus-maximization assumption implies a unit measure for the university's objectives. By speaking of surplus instead of profit, it is recognized that universities do not pay out dividends to shareholders. The results of Winston (1999, 2003) motivate the chosen surplus function of the university. Most papers quoted in this paper use similar approaches which build on an established tradition in the literature (e.g. Hansmann 1980; Clotfelter 1996, James 1990). The impact of research activities in a university is discussed at the end of the paper.

The argument can be summoned as follows. Since the university does not (directly) bear the costs of higher student inputs, it has an incentive to set input requirements very high. A program with more requirements provides more post-educational benefits to a student. However, preferences for the amount of input depend on the ability, as less able students prefer a less demanding program. In a monopoly scenario, an university uses high input requirements to deter undesired students. Input requirements are lower if the university can use tuition fees to deter those undesired students as well.

In a duopoly with regulated fees both universities choose symmetric input requirement levels. No other competitive equilibrium is achievable. In a competitive environment with autonomous tuition fee setting, universities differentiate with respect to the ability of their students, output per student, requirements in their educational programs and their tuition fees. Crucially, the better university offers lower fees, because high ability students provide more inputs and returns than low ability students. As a result, profitable low input universities emerge along side ”elite” institutions and access to higher education increases. This result implies increased inequality in educational output among the graduates.

These findings reflect some results observed by Hoxby (1997). Hoxby’s study of the US higher education sector shows that an increase in quality, fees, and product differentiation goes hand in hand with more competition between universities. However, *prima facie* evidence from American top universities seems to contradict the results of the duopoly with autonomous tuition fees. In these elite universities, tuition fees are the highest in the country and the fees increased dramatically over the last 20 years (see Ehrenberg, 2000).

The results in this paper hold because university inputs per student are set homogenously. The results on the tuition fees differ, if more demanding programs are also more expensive. Hence, a confirmation of the theoretical results requires price-cost ratios in education. Winston and Zimmerman (2000) as well as Winston (2004) report price-cost ratios which generally confirm the theoretical results of this article.

Figure 3.1 shows decreasing price-cost ratios with increasing subsidy per student. Universities in decile 1 pay the highest subsidy per student, universities in decile 2 pay the lowest. Arguably, better students go to schools with higher subsidies. That better universities have higher sticker prices for tuition fees reflects heterogeneity among universities with respect to their endowments and expenditure. This heterogeneity is not considered in the model in this paper.

The closest relations to this paper are the articles by Kemnitz (2005), Gary-Bobo and Trannoy (2005) and De Fraja and Iossa (2002). Kemnitz (2005) discusses incentive compatible university finance reform. His results for a duopoly with autonomous fees contradict the results of this paper and will be discussed more extensively below in the paper. De Fraja and Iossa (2002) model a non-cooperative game of two universities which choose their admission policy given the distance between the two universities and the transportion costs. In their model hierarchical stratification depends on mobility costs for students. High mobility costs induce symmetric behaviour of ex-ante identical universities, while intermediate costs create an asymmetric equilibrium. The authors do not find a pure strategy equilibrium given low

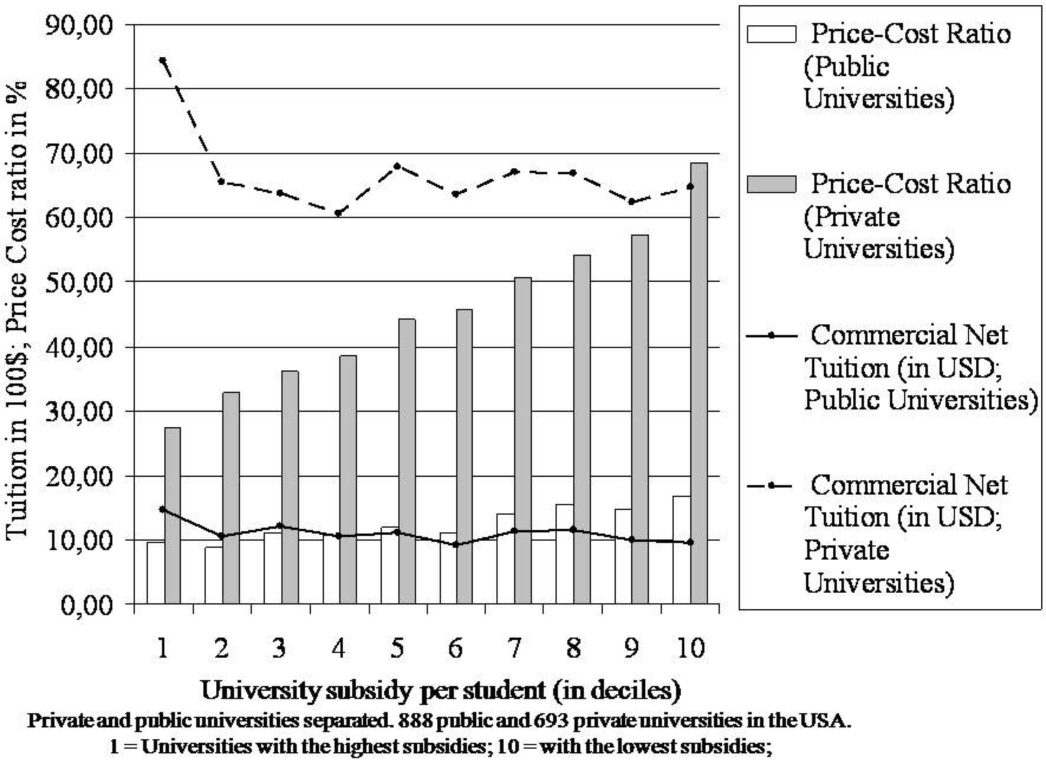

Figure 3.1: Fees and Price-Cost Ratios in US Higher Education - by Subsidy per student decile (Data from Winston (2004) pp 334-5)

mobility costs. Gary-Bobo and Trannoy (2005) discuss tuition fee-setting and admission standards in the context of asymmetric information on a student's ability. They show how universities use tuition fees for the selection of students, even if abilities are private information of the students. However, their focus is on welfare considerations, while this paper has a closer look on the behavior of universities and the industrial organization of higher education under different policy regimes.

The paper explains the existence of non-elite private schools in a new way. Previous explanations are based on business opportunities for private schools given an education policy for public schools (Brunello and Rocco (2005) or Martinez-Mora (2006)). This paper explains as a consequence of product differentiation between competitors.

This paper does not discuss agency problems within the university or between the university and its students (i.e. students dodging the input requirements). Agency problems within a university represent a cost. A comparative static analysis of changes in the cost parameter provides an implicit result for the impact of agency costs. Ortmann and Squire (2000) provide helpful contributions to this topic.

Furthermore, I assume that student actually provide the required input. De Fraja and Landeras (2006) analyze the relationship between educational competition and effort incentives for students. Competition may backfire on student effort provision as schools (in their model) compete for students. Particularly stratification in the market may be counterproductive. selection into a specific school already provides a credible signal on ability for the labor market. Students have to work less hard to obtain a signal than in a comprehensive schooling system (see also chapter 4 in this dissertation). This paper is complementary to these contributions as lower input requirements in a competitive environment are also unlikely to motivate most students for harder work.

The argument in this paper proceeds as follows: section 2 introduces the human capital production function and the model's agents. Section 3 shows how a university solves its maximization problem as a monopolist in different regulatory environments. Section 4 introduces competition. Section 5 discusses the results.

3.2 The model framework

3.2.1 The human capital production function

The university's output is measured by the value of the human capital of its graduates. This value is denoted with $V_{\theta k}$ for a student with ability θ at university k. The ability parameter θ is common knowledge. For parsimony it is uniformally distributed between 0 and 1. Of course, $V_{\theta k}$ increases *ceteris paribus* in θ. This paper only discusses vertical differentiation, e.g. all universities offer an economics program without further specialization, but with differences in the requirements for students. The variable s_k captures this level. It refers to the workload a student faces, e.g. the number and difficulty of tasks, the amount to study, thresholds for failure in exams.

Assumption 3.1 summarizes the properties of the human capital function in a simple equation. The program duration is standardized to one. The following assumption describes the human capital output:

Assumption 3.1.

$$V_{\theta k} = s_k \theta \tag{3.1}$$

For parsimony other productivity factors such as peer group effects, class size or teacher quality are ignored.

3.2.2 *The students' preferences*

Students are heterogenous regarding their ability. Given this heterogeneity the value of the acquired human capital differs among the students at university k. Every student at this university k has to pay the tuition fee T_k. The value of an alternative to higher education is zero for all students. While higher standards increase human capital they come with a cost: $\frac{s_k^2}{2}$. Hence, the utility of a student with ability θ at university k is given by:

$$U_k(\theta) = s_k\theta - \frac{s_k^2}{2} - T_k \tag{3.2}$$

The properties of the chosen approach allow to derive a single-crossing property:

Lemma 3.1. *For each student with ability θ there exists a unique preferred level of requirements $s_k^*(\theta)$.*

Proof. Assume a student faces the following hypothetical maximization problem. He can choose directly the requirement level of his university:

$$\max_{s_k} s_k\theta - \frac{s^2}{2} - T_k \tag{3.3}$$

Then, the preferred amount of requirements for a student with a given ability (s_k) reflects this ability

$$s_k^* = \theta \tag{3.4}$$

□

3.2.3 The universities' objective function

A university's objective is to maximize its surplus. I abstract from all principal-agent problems within in the university. Each university can offer only one program. The number of academic subjects is one.

The university generates revenues from tuition fees ($T_k \geq 0$) or invests in students via scholarships ($T_k < 0$). To simplify the analysis each university sets a single T_k for all of its students. Moreover, the university profits from its aggregated human capital output. These profits are not a tax but rather reflect spill-overs from the output. Examples of such spill-overs are handouts from donors who want to bask in reflected glory or alumni donations. Hence, the existence of appropriate fundraising mechanisms is assumed. In this paper the spill-overs are measured in money terms.

Assumption 3.2. *The spillover produced by each graduate of university k is a multiple of his human capital output ($\alpha V_{\theta k}$).*

The costs of educating a student are given by C. To simplify the analysis, these costs are identical across students, technologies and universities. The assumption of higher per capita revenues from increasing student ability substitutes some cost variations, such as decreasing costs with increasing ability of the students or decreasing returns to scale. Furthermore, the impact of increasing costs from increasing s_k is discussed in the paper.

Thus, we obtain the following surplus function for university k:

$$\pi_k = \int_{\underline{\theta}_k}^{\overline{\theta}_k} \left[\alpha V_{\theta k} + T_k - C\right] f(\theta) d\theta$$

The most and the least able students at the university are labelled with $\overline{\theta}_k$ and $\underline{\theta}_k$, respectively. Inserting from (3.1), the surplus function can be arranged to yield

$$\pi_k = \int_{\underline{\theta}_k}^{\overline{\theta}_k} \left[\alpha s_k \theta + T_k - C\right] f(\theta) d\theta \tag{3.5}$$

Taking the assumptions about the requirement preferences of students into account, the university's surplus maximization problem implies a trade-off between generating more graduates and receiving a higher revenue per student. More students mean more tuition fees and more graduates who will produce spill-overs. But more students also mean that expected spill-overs per student are lower because less able students do not enrol if the program is too demanding. The following two sections examines how universities solve this problem in different regulatory environments.

3.3 The monopoly case

Regulating competition is a key part of any political decision making in higher education. In Europe, universities can behave as a monopolist to a certain extent since they are shielded from competition and vicinity often dominates quality information when students choose their university. The analysis of a single may also help the reader to understand the strategic rationale of university in the chosen model framework.

Hence, it is useful to compare a monopoly with a duopoly situation as benchmark cases for regulation. Furthermore, the regulatory environment is restricted to analyzing university autonomy vs. government intervention in the fields of admission and tuition fees.

Definition 1. *Regulated admission means that the university has to enrol every applicant. If the university can select among applicants, there is autonomous admission.*

Admission regulation is a binding restriction for the university only if the university earns a negative surplus with the least able enrolled student

$$\pi_k(\underline{\theta}) < 0 \tag{3.6}$$

while the utility of these students from studying at the university is weakly positive

$$0 \leq U_k(\underline{\theta}) \tag{3.7}$$

Definition 2. *Regulated tuition fees mean T_k is given exogenously. If the university can choose T_k tuition fees are set autonomously.*

The paper analysises four regimes: Regulated and deregulated tuition fees in the monopoly and the duopoly case. Of course, each regime could go with or without regulated access. However, the results show that access regulation provides only quantitative differences.

3.3.1 Regime 1: Regulated tuition fees

In this first regime the university cannot deter undesired students with high tuition fees. Hence, the only available deterrence are the requirements for students. The university's problem can be stated as:

$$\max_{s_k} \pi_k = \int_{\underline{\theta}(s_k)}^{1} \left[\alpha V_{\theta k} + T_k - C\right] f(\theta) d\theta \tag{3.8}$$

For $\underline{\theta}(s_k)$, the participation constraint 3.7 holds.

The relevant first-order condition is

$$\frac{\partial \pi_k}{\partial s_k} = \int\limits_{\underline{\theta}(s_k)}^{1} \alpha V'_{\theta k} f(\theta) d\theta - \frac{\partial \underline{\theta}(s_k)}{\partial s_k} (\alpha V_{\theta k} + T_k - C) = 0 \tag{3.9}$$

This condition implies a trade-off. By raising s_k an university increases the revenue per student but loses students with relatively low ability. This effect allows a university to deter all undesired students. Hence, admission regulations are ineffective for a single university if the university can raise the requirements sufficiently high. As rule of thumb, tough exams substitute admission tests.

3.3.2 *Regime 2: Autonomous tuition fees*

Typically, regulations on tuition hold the fees at zero or at least artificially low. Deregulation in this area will also have an impact on requirements.

Lemma 3.2. *If a university is allowed to raise tuition fees and set their level, the university will demand less requirements than if tuition fees are abolished or held sufficiently low.*

Proof. This case is similar to the previous one. The university's optimization problem can be stated as

$$\max_{s_k, T_k} \pi_k = \int\limits_{\underline{\theta}(s_k, T_k)}^{1} \left[\alpha V_{\theta k} + T_k - C\right] f(\theta) d\theta \tag{3.10}$$

Again, the participation constraint 3.7 holds.

The first order conditions of the problem are:

$$\frac{\partial \pi_k}{\partial s_k} = \int\limits_{\underline{\theta}(s_k,T_k)}^{1} \alpha V'_{\theta k} f(\theta) d\theta - \frac{\partial \underline{\theta}(s_k)}{\partial s_k} (\alpha V_{\theta k} + T_k - C) = 0 \tag{3.11}$$

and

$$\frac{\partial \pi_k}{\partial T_k} = (1 - F(\underline{\theta})) - \frac{\partial \underline{\theta}(s_k, T_k)}{\partial T_k} T_k = 0 \tag{3.12}$$

Now suppose T_k was set to zero. Then condition 3.12 could not hold. The university has an incentive to raise the tuition fees. The increased tuition fee increases the costs of higher requirements from students (see 3.11). Hence, deregulated tuition fees imply a lower demand. □

The paper has shown so far the impact of a deregulation process for a monopolistic university. Regimes 1 reveals that (de-)regulating admission is an ineffective tool. The requirements on students are already an selection instrument. If the university wants to enrol more students it reduces these requirements anyway.

With autonomous tuition fees, the students compensate the university for the decline in per capita human capital output. In this case universities respond more strongly to the preferences of their students.

Note that the results do not depend on the chosen properties of the universities costs. Assume the cost parameter C was a function of s_k with $\frac{\partial C}{\partial s_k} > 0$ and $\frac{\partial^2 C}{\partial s_k^2} > 0$. Then lower requirement levels would be obtained. In regime 1 with regulated fees, the university may not be able to deter undesired students. A mere increase in the costs C, e.g. by accounting for agency costs, implies an increase in student requirement.

3.4 The duopoly case

3.4.1 *Regime 3: Regulated tuition fees*

The two competing universities 1 and 2 are ex ante identical, in particular regarding the assigned tuition fees (i.e. $T_1 = T_2$). Hence, they solve the same maximization problem.

$$\max_{s_k} \pi_k = \int_{\underline{\theta}_k(s_k)}^{\overline{\theta}_k(s_k)} \left[\alpha\left(V_{\theta k}\right) + T_k - C\right] f(\theta) d\theta \tag{3.13}$$

with $k \in \{1, 2\}$ and s.t.

$$U_k(\underline{\theta}_k) = V_{\underline{\theta}_k k} - \frac{s_k^2}{2} - T_k \geq 0 \tag{3.14}$$

The expression $\overline{\theta}_k$ denotes the ability of the most able student at school k, while $U_k(\underline{\theta}_k)$ represents the utility of the least able student at school k from his enrolment at this school.

In the case of an asymmetric equilibrium a second constraint exists at the boundary between the universities. Assume university 1 to require more from its students than university 2 ($s_1 > s_2$). Then, the least able students at the first university and the most able ones at the second one have the same utility. Hence, they are indifferent regarding the choice of their university:

$$U_2(\overline{\theta}_2) = U_1(\overline{\theta}_2) = U_1(\underline{\theta}_1) = U_2(\underline{\theta}_2) \tag{3.15}$$

The equation states that the utility of the most able student at school 2 from enrolment at this school is equal to the his utility from being enrolled at school 1. The same holds vice versa for the least able student at school 1, such that both types get an identical utility from enrolment in each school.

The applied analytical method is a simple single-period two-agents non-cooperative game. At the beginning of the one period, both universities simultaneously choose s_k. At the end of the period, the outcome for the universities is observed.

Proposition 3.1. 1. *The resulting equilibrium is symmetric.*

2. *Under the autonomous admission regime, both universities require less from their students than under the regulated one.*

Proof. See Appendix □

The universities cannot choose different requirement levels because of two reasons: Either the university with lower requirements has an incentive to increase the student inputs. Such a step attracts more high ability students from that other school. Alternatively, the university with a higher demand on students wants to decrease this share to get more students. One of both reasons is always applicable in case of an asymmetric outcome.

Compared with the respective monopoly cases, one can observe the effects of competition. The universities are induced to account for their students' preferences because they have to compete for them.

3.4.2 Regime 4: Autonomous tuition fees

In this case the maximization problem of both universities remains identical to the regulated environment but universities can use the additional choice variable T_k :

$$\max_{s_k,T_k} \pi_k = \int_{\underline{\theta}_k}^{\overline{\theta}_k} [\alpha\, V_{\theta k} + T_k - C]\, f(\theta) d\theta \tag{3.16}$$

with $k \in \{1, 2\}$. Assume for the moment that admission is not regulated. The maximization is subject to a zero utility constraint and a zero profit constraint (like equation (3.6)) for the least able student in each university:

$$U_k(\underline{\theta}_k) = V_{\underline{\theta}_k k} - \frac{s_k^2}{2} - T_k \geq 0 \tag{3.17}$$

$$\pi_k(\underline{\theta}_k) = \alpha V_{\underline{\theta}_k k} + T_k - C \geq 0 \tag{3.18}$$

The structure of the game is like in the previous subsection. Both universities simultaneously decide on their choice variables and receive the outcome at the end of the single period.

Proposition 3.2. *The resulting equilibrium is asymmetric. There exists an university* 1*, with a large* s_1*, low* T_1*, and a more able student body. And there exists another university* 2 *with a small* s_2*, high* T_2*, and less able student body.*

Proof. See appendix □

The logic behind the result is straightforward. Price competition alone reduces the surpluses to zero. With profits being at zero, each university likes to attract the better students and increase revenue per student by increasing s_k. Then low ability students are ready to pay a somewhat higher fee to face less requirements. Hence, given a sufficiently low T_1 and sufficiently high s_1, an alternative strategy 2 with relatively high fees and a low s_2 provides identical surplus. Product differentiation allows each university to make a positive surplus.

Regulated admission provides qualitatively similar results, although the chosen levels of requirements are higher for both universities. The quantitative difference stem from the fact that competition attracts undesired students. Hence, the alternative strategy 2 requires a much higher level of requirements at the 1st university to become profitable.

In terms of market outcome in a duopoly, Kemnitz (2005) gets a contradicting result. He identifies one large university with highly rewarding education and high tuition fees as well as one small university with a low rewarding education and low tuition fees. This small university can only stay in the market, because Kemnitz assumes that marginal costs of education quality increase in quality and admission thresholds are given exogenously. The result is driven by the assumption that all students prefer higher quality for a given tuition fee because failure rates are lower.

The result is plausible if more rewarding education means better instruction of a given content, though failure rates also depend on the peer group composition (see Meier, 2004). Unlike Kemnitz, this paper assumes that rewards depend on the requirements for customer inputs (and didactic quality remains unchanged). More requirements increase both rewards and cost of the education, but they are more harmful for low ability students.

3.5 Discussion and Conclusion

The paper has shown how the introduction of competition, a deregulation on tuition fees and the permission of university-defined admission standards shape a university's attitude towards its students. Deregulation induces the universities to pay more attention to the characteristics and preferences of a student. With the possibility to set tuition fees autonomously universities respond to the different demands from heterogeneous students by offering different educational products for different prices.

The characteristics of competitive outcomes depend on the regulations of tuition fees. Competition without autonomous tuition fees generates a symmetric equilibrium while autonomy on fees sets incentives for product differentiation. In particular, the model gives an explanation for the existence of profitable down-market universities next to elite schools. These universities charge higher fees but demand lower inputs.

The model in this paper takes two ex-ante identical universities as actors in order to show the effects of the deregulation in a country with rather homogenous providers of higher education, as it is the case in many European countries. In particular cost per students are standardized. Reported price-cost ratios (Winston, 2004) confirm the results.

The analysis was restricted to uniform tuition fees for all students within a university. This simplification does not change the qualitative results. Of course, each university could get a higher surplus if it charged each student until his utility from enrolling is just slightly better than his outside option. However, this represents just a redistribution. Each student will still face a choice between high input requirements and low fees on the one side and high fees and low requirements on the other side.

The competitive analysis was restricted to a duopoly case because it captures the main results. The logic behind proposition 3.2 also applies to a setting with more than two universities. Each universities will demand a different input for a different price. Prices will increase across universities with falling input requirements to satisfy the equal profit condition. The symmetric equilibrium in proposition 3.1. will not hold in the case of three or more universities, because a small increase (decrease) in requirements would attract all the high (low) ability students and capture half the profits. However, if any equilibrium exists with three or more universities and regulated tuition fees it will have less differentiation than a similar equilibrium with deregulated fees. At least two universities with the same input requirements will coexist because the university with the lowest requirements has an incentive to increase requirements to the level of the next university (see equation 3.A.1 in the appendix).

The model does not consider the impact of a university's research activities on the teaching policy. With respect to this ducal task function, Warning (2004) provides an interesting empirical insight into strategic positioning in German higher education. Research activities affect both the costs of teaching (via the opportunity costs) and the benefits for students, e.g. if better researchers offer students an insight into the state-of-the-art in their discipline. Then fees and input requirements would rise, if both universities are provided with strong researchers.

The different regimes described in this paper suggest that higher education in many European countries shows the properties of the symmetric equilibrium, while the regime 4 shows the US American system. With restrictions on tuition fees and admission thresholds in place, this regime 4 is ”ante portas” in Europe. Whether regime 4 will or should be implemented cannot be addressed by a mere positive analysis. In the simple model framework of this paper, it obviously is the efficient solution.

Appendix

Proof of proposition 3.1.

Suppose that an asymmetric equilibrium exists with $s_1 > s_2$ and $\pi_2 = \pi_1 \geq 0$. The constraint (3.15) is binding. The students at the boundary between the universities are indifferent. Therefore, the upper boundary at university 2 ($\overline{\theta}_2$) is a function of both the s_1 chosen by school 1 (s_1^*) and s_2 which implies the following optimization problem for university 2:

$$\max_{s_2} \pi_2 = \int_{\underline{\theta}_2(s_2)}^{\overline{\theta}_2\left(s_1^*,s_2\right)} \left[\alpha V(\theta, s_2) + T_2 - C\right] f(\theta) d\theta$$

With the endogenous upper boundary, we get

$$\frac{\partial \pi_2}{\partial s_2} = \alpha V_{\overline{\theta}_2 2} \frac{\partial \overline{\theta}_2}{\partial s_2} + \int_{\underline{\theta}_2(s_2)}^{\overline{\theta}_2\left(s_1^*,s_2\right)} \alpha V'(\theta, s_2) f(\theta) d\theta - \alpha V_{\underline{\theta}_2 2} \frac{\partial \overline{\theta}_2}{\partial s_2}. \tag{3.A.1}$$

This expression is strictly positive given the uniform distribution of θ. In case of a normal distribution it is possible that $\frac{\partial^2 \pi_2}{\partial (q_2)^2} < 0$. However, the reduction of the second university's incentive to converge to the standard of the first university means an increased incentive for the latter to converge. Additionally, it is strictly positive whether there exists a regulation on admission or not. The consequence for university 1 is a loss in profits and a reduction of s_1 as a reaction. Thus, the case of an asymmetric equilibrium can be excluded. (1st statement)

The resulting solution for $s_1 = s_2$ becomes smaller if $\pi_k(\underline{\theta}_k) < 0 \leq U_k(\underline{\theta}_k)$ does not hold or universities can regulate admission autonomously. A lower s_k is less harmful because it does not attract undesired students. (3rd statement)

Proof of proposition 3.2.

Assume a symmetric equilibrium with $0 < s_1 = s_2 = s < 1$, $T = T_1 = t_2$ and $\pi_2 = \pi_1 \geq 0$. Then $\frac{\partial \pi_k}{\partial T_k} < 0$ holds until $\pi_2 = \pi_1 = 0$ because a marginal reduction of the fees allows each university to attract all students. The least able are not enrolled since they create losses for the university.

At $\pi_2 = \pi_1 = 0$, $\frac{\partial \pi_k}{\partial s} > 0$ holds, since the high-ability students prefer a more demanding education. With increasing s and decreasing T the incentive to implement an alternative 2 with $s_2 < s$ and $T_2 > T$ increases. Equation (3.A.2) states the surplus of such an alternative 2. The notation s is replaced with s_1 and T with T_1.

$$\pi_2 = \int\limits_{\underline{\theta}_2(s_2,T_2)}^{\overline{\theta}_2(s_1,T_1,s_2,T_2)} \left[\alpha V(\theta, s_2) + T_2 - C\right] f(\theta) d\theta \tag{3.A.2}$$

The upper boundary is defined by the zero-profit condition for the least able student at university 1 $(\pi_1(\underline{\theta}_1) = 0)$, which again refers to the students' utility function (3.2). As an approximation one obtains $\overline{\theta}_2 = \underline{\theta}_1$, but $0 < U_2(\overline{\theta}_2) < U_1(\underline{\theta}_1)$. The least able students at university 1 are strictly better off. However, $\frac{\partial U_1(\underline{\theta}_1)}{\partial s_1} < 0$ holds.

The equilibrium is determined by four equations and two binding constraints for the lower boundaries which can be solved for $s_1, T_1, s_2, T_2, \underline{\theta}_2$ and $\underline{\theta}_1$. Two equations reflect the first order conditions if π_2 from equation (3.A.2) is maximized with respect to s_2 and T_2 and subject to the constraint (3.18):

$$\frac{\partial \pi_2}{\partial s_2} = \frac{\partial \overline{\theta}(s_2)}{\partial s_2} (\alpha V_{\theta 2} + T_2 - C) + \int\limits_{\underline{\theta}(s_2,T_2)}^{\overline{\theta}(s_2,T_2,s_1^*,T_1^*)} \alpha V'_{\theta 2} f(\theta) d\theta - \frac{\partial \underline{\theta}(s_k)}{\partial s_k} (\alpha V_{\theta k} + T_k - C) = 0 \tag{3.A.3}$$

and

$$\frac{\partial \pi_2}{\partial T_2} = \frac{\partial \overline{\theta}}{\partial T_2} T_2 + \left(F(\overline{\theta}_2) - F(\underline{\theta}_2)\right) - \frac{\partial \underline{\theta}}{\partial T_2} T_2 = 0 \tag{3.A.4}$$

Note that $\frac{\partial \overline{\theta}(s_2)}{\partial s_2} = \frac{\partial \overline{\theta}}{\partial T_2} T_2 = 0$ because $0 < U_2(\overline{\theta}_2) < U_1(\underline{\theta}_1)$ holds. A marginal change in requirement or tuition policy at university 2 does not induce anyone to leave university 1, because all admitted students are strictly better off at this university.

Constraint (3.14) is not binding for a university implementing alternative 2 as its problem is similar to the problem of a monopolistic university in the same regulatory environment (see equation (3.10)). The existence of positive profits in a monopoly situation guarantees an interior solution.

In equilibrium universities are indifferent between alternative 1 and alternative 2, if a marginal reduction of T_1 (say by $0 < \varepsilon_T$) and a marginal increase of s_1 (by $0 < \varepsilon_s$) yield the same surplus as the implementation of alternative 2. This means:

$$\lim_{\varepsilon_T \to 0} \pi_1 (s_1, T_1 - \varepsilon_T) = \pi_2 (s_2, T_2) \tag{3.A.5}$$

$$\lim_{\varepsilon_s \to 0} \pi_1 (s_1 + \varepsilon_s, T_1) = \pi_2 (s_2, T_2) \tag{3.A.6}$$

The two equations imply that, at $s_1 > s_2$ and $T_1 < T_2$, the surpluses in equilibrium are equal and positive:

$$\pi_1 (s_1, T_1) = \pi_2 (s_2, T_2) > 0.$$

The analysis of a duopoly with autonomous fees but regulated admission yields similar results. Regulated admission reduces the expected surpluses, students can join the university they prefer, irrespective of the losses for the university. Unlike in the monopoly case, tuition fees cannot be set high enough to deter these students. Such a scenario implies $\overline{\theta}_2 = \underline{\theta}_1$ and $U_{\overline{\theta}_2} = U_{\underline{\theta}_1}$.

Again, universities compete by lowering tuition fees and increasing the riskiness of its program. A marginal increase of s in one university induces the most able, and most surplus-generating, students to join the university and the less able and loss creating students to stay put in the other one. This implies opposite effects on the size of the student body. While lower tuition fees attracts more low-skill students, the increased requirements deters them.

References

Brunello, G. and L. Rocco. 2005. "Educational standards in private and public schools." CESifo.

Clotfelter, Charles T. 1996. *Buying the best. Cost escalation in elite higher education*. Princeton: Princeton University Press.

De Fraja, Gianni and Elisabetta Iossa. 2002. "Competition among universities and the emergence of the élite institution." *Bulletin of Economic Research*, 54:3, pp. 275-93.

De Fraja, Gianni and Pedro Landeras. 2006. "Could do better: the effectiveness of incentives and competition in schools." *Journal of Public Economics*, 90, pp. 189-213.

Ehrenberg, Ronald G. 2000. *Tuition rising. Why college costs so much*. Cambridge, MA: Harvard University Press.

Epple, Dennis, Richard E. Romano, and Holger Sieg. 2003. "Peer effects, financial aid, and selection of students into colleges and universities: An empirical analysis." *Journal of Applied Econometrics*, 18:5, pp. 501-25.

Gary-Bobo, Robert J. and Alain Trannoy. 2005. "Efficient tuition fees, examinations and subsidies." CEPR: 31.

Hansmann, Henry B. 1980. "The Role of Nonprofit Enterprise." *The Yale Law Journal*, 89:5, pp. 835-901.

Hoxby, Caroline M. 1997. "How the changing market structure of U.S. higher education explains college tuition." NBER: Cambridge, MA.

James, Estelle. 1990. "Decision processes and priorities in higher education," in "*The economics of American universities*" Stephen A. Hoenack and Eileen L. Collins eds. Albany: State University of New York Press, pp. 77-106.

Kemnitz, Alexander. 2005. "University finance reform, competition and teaching quality." University of Mannheim.

Martínez-Mora, Francisco. 2006. "The existence of non-elite private schools." *Journal of Public Economics*, 90, pp. 1505-18.

Meier, Volker. 2004. "Choosing between School Systems: The Risk of Failure." *Finanzarchiv*, 60:1, pp. 83-93.

Ortmann, Andreas and Richard Squire. 2000. "A game-theoretic explanation of the administrative lattice in institutions of higher learning." *Journal of Economic Behavior and Organization*, 43, pp. 377-91.

Rothschild, Michael and Lawrence J. White. 1995. "The analytics of the pricing of higher education and other services in which the customers are inputs." *Journal of Political Economy,* 103:3, pp. 573-86.

Warning, Susanne. 2004. "Performance Differences in German Higher Education: Empirical Analysis of Strategic Groups." *Review of Industrial Organization*, 24:4, pp. 393-408.

Winston, Gordon C. 1999. "Subsidies, hierarchy and peers: the awkward economics of higher education." *Journal of Economic Perspectives,* 13:1, pp. 13-36.

Winston, Gordon C. 2003. "Towards a theory of tuition: Prices, peer wages, and competition in higher education." Williams College - Wiliams Project on the Economics of Higher Education: Williamstown.

Winston, Gordon C. 2004. "Differentiation among US Colleges and Universities." *Review of Industrial Organization,* 24, pp. 331-54.

Chapter 4

Student Selection and Incentives

4.1 Introduction

Theoretical models with standard peer effect measures (e.g. average ability of peers) often imply that ability grouping in educational organizations is efficient. Yet robust empirical foundations for the measures and their implications are hard to come by. Hence, it is tempting to argue that the standard assumptions about peer effects are not correct.

The paper offers an explanation for the performance of comprehensive systems with a standard peer effect measure. It also shows that performance scores are not sufficient to judge the efficiency of an educational system, even with constant monetary inputs. The incentives for a student differ between a selective and a comprehensive system. Therefore, effort provision of students (or their parents) has to be taken into account as well. If selection takes place in secondary education, most students invest more effort in primary education to get to the better school. This result has implications for the interpretation of empirical results about tracking policies.

Many contributions to the economics of education make assumptions about how students influence their fellow students, the peer effects. Popular measures for the peer effect are average ability in a group (e.g. Epple et al., 2003), probability of interruption (Lazear, 2001) or homogeneity in a class (e.g. Dobbelsteen et al., 2002). These approaches typically imply that *ceteris paribus* the (self)selection of students in homogeneous groups -with respect to ability or the propensity to interrupt- is efficient and maximizes educational output (see e.g. Epple et al, 2003). Good students provide more positive external effects. They also get higher benefit from these effects because of higher marginal productivity.

Perfect ability grouping (or synonymously tracking, selection or sorting) may not lead to higher qualifications, even if the direct benefits from sorting are larger for the more able students than the respective losses for the less able students. Consider two main functions of education. Firstly, students acquire skills which increase their post-educational productivity in the labor market. More skills improve the post-educational wages, which induces students to put some effort into their education. Secondly, schooling provides a signal for the students' "innate" and unobservable abilities, as more able students are likely to have a better educational degree than less able ones. This "raw" ability also affects productivity, e.g. through the acquisition of job-specific human capital or general problem-solving skills. Hence, employers want to identify it and adjust wage offers accordingly.

However, sorting students according to (expected) educational results provides an information on ability before the education has been completed. Students in a good school are less likely to have low ability than students in a rather bad school. After selection, qualification will have less impact on the belief of an employer about the ability of student. Therefore, students have *ceteris paribus* a lower incentive to invest in education. A reduction in incentives can explain why potential gains in peer group effects from selection are so difficult to observe and why selective systems do not produce higher test scores.

Equal qualifications do not mean that tracking and non-tracking systems provide the same level of welfare. The paper shows that comprehensive schooling requires more effort from students to get the same level of qualification. Selection will always be the dominant solution if education does not generate some sufficiently large positive external effects.

Selection after some time (say in secondary education) has an impact on the incentives in a pre-selection period (i.e. comprehensive primary education). Sorting transfers incentives to the pre-selection period. If selection takes place in secondary education, most students invest more in primary education to get to the better school. A differences-in-differences methodology like in Hanushek and Wößmann (2006) or a similar approach discussed in Pischke and Manning (2006) underestimate the impact of selection on educational performance. A part of section 4 is dedicated to this measurement problems.

Econometric problems also restrict potential tests of the assumptions about peer effects. The problem in many empirical studies (e.g. Hoxby, 2000 & 2001, McEwan, 2003, Hanushek et al., 2003, Cullen et al., 2003, Ammermüller and Pischke, 2006) is the identification of exogenous ability measures to get around the so called "reflection problem" (see Manski, 1993). Then, the (self) selection of students into different peer groups does not allow to estimate what would have happened to the same student in another group. (Quasi-)Experimental studies may bring more insight but are still in short supply. Sacerdote (2001) provides evidence for peer effects in dorms at Dartmouth College, where assignment to the rooms is a random process. A first experiment with Swiss students shows that good students benefit more strongly from other good students than bad students do (Eisenkopf, 2007). If this result can be confirmed in further experiments, it supports the standard peer effect assumptions.

The discussion about optimal education policy has largely ignored incentives for students, although empirical studies show that incentives matter (De Fraja et al (2005), Kremer et al. (2004) with students from Kenya, Angrist and Lavy (2002) with Israeli student data). De Fraja and Landeras (2006) provide an inspiring theoretical contribution to the research of incentives in education. They show that increasing teacher incentives via more competition among schools may backfire on student incentives in a similar way. A greater emphasis has been put on institutional aspects, such as school choice and school competition (e.g. Hoxby, 2003), class size (West and Wößmann, 2005), peer group effects, or the sorting of students.

Furthermore, few studies on education policy consider the well established literature about education as a signal in the labor market (starting with Spence, 1973). Beside De Fraja and Landeras (2006), Brunello and Giannini (2005) provide an exception. Brunello and Giannini (2005) also show why neither perfect sorting nor comprehensive schooling provides a strictly dominant solution, although with an entirely different treatment of the signalling problem. For them (Brunello and Giannini, 2005, p. 190)

> ”[s]tratified systems trade the advantages of specialization and signalling against the disadvantages of producing skills with limited flexibility and versatility.”

In Brunello and Giannini (2005), the students cannot influence the admission result and the subsequent signal with more effort. ”Their” students have technical and academic abilities and schools provide different types of education. In contrast to Brunello and Giannini (2005), this paper provides a policy instrument to manipulate the sorting process. I do not discuss when students should optimally be separated. This problem is analyzed in Brunello et al. (2004). Finally, other reasons exist why a *sui generis* positive peer group effect can produce negative external effects on the outcome from ability grouping, e.g. higher failure rates for high-ability students (Meier, 2004).

To support the argument, section 2 introduces the analytic framework. Section 3 identifies a (not achievable) first-best solution and compares second-best outcomes. In section 4 intertemporal effects are analyzed. The section shows that selection in secondary education does not entirely reduce signalling incentives but shifts them to primary education. Section 4 also discusses the econometric implication of this finding. Section 5 derives empirically testable hypotheses from the model and concludes.

4.2 The model

Suppose the existence of one region with two schools $k \in \{a, b\}$ of equal size. Each school must enrol half of the students in the region. The mass of students in the region is normalized to one. Students are heterogeneous with respect to their ability θ. The ability is distributed symmetrically around $\hat{\theta}$ according to the differentiable distribution function $F(\theta)$, with $F'(\theta) = f(\theta)$. The ability of student i is known by the student, but not by the future employer of the student. Employers know the distribution of ability among the students. Students can invest in effort e to improve their educational qualification. The costs of effort follow the function $C(e)$, with $C'(e) > 0$ and $C''(e) > 0$. Only the student knows about his choice of the effort level.

Policy makers can choose between a comprehensive and a selective educational system, i.e. between sorting and integrating students. Under a sorting (or tracking) policy, high ability students go to one school and low ability students to the other one. Otherwise, the ability distribution is identical in both schools. For the moment it is assumed that sorting is perfect. The choice between a selective and a comprehensive system has an impact on average ability $\overline{\theta}_k$ in a school (the peer effect). Thus the contributions to educational output are defined: effort, ability and the peer effect.

Assumption 4.1. *A student's educational output or qualification is $q(e, \theta; \overline{\theta}_k)$, and satisfies $\frac{\partial q}{\partial e}, \frac{\partial q}{\partial \theta}, \frac{\partial q}{\partial \overline{\theta}_k} > 0$ as well as $\frac{\partial^2 q}{\partial e^2}, \frac{\partial^2 q}{\partial \theta^2} < 0$. All cross derivatives are strictly positive.*

Unlike in De Fraja and Landeras (2006) the observation of the qualification is distorted, i.e. the marks do not accurately reflect the actual skills.

Assumption 4.2. *A student's* observed *educational output or qualification is*

$$Q = q + \epsilon \tag{4.1}$$

The random variable ϵ is iid across students and follows the distribution $G(\epsilon)$ with the density function $g(\epsilon)$. The expected value of ϵ is zero and $g(\epsilon)$ is positive for all ϵ. The distribution is single-peaked and symmetric around the expected value.

Imperfect observation makes education risky. However, the paper does not discuss the implications of risk aversion but the impact of this imperfection on risk neutral students. Therefore, students and employers are risk-neutral. After education, a student enters the labor market.

Assumption 4.3. *A student's labor market output depends on his educational qualification (q) and his ability (θ). The labor market output is $\Pi(\theta, q)$, with $\frac{\partial \Pi}{\partial \theta}$, $\frac{\partial \Pi}{\partial q} > 0$, and $\frac{\partial^2 \Pi}{\partial \theta^2}$, $\frac{\partial^2 \Pi}{\partial q^2} < 0$. $\frac{\partial^2 \Pi}{\partial \theta^2}$ is sufficiently small to compensate any complementarity of qualification and ability on labor market output as well as any impact of a change in ability on qualification.*

Since employers cannot observe the ability θ they have to estimate θ and subsequently the expected labor market output from the observable qualification. The term $\mu(\theta; Q)$ denotes the density of an employer's belief about the ability of an individual whose observed qualification is Q. Belief formation takes place after the qualifications have been published. As in De Fraja and Landeras (2006), the analysis is restricted to a pure strategy equilibrium in which student effort does not decrease in ability. Hence, the employer formulates beliefs about a student's ability which are, on average, consistent with the actual ability.

In a competitive labor market initial wage offers for a student with observed qualification Q are equal to expected productivity $B(Q)$.

$$w = B(Q) = \int_{\theta} \int_{\epsilon} \Pi(\theta, Q - \epsilon)\mu(\theta; Q - \epsilon)g(\epsilon)d\theta d\epsilon \qquad (4.2)$$

Note that the expected wage offer for distorted qualification signals is lower than the offer for accurate ones. Labor market output $\Pi(\theta, q)$ is concave in both arguments and the actual qualification q again is a concave function of θ. The random variable ε, which distorts observed qualification Q, is symmetrically distributed around its expected value. Due to the concavity of the labor market output the employer's losses from $q < Q$ are greater than the gains from $q > Q$. Hence the employer has to offer a lower wage than in the case of perfectly observable qualifications. Otherwise the losses of students with a high ability from relatively poor observable qualifications ($Q < q$) would be smaller than the gains of less able ones from relatively good observable qualifications ($Q > q$).

If the employer can learn the actual ability some time after employment has started, any new wage offer reflects the actual output $\Pi(\theta, q)$. For simplicity reasons, such a possibility is not considered in the analysis.

4.3 To sort or not to sort?

The analysis starts with the comparison of the two extreme cases. In the comprehensive case (no sorting) students are randomly assigned to the two schools. In the (perfect) sorting case, the more able students go to one school, the less able to the other one. A perfect selection test does exist. I will discuss imperfect selection later in the paper.

4.3.1 No sorting

The ability composition in both schools is identical, which is denoted by the superscript ns. The respective peer effect is given by $\overline{\theta}_k^{ns}$, such that $\overline{\theta}_a^{ns} = \overline{\theta}_b^{ns} = \hat{\theta}$.

The objective of a student is to maximize the difference between the expected wage and his costs of effort by his choice of effort. The student correctly anticipates the wage formation of future employers and the effort supply of his fellow students. As stated above the social externality of his qualification is irrelevant to any student. The maximization problem is

$$\max_e B(Q^{ns}) - C(e) \tag{4.3}$$

with

$$B(Q^{ns}) = \int_\theta \int_\epsilon \Pi(\theta, Q^{ns} - \epsilon)\mu(\theta; Q^{ns} - \epsilon)g(\epsilon)d\theta d\epsilon$$

The transformation of the first order condition yields

$$\left(\frac{\partial B(Q^{ns})}{\partial Q^{ns}} + \frac{\partial B(Q^{ns})}{\partial \int_\theta \mu(\theta; Q^{ns})d\theta} \frac{\partial \int_\theta \mu(\theta; Q^{ns})d\theta}{\partial Q^{ns}} \right) \frac{\partial Q^{ns}}{\partial e} = C'(e) \tag{4.4}$$

The first summand shows the marginal increase in productivity, the second one the marginal change in the employers' beliefs. Better educational results increase the employers' expectation about the student's ability. However, all fellow students provide additional effort as well and the expected labor market output per student still increases strictly in ability.

4.3.2 Sorting

In a selective schooling system, the students with above-average ability ($\theta > \hat{\theta}$) go to school a, the others go to school b. Assume for simplicity that a costless test is available to identify perfectly if a student is above or below this threshold. Of course, after sorting peer effects at schools a and b have the following property: $\overline{\theta}_a^s > \overline{\theta}_b^s$. The superscript s indicates that student have been selected. Employers know that the ability of a student at a school a cannot be below $\hat{\theta}$ and adjust their beliefs accordingly. Now, the term $\mu(\theta; Q^s \mid \theta > \hat{\theta})$ denotes the density of an employer's belief about the ability of a student at school a whose observed qualification is Q^s. The density function is truncated at $\hat{\theta}$. For students at school a, the truncation is at the lower end. For students at school b, the truncation is at the upper end.

The maximization of a student at school a changes into:

$$\max_e \int_{\hat{\theta}} \int_{\epsilon} \Pi(\theta, Q^s)\mu(\theta; Q^s \mid \theta > \hat{\theta})g(\epsilon)d\theta d\epsilon - C(e) \tag{4.5}$$

Let $B_a(Q^s, \theta > \hat{\theta})$ denote $\int_{\hat{\theta}} \int_{\epsilon} \Pi(\theta, Q^s)\mu(\theta; Q^s, \theta > \hat{\theta})g(\epsilon)d\theta dq^s$. The first order condition turns into

$$\left(\frac{\partial B_a(Q^s, \theta > \hat{\theta})}{\partial Q^s} + \frac{\partial B_a^s(Q^s, \theta > \hat{\theta})}{\partial \int_{\hat{\theta}}^{\infty} \mu(\theta; Q^s \mid \theta > \hat{\theta})d\theta} \frac{\partial \int_{\hat{\theta}}^{\infty} \mu(\theta; Q^s \mid \theta > \hat{\theta})d\theta}{\partial Q^s} \right) \frac{\partial Q^s}{\partial e} = C'(e) \tag{4.6}$$

For a student at school b, optimal effort is like-wise

$$\left(\frac{\partial B_b^s(Q^s \mid \theta < \hat{\theta})}{\partial Q^s} + \frac{\partial B_b^s(Q^s \mid \theta < \hat{\theta})}{\partial \int_{-\infty}^{\hat{\theta}} \mu(\theta; Q^s \mid \theta < \hat{\theta})d\theta} \frac{\partial \int_{-\infty}^{\hat{\theta}} \mu(\theta; Q^s \mid \theta < \hat{\theta})d\theta}{\partial Q^s} \right) \frac{\partial Q^s}{\partial e} = C'(e) \tag{4.7}$$

Proposition 4.1. *A perfect selection is the dominant policy.*

Proof. The marginal increase in observed qualification is greater positive if students are tracked according to their ability.

$$\left(\int_{\theta} Q_e^s f(\theta) d\theta > \int_{\theta} Q_e^{ns} f(\theta) d\theta \right)$$

with

$$Q_e^s = \frac{\partial Q(q(e(\theta), \theta, \overline{\theta}_k^s))}{\partial e}$$

$$Q_e^{ns} = \frac{\partial Q(q(e(\theta), \theta, \overline{\theta}_k^{ns}))}{\partial e}$$

Students at school a in the sorting regime face a better peer group than they would in a comprehensive schools. Students at school b face a worse peer group. Ability is symmetrically distributed around $\hat{\theta}$, such that average ability in both schools is equidistant from the average ability in the entire population ($\hat{\theta} - \overline{\theta}_b = \overline{\theta}_a - \hat{\theta}$) and students at the better school have a greater marginal productivity ($\frac{\partial^2 q}{\partial \overline{\theta}_k \partial \theta} > 0$, see assumption 4.1). Hence, the increase in qualification for students at school a is greater than the loss for students at school b. Students in comprehensive schools have to provide more effort to get the same qualification. The logic is applicable to a comparison between a perfect selection and any imperfect one. □

Proposition 1 does not rule out that comprehensive education leads to higher qualifications. The impact of student effort on the employers' belief formation is lower in a sorting regime than in a comprehensive one. Both $\frac{\partial \int_{-\infty}^{\hat{\theta}} \mu(\theta; Q^s, \theta < \hat{\theta}) d\theta}{\partial Q^s}$ in equation (4.7) and $\frac{\partial \int_{\hat{\theta}}^{\infty} \mu(\theta; Q^s, \theta > \hat{\theta}) d\theta}{\partial Q^s}$ in equation (4.6) are smaller than $\frac{\partial \int_{\theta} \mu(\theta; Q^{ns}) d\theta}{\partial Q^{ns}}$ in equation (4.4). This inequality increases if student ability is close to $\hat{\theta}$. Students at school b with good observed qualifications will never be as able as students with a poor record at school a. Hence, comprehensive schools may elicit more effort and higher qualifications than schools in a selective system.

Higher qualifications from higher effort levels may increase the wage offers for a student such that $\int_\theta B_k^{ns}(\theta)f(\theta)d\theta > \int_\theta B_k^s(\theta)f(\theta)d\theta$ holds. Yet, the additional costs of effort outweigh this increase in wage offers.

$$\left(\int_\theta C^{ns}(e(\theta),\overline{\theta}_k^{ns}) - C^s(e(\theta,\overline{\theta}_k^{ns})) - B_k^{ns}(\theta) + B_k^s(\theta)f(\theta)d\theta < 0\right)$$

Otherwise, equations (4.6) and (4.7) would be violated. Comprehensive education can be the optimal solution, if qualification or just the effort provision generate positive external effects, e.g. happy parents.

The properties of the peer effect also show which students prefer which educational system.

Corollary 1. *Students with $\theta > \hat{\theta}$ prefer a selective schooling system, students with $\theta < \hat{\theta}$ prefer comprehensive schools. Sorting increases the inequality in educational qualification between these two types of students.*

The marginal productivity is greater for students with $\theta > \hat{\theta}$ in the selective system and greater for students with $\theta < \hat{\theta}$ in the comprehensive system. Therefore, students with a high expected qualification will leave a comprehensive school, if they can also go to a selective one, e.g. a private school or a selective school in a neighboring jurisdiction. Of course, such an outside option undermines potential benefits from comprehensive schooling. The inequality in qualification increases in a selective system as good students gain from improved peer effects and low ability students lose accordingly.

4.4 Intertemporal incentives

Almost all educational systems do not sort in elementary school but at some later point. In this section, the incentives from a selective school system for pre-selection students are analyzed. The results from the analysis have implications for econometric research which will be discussed in a subsection.

Consider a two period process. In period $t = 1$ (or in primary education) all students are educated in identical schools. In period $t = 2$ (or secondary education) the students are either sorted or they are still kept in identical schools. In the selective case, sorting depends on the observed output in the first period (Q_1). This criterion allows for a cheap and simple selection mechanism. I assume all other alternatives as prohibitively expensive. If Q_1 is greater than the $\hat{q}$, the student will be sorted into better peer group. The term $\hat{q}$ denotes the expected output of a student with average ability in equilibrium. Of course, this is an endogenous variable. Given an infinite number of students and effort provision increasing in ability this expected output constitutes a quasi-exogenous threshold value.

Students can choose their respective effort level at the start of each period. The qualification in period 2 additionally depends on the actual qualification after period 1, such that

$$Q_2 = q(e_2, \theta, q_1; \overline{\theta}_k)) + \varepsilon_2 \tag{4.8}$$

The separation into two periods changes the incentive structure of education as students in a tracking system have higher incentives in the first period.

The analysis of intertemporal incentives is backwards. Period 2 is analogous to what has been discussed in the previous section, with two notable exceptions. Firstly, the peer group effects are different because sorting is based on the observable output Q_1. which depends on a random variable. Hence, some less able or less engaged students will slip into the better peer group and vice versa. This exception does not imply changes in the qualitative results of the previous section. Secondly, the impact of the output in period 1 has to be taken into account. Therefore, further qualitative differences between a selective and comprehensive system stem from differences in the pre-selection period.

In this period 1, consider first the non-sorting case. The problem of a student is the following:

$$\max_{e_1} B^{ns}(Q_2(q_1)) - C(e_1) \tag{4.9}$$

which implies

$$\left(\frac{\partial B^{ns}(Q_2)}{\partial Q_2} + \frac{\partial B^{ns}(Q_2)}{\partial \int_\theta \mu(\theta; Q_2)d\theta} \frac{\partial \int_\theta \mu(\theta; Q_2)d\theta}{\partial Q_2}\right) \frac{\partial Q_2}{\partial Q_1} \frac{\partial Q_1}{\partial e_1} = C'(e) \tag{4.10}$$

The school indicator k has been ignored in the notation because all schools are identical in both periods. In the comprehensive school system students in elementary school are only motivated by the effect on their post-educational wages.

For the sorting case, a student has to solve

$$\begin{aligned} &\max_{e_1} \Pr(Q_1 > \hat{q})\, B^s(Q_2(q_1) \mid Q_1 > \hat{q}) \\ &\qquad + (1 - \Pr(Q_1 > \hat{q}))\, B^s(Q_2(q_1) \mid Q_1 < \hat{q}) - C(e_{i,1}) \end{aligned} \tag{4.11}$$

From the relevant first order condition follows

$$g\left(Q_1-\hat{q}\right)\Delta B+\Pr\left(Q_1>\hat{q}\right)\frac{\partial B^s(Q_2(q_1)\mid Q_1>\hat{q})}{\partial q_1}\frac{\partial q_1}{\partial e_1}$$
$$+\left(1-\Pr\left(Q_1>\hat{q}\right)\right)\frac{\partial B^s(Q_2(q_1)\mid, Q_1<\hat{q})}{\partial q_1}\frac{\partial q_1}{\partial e_1}=C'(e_{i,1}) \quad (4.12)$$

Notation in the equation is simplified, with

$$\Delta B=B^s(Q_2(q_1)\mid Q_1>\hat{q})-B^s(Q_2(q_1)\mid Q_1<\hat{q})$$

$$\frac{\partial B^s(Q_2(q_1)\mid Q_1>\hat{q})}{\partial q_1}=$$
$$\left(\frac{\partial B^s(Q_2(q_1)\mid Q_1>\hat{Q})}{\partial Q_2}+\frac{\partial B^s(Q_2(q_1)\mid Q_1>\hat{Q})}{\partial\int_\theta\mu(\theta;Q_2,Q_1>\hat{Q})d\theta}\frac{\partial\int_\theta\mu(\theta;Q_2,Q_1>\hat{Q})d\theta}{\partial Q_2}\right)\frac{\partial Q_2}{\partial q_1}$$

and

$$\frac{\partial B^s(Q_2(q_1)\mid Q_1<\hat{q})}{\partial q_1}=$$
$$\left(\frac{\partial B^s(Q_2(q_1)\mid Q_1<\hat{Q})}{\partial Q_2}+\frac{\partial B^s(Q_2(q_1)\mid Q_1<\hat{Q})}{\partial\int_\theta\mu(\theta;Q_2,Q_1<\hat{Q})d\theta}\frac{\partial\int_\theta\mu(\theta;Q_2,Q_1<\hat{Q})d\theta}{\partial Q_2}\right)\frac{\partial Q_2}{\partial q_1}$$

Recall that $g(.)$ is the density function of the random variable. The selective system motivates a student in period 1 by increasing its chance to be promoted to the good school $(g\left(Q_{i1}-\hat{q}\right)\Delta w_{i,2})$ as well as via the wage increases from overall increased educational output.

Proposition 4.2. *Assume that - for a given qualification q_1 from period 1 - the expected additional qualification in secondary education $\Delta(s,ns)$ is identical in the sorting and in* *the non-sorting regime.*

$$\Delta(s, ns) = \int\limits_{\theta} Q_2^s(\theta) - Q_1^s(\theta) d\theta - \int\limits_{\theta} Q_2^{ns}(\theta) - Q_1^{ns}(\theta) d\theta = 0 \tag{4.13}$$

Then, the following results hold:

1. *First period qualification in the sorting regime is lower for the students with sufficiently low ability*

   $$\lim_{\theta \to -\infty} (q(\theta)_1^s - q(\theta)_1^{ns}) < 0$$

   and higher for the average student and for students with very high expected ability.

   $$q(\theta \geq \hat{\theta})_1^s > q(\theta \geq \hat{\theta})_1^{ns}$$

2. *Increasing uncertainty (increasing σ^2) reduces the incentives for the students with average ability and increases the incentives for the students at the margins of the ability distribution.*

Proof. The term $g\left(Q_1 - \hat{q}\right) \Delta B$ in equation (4.12) denotes the increase in probability to get the higher returns from better schooling. Due to the properties of the error term, the following conditions hold for students with very high expected output and for students with very low expected output:

$$\lim_{q \to -\infty} g\left(Q_1 - \hat{q}\right) = \lim_{q \to \infty} g\left(Q_1 - \hat{q}\right) = 0 \tag{4.14}$$

$$\lim_{q \to -\infty} \Pr\left(Q_1 > \hat{q}\right) = 1 - \lim_{q \to \infty} \Pr\left(Q_1 > \hat{q}\right) = 0 \tag{4.15}$$

Very able students are almost certain to go to school a. Students with very low ability are bound for school b. Additional effort basically does not affect the selection outcome.

The peer group effects imply that very good students benefit from increased marginal productivity in a selective system, while low ability students face a loss in productivity (see proof of proposition 1):

$$\frac{B^s(Q_2(q_1) \mid Q_1 > \hat{q})}{\partial q_1}\frac{\partial q_1}{\partial e_1} > \frac{B^{ns}(Q_2, q_1)}{\partial q_1}\frac{\partial q_1}{\partial e_1}$$

and

$$\frac{B^s(Q_2(q_1) \mid Q_1 < \hat{q})}{\partial q_1}\frac{\partial q_1}{\partial e_1} < \frac{B^{ns}(Q_2(q_1))}{\partial q_1}\frac{\partial q_1}{\partial e_1}$$

These implications establish the results for the high ability students and the low ability students.

For the student with $\theta = \hat{\theta}$,

$$g\left(Q_1 - \hat{q}\right)\Delta B + \Pr\left(Q_1 > \hat{q}\right)\frac{\partial B^s(Q_2(q_1) \mid Q_1 > \hat{q})}{\partial e_1} + \left(1 - \Pr\left(Q_1 > \hat{q}\right)\right)\frac{\partial B^s(Q_2(q_1) \mid Q_1 < \hat{q})}{\partial e_1} > \frac{\partial B^{ns}(Q_2(q_1))}{\partial e_1} \quad (4.16)$$

has to hold in order to satisfy the proposition. Since both schools have the same amount of students, the expected qualification in period 1 for a student is defined by the threshold value $\hat{q}$. such that $g\left(E(Q_1) - \hat{q}\right) = g(0) > 0$. The peer effects imply

$$\Pr\left(Q_1 > \hat{q}\right)\frac{\partial B^s(Q_2(q_1) \mid Q_1 > \hat{q})}{\partial e_1} + \left(1 - \Pr\left(Q_1 > \hat{q}\right)\right)\frac{\partial B^s(Q_2(q_1) \mid Q_1 < \hat{q})}{\partial e_1} \not< \frac{\partial B^{ns}(Q_2)}{\partial e_1}$$

which establishes equation (4.16).

For result 2 notice that $g\left(Q_1 - \hat{q}\right)$ increases with increasing uncertainty at the margins of the distributions of possible output and decreases at $E(Q_1) = \hat{q}$. □

Result 1 is similar to results from bonus or tournament models. Increased uncertainty increases the likelihood of passing or failing at the margins. Hence, these students work harder to achieve the desired result. For average students, the impact of effort on the outcome decreases, hence the incentives are lower.

4.4.1 Econometric Implications

The implication of result 1 is crucial for the empirical literature. The evaluation of tracking policies (e.g. Argys et al. 1996; Betts and Skolnick 2000; Figlio and Page 2002) often suffers from a lack of independent observations. It is also difficult because unmeasured factors can bias the estimation results, e.g. the impact of other policies or differences in teacher quality. The differences-in-differences approach by Hanushek and Wößmann (2006) controls for these institutional differences across countries, as long as tracking and non-tracking countries do not differ systematically in other relevant policies. The approach and the available data set of Hanushek and Wößmann do provide a valuable insight into the impact of tracking on the distribution of educational outcomes. Yet, their methodological improvement in one area raises another objection with respect to overall outcome. Hanushek and Wößmann (2006, p. 66) claim that

> ”[t]he impact of tracking can (...) be estimated by comparing the average achievement gain in tracked countries to that in untracked countries.”

Such a differences-in-differences methodology tests if equation (4.13) in proposition 2 is larger, smaller or equal to zero. A similar approach is provided by Pischke and Manning (2006). However, due to result 1 in the previous proposition, $\Delta(s, ns)$ does not capture that tracking ceteris paribus leads to higher performance in primary education ($\int_\theta Q_1^s(\theta)d\theta > \int_\theta Q_1^{ns}(\theta)d\theta$). Equation (4.13) just captures the impact of tracking on the obtained qualifications in *secondary* education. The full impact of tracking is underestimated.

4.5 Conclusion

The paper has discussed a reason why positive effects from the selection of students are not observable, even if standard assumptions about peer group effects are not rejected. Sorting already provides a signal about unobservable ability. Hence, student performance in selective secondary education has less impact on the beliefs of an employer than in a comprehensive system. Introducing ability grouping transfers incentives from secondary into primary education. If the assumptions about peer effects are correct, then sorting is the dominant policy. Equality in test scores with identical financial inputs does not reflect unobservable effort provision.

The results allow to derive empirically testable hypotheses. Firstly, performance in primary education is higher in systems where selection takes place in secondary education. Primary education has a higher impact on labor market success in selective systems. Hence, differences-in-differences estimations do not capture the full impact of tracking.

Secondly, proposition 1 implies that comprehensive education requires higher effort contributions from students to acquire the same qualification. If this hypotheses can be confirmed it questions a popular assumption that improving qualification with given financial resources also improves welfare. This assumption only holds if the social benefits from education are sufficiently large.

The argument in the paper explains why average ability in a peer group does not need to be rejected as a measure for peer effects. However, it also does not eliminate the case for further research on the properties of peer effects.

References

Ammermüller, Andreas and Jörn-Steffen Pischke. 2006. "Peer Effects in European Primary Schools: Evidence from PIRLS." ZEW Discussion Paper: Mannheim.

Angrist, J. and Victor Lavy. 2003. "The Effect of High School Matriculation Awards: Evidence from Randomized Trials." NBER.

Argys, Laura M., Daniel I. Rees, and Dominic J. Brewer. 1996. "Detracking America's schools: Equity at zero cost?" Journal of Policy Analysis and Management, 15:4, pp. 623-45.

Betts, Julian R. and Jamie L. Shkolnik. 2000. "The effects of ability grouping on student achievement and resource allocation in secondary schools." Economics of Education Review, 19, pp. 1-15.

Brunello, Giorgio and Massimo Giannini. 2005. "Stratified or comprehensive? The economic efficiency of school design." Scottish Journal of Political Economy, 51:2, pp. 173-93.

Brunello, Giorgio, Massimo Giannini, and Kenn Ariga. 2004. "The Optimal Timing of School Tracking." IZA.

Cullen, Julie, Brian Jacob, and Steven Levitt. 2003. "The effect of school choice on student outcomes: Evidence from randomized lotteries." NBER Working Paper.

De Fraja, Gianni and Pedro Landeras. 2006. ”Could do better: the effectiveness of incentives and competition in schools.” Journal of Public Economics, 90, pp. 189–213.

De Fraja, Gianni, Tania Oliveira, and Luisa Zanchi. 2005. ”Must try harder. Evaluating the role of effort on examination results.” CEPR.

Dobbelsteen, Simone, Jesse Levin, and Hessel Oosterbeek. 2002. ”The Causal Effect of Class Size on Scholastic Achievement: Distinguishing the Pure Class Size Effect from the Effect of Changes in Class Composition.” Oxford Bulletin of Economics and Statistics, 64:1, pp. 17-38.

Eisenkopf, Gerald. 2007. ”Learning and Peer Effects - Experimental Evidence.” University of Konstanz.

Epple, Dennis, Elizabeth Newlon, and Richard E. Romano. 2002. ”Ability tracking, school competition, and the distribution of educational benefits.” Journal of Public Economics, 83, pp. 1-48.

Epple, Dennis, Richard E. Romano, and Holger Sieg. 2003. ”Peer effects, financial aid, and selection of students into colleges and universities: An empirical analysis.” Journal of Applied Econometrics, 18:5, pp. 501-25.

Figlio, David N. and Marianne E. Page. 2002. ”School choice and the distributional effects of ability tracking: does separation increase inequality?” Journal of Urban Economics, 51:3, pp. 497–514.

Hanushek, Eric A., John F. Kain, Jacob M. Markman, and Steven Rivkin. 2003. ”Does Peer Ability Affect Student Achievement?” Journal of Applied Econometrics, 18:5, pp. 18 (5), 527-44.

Hanushek, Eric A. and Ludger Wößmann. 2006. ”Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence across Countries.” Economic Journal, 116, pp. C63–C76.

Hoxby, Caroline. 2003. ”School choice and school competition: Evidence from the United States.” Swedish Economic Policy Review, 10, pp. 11-67.

Hoxby, Caroline M. 2000. ”Does competition among public schools benefit students and taxpayers?” American Economic Review, 90:5, pp. 1209-38.

Hoxby, Caroline M. 2001. ”Peer effects in the classroom: Learning from gender and race variation.” NBER.

Kremer, Michael, Edward Miguel, and Rebecca Thornton. 2004. ”Incentives to Learn.” Harvard University: Cambridge, MA.

Lazear, Edward P. 2001. ”Educational Production.” The Quarterly Journal of Economics, 116:3, pp. 777-803.

Manski, Charles. 1993. ”Identification of endogenous social effects: The reflection problem.” Review of Economic Studies, 60:3, pp. 531-42.

McEwan, Patrick. 2003. ”Peer effects on student achievement: Evidence from Chile.” Economics of Education Review, 22:2, pp. 131-41.

Meier, Volker. 2004. ”Choosing between School Systems: The Risk of Failure.” Finanzarchiv, 60:1, pp. 83-93.

Pischke, Jörn-Steffen and Alan Manning. 2006. ”Comprehensive versus Selective Schooling in England in Wales: What Do We Know?” NBER.

Sacerdote, Bruce. 2001. ”Peer effects with random assignment: Results for Dartmouth roommates.” Quarterly Journal of Economics, 116:2, pp. 681–704.

Spence, Michael. 1973. ”Job Market Signaling.” The Quarterly Journal of Economics, 87:3, pp. 355-74.

West, Martin R. and Ludger Wößmann. 2005. ”Which school systems sort weaker students into smaller classes? International evidence.” European Journal of Political Economy.

Wößmann, Ludger. 2004. ”How equal are educational opportunities? Family background and student achievement in Europe and the US.” CESifo.

Complete References

Ammermüller, Andreas and Jörn-Steffen Pischke. 2006. ”Peer Effects in European Primary Schools: Evidence from PIRLS.” ZEW Discussion Paper: Mannheim.

Angrist, J. and Victor Lavy. 2003. ”The Effect of High School Matriculation Awards: Evidence from Randomized Trials.” NBER.

Argys, Laura M., Daniel I. Rees, and Dominic J. Brewer. 1996. ”Detracking America’s schools: Equity at zero cost?” Journal of Policy Analysis and Management, 15:4, pp. 623-45.

Baker, George. 2002. ”Distortion and Risk in Optimal Incentive Contracts.” Journal of Human Resources, 37:4, pp. 728-51.

Baker, G., R. Gibbons, and K. J. (1994) Murphy. 1994. ”Subjective Performance Measures in Optimal Incentive Contracts.” The Quarterly Journal of Economics, 109:4, pp. 1125-56.

Banker, R. D. and Srikant Datar. 1989. ”Sensitivity, precision, and linear aggregation of signals for performance evaluation.” Journal of Accounting Research, 27:1, pp. 21-39.

Bergstresser, Daniel and Thomas Philippon. 2006. ”CEO Incentives and Earnings Management.” Journal of Financial Economics, 80:3, pp. 511-29.

Betts, Julian R. and Jamie L. Shkolnik. 2000. ”The effects of ability grouping on student achievement and resource allocation in secondary schools.” Economics of Education Review, 19, pp. 1-15.

Brunello, Giorgio and Massimo Giannini. 2005. ”Stratified or comprehensive? The economic efficiency of school design.” Scottish Journal of Political Economy, 51:2, pp. 173-93.

Brunello, Giorgio, Massimo Giannini, and Kenn Ariga. 2004. ”The Optimal Timing of School Tracking.” IZA.

Brunello, G. and L. Rocco. 2005. ”Educational standards in private and public schools.” CESifo.

Clotfelter, Charles T. 1996. Buying the best. Cost escalation in elite higher education. Princeton: Princeton University Press.

Courty, Pascal and Gerald Marschke. 2003. "Dynamics of Performance Measurement Systems." Oxford Review of Economic Policy, 19:2, pp. 268-84.

Cullen, Julie, Brian Jacob, and Steven Levitt. 2003. "The effect of school choice on student outcomes: Evidence from randomized lotteries." NBER Working Paper.

Currie, Janet and Enrico Moretti. 2003. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings." Quarterly Journal of Economics, 118:4, pp. 1495-532.

Datar, Srikant, Susan Cohen Kulp, and Richard A. Lambert. 2001. "Balancing Performance Measures." Journal of Accounting Research, 39:1, pp. 75-92.

De Fraja, Gianni and Elisabetta Iossa. 2002. "Competition among universities and the emergence of the élite institution." Bulletin of Economic Research, 54:3, pp. 275-93.

De Fraja, Gianni and Pedro Landeras. 2006. "Could do better: the effectiveness of incentives and competition in schools." Journal of Public Economics, 90, pp. 189–213.

De Fraja, Gianni, Tania Oliveira, and Luisa Zanchi. 2005. "Must try harder. Evaluating the role of effort on examination results." CEPR.

Dee, Thomas S. 2004. "Are there civic returns to education?"" Journal of Public Economics, 88:9-10, pp. 1697-720.

Dikolli, S. 2001. "Agent employment horizons and contracting demand for forward-looking performance measures." Journal of Accounting Research, 39:3, pp. 481-94.

Dittrich, Dennis and Martin Kocher. 2006. "Monitoring and Pay: An Experiment on Employee Performance under Endogenous Supervision." Max Planck Institute of Economics.

Dobbelsteen, Simone, Jesse Levin, and Hessel Oosterbeek. 2002. "The Causal Effect of Class Size on Scholastic Achievement: Distinguishing the Pure Class Size Effect from the Effect of Changes in Class Composition." Oxford Bulletin of Economics and Statistics, 64:1, pp. 17-38.

Dutta, S. and S. Reichelstein. 2003. "Leading Indicators, performance measurement and long-term versus short-term contracts." Journal of Accounting Research, 41:5, pp. 837-66.

Dutta, Sunil and Stefan Reichenstein. 2005. "Stock Price, Earnings, and Book Value in Managerial Performance Measures." Accounting Review, 80:4, pp. 1069-100.

Ehrenberg, Ronald G. 2000. Tuition rising. Why college costs so much. Cambridge, MA: Harvard University Press.

Eisenkopf, Gerald. 2007. "Learning and Peer Effects - Experimental Evidence." University of Konstanz.

Epple, Dennis, Elizabeth Newlon, and Richard E. Romano. 2002. "Ability tracking, school competition, and the distribution of educational benefits." Journal of Public Economics, 83, pp. 1-48.

Epple, Dennis, Richard E. Romano, and Holger Sieg. 2003. "Peer effects, financial aid, and selection of students into colleges and universities: An empirical analysis." Journal of Applied Econometrics, 18:5, pp. 501-25.

Ewert, Ralf and Alfred Wagenhofer. 2005. "Economic Effects of Tightening Accounting Standards to Restrict Earnings Management." Accounting Review, 80:4, pp. 1101-24.

Feltham, G. and J. Xie. 1994. "Performance measure congruity and diversity in multi-task principal-agent relations." Accounting Review, 69:3, pp. 429-53.

Figlio, David N. and Marianne E. Page. 2002. "School choice and the distributional effects of ability tracking: does separation increase inequality?" Journal of Urban Economics, 51:3, pp. 497–514.

Gary-Bobo, Robert J. and Alain Trannoy. 2005. "Efficient tuition fees, examinations and subsidies." CEPR: 31.

Gibbs, Michael, Kenneth A. Merchant, Wim A. Van der Stede, and Mark E. Vargus. 2004. "Determinants and Effects of Subjectivity in Incentives " Accounting Review, 79:2, pp. 409-36.

Groot, W. and H. M. van den Brink. 2000. "Overeducation in the labor market: a meta-analysis." Economics of Education Review, 19:2, pp. 149-58.

Hansmann, Henry B. 1980. "The Role of Nonprofit Enterprise." The Yale Law Journal, 89:5, pp. 835-901.

Hanushek, Eric A., John F. Kain, Jacob M. Markman, and Steven Rivkin. 2003. "Does Peer Ability Affect Student Achievement?" Journal of Applied Econometrics, 18:5, pp. 18 (5), 527-44.

Hanushek, Eric A. and Ludger Wößmann. 2007. "The role of school improvement in economic development." NBER.

Hanushek, Eric A. and Ludger Wößmann. 2006. "Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence across Countries." Economic Journal, 116, pp. C63–C76.

Holmstrom, B. and Milgrom P. 1991. "Multi-task Principal-Agent Analysis: Incentive Contracts, Asset Ownership, and Job-Design." Journal of Law, Economics and Organization:7 (special issue), pp. 24-52.

Hoxby, Caroline. 2003. "School choice and school competition: Evidence from the United States." Swedish Economic Policy Review, 10, pp. 11-67.

Hoxby, Caroline M. 1997. "How the changing market structure of U.S. higher education explains college tuition." NBER: Cambridge, MA.

Hoxby, Caroline M. 2000. "Does competition among public schools benefit students and taxpayers?" American Economic Review, 90:5, pp. 1209-38.

Hoxby, Caroline M. 2001. "Peer effects in the classroom: Learning from gender and race variation." NBER.

Ittner, Christopher D., David F. Larcker, and Marshall W. Meyer. 2003. "Subjectivity and the Weighting of Performance Measures: Evidence from a Balanced Scorecard." Accounting Review, 78:3, pp. 725-58.

Jacob, Brian A. 2005. "Accountability, Incentives and Behavior: The Impact of High-Stakes Testing in the Chicago Public Schools." Journal of Public Economics 89:5-6, pp. 761-96.

James, Estelle. 1990. "Decision processes and priorities in higher education," in The economics of American universities. Stephen A. Hoenack and Eileen L. Collins eds. Albany: State University of New York Press, pp. 77-106.

Kaplan, R. S and D. P. Norton. 1992. "The Balanced Scorecard - measures that drive performance." Harvard Business Review.

Kaplan, R. S and D. P. Norton. 1993. "Putting the Balanced Scorecard to work. ." Harvard Business Review, pp. 134-47.

Kemnitz, Alexander. 2005. "University finance reform, competition and teaching quality." University of Mannheim: Mannheim.

Kerr, Steven. 1975. "On the Folly of Rewarding for a While Hoping for B." Academy of Management Journal, 18:4, pp. 769-83.

Kremer, Michael, Edward Miguel, and Rebecca Thornton. 2004. "Incentives to Learn." Harvard University: Cambridge, MA.

Krishnan, Ranjani, Joan L. Luft, and Michael D. Shields. 2005. "Effects of Accounting-Method Choices on Subjective Performance-Measure Weighting

Decisions: Experimental Evidence on Precision and Error Covariance." Accounting Review, 80:4, pp. 1163-92.

Lazear, Edward P. 2001. "Educational Production." The Quarterly Journal of Economics, 116:3, pp. 777-803.

Lazear, Edward P. 2006. "Speeding, Terrorism, and Teaching to the Test." Quarterly Journal of Economics, 121:3, pp. 1029-61.

Liang, P. J. 2004. "Equilibrium Earnings Management, incentive contracts, and accounting standards." Contemporary Accounting Research, 21, pp. 685-717.

Lochner, Lance and Enrico Moretti. 2004. "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports." American Economic Review, 94:1, pp. 155-89.

Manski, Charles. 1993. "Identification of endogenous social effects: The reflection problem." Review of Economic Studies, 60:3, pp. 531-42.

Martínez-Mora, Francisco. 2006. "The existence of non-elite private schools." Journal of Public Economics, 90, pp. 1505-18.

McEwan, Patrick. 2003. "Peer effects on student achievement: Evidence from Chile." Economics of Education Review, 22:2, pp. 131-41.

Meier, Volker. 2004. "Choosing between School Systems: The Risk of Failure." Finanzarchiv, 60:1, pp. 83-93.

Moers, Frank. 2006. "Performance Measure Properties and Delegation." Accounting Review, 81:4, pp. 897-924.

Milligan, Kevin, Enrico Moretti, and Philip Oreopoulos. 2004. "Does education improve citizenship? Evidence from the United States and the United Kingdom." Journal of Public Economics, 88:9-10, pp. 1667-95.

Murphy, Kevin J. and Paul Oyer. 2003. "Discretion in Executive Incentive Contracts." University of Southern California.

Nagin, D. S., J. B. Rebitzer, S. Sanders, and L. J. Taylor. 2002. ”Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment.” American Economic Review, 92:4, pp. 850-73.

Ortmann, Andreas and Richard Squire. 2000. ”A game-theoretic explanation of the administrative lattice in institutions of higher learning.” Journal of Economic Behavior and Organization, 43, pp. 377-91.

Pischke, Jörn-Steffen and Alan Manning. 2006. ”Comprehensive versus Selective Schooling in England in Wales: What Do We Know?” NBER.

Plug, Erik and Wim Vijverberg. 2003. ”Schooling, Family Background, and Adoption: Is It Nature or Is It Nurture?” Journal of Political Economy, 111:3, pp. 611-41.

Prendergast, Canice. 1999. ”The provision of incentives in firms.” Journal of Economic Literature, 37, pp. 7-63.

Rajan, Madhav V. and Stefan Reichelstein. 2006. ”Subjective Performance Indicators and Discretionary Bonus Pools ” Journal of Accounting Research, 44:3, pp. 585-618.

Rothschild, Michael and Lawrence J. White. 1995. ”The analytics of the pricing of higher education and other services in which the customers are inputs.” Journal of Political Economy, 103:3, pp. 573-86.

Sacerdote, Bruce. 2001. ”Peer effects with random assignment: Results for Dartmouth roommates.” Quarterly Journal of Economics, 116:2, pp. 681–704.

Schipper, K. 2003. ”Principle-based accounting standards.” Accounting Horizons, pp. 61-72.

Sliwka, Dirk. 2002. ”On the use of nonfinancial performance measures in management compensation.” Journal of Economics and Management Science, 11:3, pp. 487-511.

Smith, Michael J. 2002. "Gaming Nonfinancial Performance Measures." Journal of Management Accounting Research, 14, pp. 119-33.

Spence, Michael. 1973. "Job Market Signaling." The Quarterly Journal of Economics, 87:3, pp. 355-74.

Ursprung, H. W. 2003. "Schneewittchen im Land der Klapperschlangen: Evaluation eines Evaluators"." Perspektiven der Wirtschaftspolitik:Band 2, pp. 177-89.

Warning, Susanne. 2004. "Performance Differences in German Higher Education: Empirical Analysis of Strategic Groups." Review of Industrial Organization, 24:4, pp. 393-408.

West, Martin R. and Ludger Wößmann. 2005. "Which school systems sort weaker students into smaller classes? International evidence." European Journal of Political Economy.

Winston, Gordon C. 1999. "Subsidies, hierarchy and peers: the awkward economics of higher education." Journal of Economic Perspectives, 13:1, pp. 13-36.

Winston, Gordon C. 2003. "Towards a theory of tuition: Prices, peer wages, and competition in higher education." Williams College - Wiliams Project on the Economics of Higher Education: Williamstown.

Winston, Gordon C. 2004. "Differentiation among US Colleges and Universities." Review of Industrial Organization, 24, pp. 331-54.

Wößmann, Ludger. 2004. "How equal are educational opportunities? Family background and student achievement in Europe and the US." CESifo.

Erklärung

Ich versichere hiermit, dass ich die vorliegende Arbeit mit dem Thema

The Organization of Education

ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Weitere Personen, insbesondere Promotionsberater, waren an der inhaltlich materiellen Erstellung dieser Arbeit nicht beteiligt. Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Konstanz, 6. Februar 2007

(Gerald Eisenkopf)