Discussion Paper No. 07-016

Double Standards in Educational Standards – Are Disadvantaged Students Being Graded More Leniently?

Oliver Himmler and Robert Schwager

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 07-016

Double Standards in Educational Standards – Are Disadvantaged Students Being Graded More Leniently?

Oliver Himmler and Robert Schwager

Non-Technical Summary

An effective and equitable schooling system is widely seen as an important instrument to promote social policy goals. Since a good education is hardly valued on the labour market without a formal degree, the grading and examination system is crucial for the success of such policies. The purpuse of the present paper is therefore to examine, both theoretically and empirically, the interaction of the social status of a school's students and the standards applied at examination.

We present a theoretical model where schools set graduation standards and students decide on how much learning effort to provide. Graduates from disadvantaged backgrounds are assumed to obtain a lower wage than students from other social classes, because they are discriminated against, or because they lack, for example, rhetorical abilities, connections to personal networks, or other "soft skills". We show that in this setup, schools with a student body from lower social classes set less demanding standards than other schools, even if the abilities of the disfavoured students are identical to those of others. This result obtains since graduates from unfavourable backgrounds have lower returns to education, and therefore are less willing to satisfy any given standard than students from higher social origins. To make up for the resulting loss in the numbers of graduates schools with disadvantaged students award degrees more easily.

The predictions from the theoretical model are then tested using school level data from the Netherlands. In that country, students in most subjects must pass central exams as well as school specific exams in order to receive a diploma. Thus, the difference between the average grades obtained by the final class of a school in the school specific and in the central examination provides a measure for the grading standard applied by this school. Using this difference as the dependent variable, we find that in most specifications, the percentage of cultural minority students, and the percentage of students eligible for financial aid, have a significant positive impact on the grade difference. Thus, the empirical analysis backs the main prediction of the theoretical model: Schools with a higher percentage of disadvantaged students use a more lenient grading scheme than other schools.

The flavour of this result differs strikingly from the findings of the latest PISA report on student achievement in Germany (see Prenzel et al., 2005). According to this study, children from lower social classes are much less likely to attend the branch of secondary schooling preparing for university than those from higher classes, even after controlling for ability. Whereas this seems to suggests that grading and examination in Germany are skewed against disadvantaged students, we show that in the Netherlands such students are held to less demanding standards than their counterparts from average backgrounds. Given the different sort of data, a different country, and a different approach, one can obviously not draw any more general conclusions from the divergence of these results. Nevertheless, our results point out that, in order to explain the treatment of lower class students by the schooling system, more than a simple appeal to discrimination is needed.

Double Standards in Educational Standards – Are Disadvantaged Students Being Graded More Leniently?[a]

Oliver Himmler

Georg-August-Universität Göttingen[b]

and

Robert Schwager

Georg-August-Universität Göttingen[c]
and Zentrum für Europäische Wirtschaftsforschung, Mannheim

First version: February 2006
This version: March 2007

Abstract

A simple model of decentralised graduation standards is presented. It is shown that a school whose students are disadvantaged on the labour market applies less demanding standards because such students have less incentives to graduate. The model's predictions are tested using Dutch school-level data. Since students in the Netherlands have to participate both in a central and in a school specific examination, we can identify the grading policy of individual schools. We find that schools which harbour greater shares of disadvantaged students tend to set lower standards. This effect is largest in the branch of secondary schooling preparing for university.

Keywords: education, grading, social status, schools, Netherlands
JEL: I21, J15

---

[a]We are grateful to the *Deutsche Forschungsgemeinschaft* for funding part of this research in the priority programme 1142 "Institutional Design of Federal Systems: Theory and Empirical Evidence". We thank Fabienne Plieschke for her assistance in assembling the data. We are indebted to Thiess Büttner, Dennis Epple, and Volker Meier as well as seminar participants at the *Verein für Socialpolitik* Education Economics Group meeting in Kreuzlingen, Athens University of Economics and Business, the priority programme workshop in Bonn, the IIPF Congress in Paphos, the CESifo Group Seminar in Munich and the Århus School of Business for helpful comments. All remaining error are our own.

[b]Volkswirtschaftliches Seminar, Platz der Göttinger Sieben 3, D-37073 Göttingen, phone: +49/551/397290, email: ohimmle@uni-goettingen.de

[c]phone: +49/551/397293, email: rschwag@uni-goettingen.de

1 Introduction

Education policy is widely seen as a means to promote social policy goals. Good schooling is supposed to help the children of disfavoured members of society to earn higher incomes than their parents and to gain social status. In most countries, however, students do not achieve the benefits of formal education just by attending. Instead, they need to obtain the appropriate degree, preferably with good grades. Therefore, any social impact of education policy is filtered through the grading and examination system. Whether good schools will contribute much to social mobility depends on the way standards are chosen, and whether this choice depends on the social origin of students. In this paper, we therefore examine, both theoretically and empirically, the interaction of the social status of a school's students and the standards applied at examination.

In the first part of the paper we introduce status into a model of the choice of examination standard provided by Costrell (1994, 1997). In this model, each school sets its graduation standard so as to maximise the sum of the wages earned by its students. This decision is governed by the trade-off between the number of graduates, which decreases if the standard is more demanding, and the wage earned by each graduate, which increases in the standard. We extend Costrell's formulation by assuming that, in addition to the standard, also the social origin affects the wage earned by graduates. For a given standard, students from disadvantaged backgrounds obtain a lower wage than students from other social classes. We show that in this setup, schools with a disadvantaged student body set lower standards than other schools, even if the abilities of the disfavoured students are identical to those of others. Standards are inflated in this way because the wage discount experienced by graduates from unfavourable backgrounds depresses the return to learning effort for these students. They are thus less willing to satisfy any given standard than students from average social origins. To make up for the resulting loss in the numbers of graduates schools with disadvantaged students choose less demanding standards.

If the standard applied by a single school is not observable by employers, the graduates from several schools are pooled together in a common labour market, earning the same wage. We show that in such a scenario the equilibrium standard is decreasing in the size of the relevant labour market, that is, the number of schools whose graduates are pooled together. Our model thus confirms the well-known grading externality induced by locally determined but unobservable standards: Schools have an incentive to free-ride on high wages brought about by the other schools' tough grading. This mechanism has an implication for social policy, which is our focus here. It is plausible that the students from different social backgrounds are not equally mobile when applying for jobs. Specifically, it may be that disadvantaged students on average stay closer to their original residence. We show that, if this is true, the externality will be smaller in the case of disadvantaged schools, counteracting the tendency to set lower standards induced by unequal job prospects.

In the second part of the paper, we test the theoretical model using data from the Netherlands. This choice of subject is motivated by several features of the Dutch education system. Most importantly, students must pass central exams as well as school specific exams in order to receive a diploma. The average grades obtained by the final class in both examinations in each subject are published annually for all schools in the Netherlands in the so-called *Kwaliteitskaarten* (quality cards). Thus, we are able to use the grades earned in the central examination as a benchmark against which to measure standards employed by individual schools in the school specific examination. In addition, secondary education in the Netherlands is organised in several branches directed towards different further careers, from pre-university education to practical vocational training. This allows us to differentiate our analysis of grading standards according to different labour markets targeted by the different branches.

The empirical analysis aims at explaining differences in standards chosen by different schools. To do this, we use the difference between the average grade of the school specific and the central examination as the dependent variable. The key explanatory variables are two proxies for the social status of a school's students, the percentage of cultural minority students and the percentage of students eligible for financial aid. It turns out that these variables in most specifications indeed have a significant positive impact on the grade difference. Thus, the empirical analysis generally backs the main prediction of the theoretical model: Schools with a higher percentage of disadvantaged students use a more lenient grading scheme than other schools. The effect is largest for the school branch which leads to university, smaller but still significant for the branches of intermediate academic level, and insignificant in the case of the most practically oriented branch. Since it is plausible that graduates are the more mobile the higher the academic level achieved, we take these results as an, albeit weak, evidence for the importance of a school's market size on the grading standard applied.

The present paper contributes to the broad literature in education economics which analyses the effects of the social composition of schools (see, for example, Epple and Romano, 1998, Nechyba, 1999, Epple, Newlon, and Romano, 2002, and Hanushek, 2002: 2078-2081). More specifically, our work is related to several studies analysing examination standards. Most of this research is concerned with the impact of different institutional arrangements for testing and examining students on students' achievement. It has been well established by this strand of research that central standard setting in education paired with centrally devised and graded examinations leads to higher achievements in standardised tests. Theoretical foundations for this claim can be found in Costrell (1997) and Jürges, Richter, and Schneider (2005). Empirically, the performance enhancing effect of central standards has been established, among others, by Bishop et al. (2001), Betts (1998b), and Jacob (2005) for the United States,[1] and by Bishop (1997, 1999), Wößmann (2003), and Jürges, Schneider, and Büchel (2005) for other countries.

Much rarer are studies which aim at explaining how standards are set, and why. The basic

[1]See also the survey by Betts and Costrell (2001).

theoretical approach followed in the present paper has been advanced by Betts (1998a) and Costrell (1994, 1997). More recently, Chan, Li, and Suen (2005) endogenise pooling across several types of schools in a signalling model of grading standards. None of these papers, however, addresses the issue of social class.

This is a major issue in the latest PISA report on student achievement in the OECD. According to Prenzel at al. (2005: 24), children from lower social classes in Germany are much less likely to attend the branch of secondary schooling preparing for university than those from higher classes. Since this allegedly holds even if one controls for the individual student's intelligence, the result suggests, in opposite to our findings, that the grading and examination system discriminates against disadvantaged students. Given the different sort of data and a different, rather descriptive approach, it is, however, difficult to discern the origin of these diverging conclusions.

Finally, Wikström and Wikström (2005) analyse the determinants of grading standards in Sweden. Their approach is similar to ours since it also uses a central test as a benchmark against which local grading is measured. The Dutch central examination differs from this test, and is possibly more attractive as a benchmark, since it is compulsory, avoiding self-selection issues, and since its grades are measured on the same scale as the score of the local examination. Moreover, Wikström and Wikström focus on variables which differ across municipalities such as the intensity of competition among schools, whereas we directly address differences in the characteristics of the schools' student populations.

The remainder of the paper is organised as follows. The following two sections contain the theoretical analysis. The model is presented in section 2 and the optimal standards are derived in section 3. In section 4 we give a brief overview of the institutional setup of the Dutch education system, describe the data, and present the estimation approach. Section 5 then contains the results of the empirical analysis. The concluding section 6 discusses policy implications and possible future lines of research.

2 The Model

In our model, schools set graduation standards which determine wages, and students choose how much learning effort to expand. Students have identical preferences over the wage $w \geq 0$ they will receive after leaving school and the learning effort $e \geq 0$ they expand at school. The learning effort is meant to reflect not only time spent in school or doing homework but also, and possibly more importantly, the intensity of unpleasant school-related activities such as paying attention to the teacher, behaving well in class, thinking hard, etc. The utility function is quasilinear and given by $u(w, e) = w - c(e)$. The function $c(e) = e^{\eta}$ describes the cost of learning effort, with a constant elasticity $\eta > 1$.

Students differ in their ability to transform effort into examination results, as expressed by a student's learning productivity γ. At all schools, γ is distributed according to a uniform distribution over the interval $[0, \overline{\gamma}]$, with density $f = 1/\overline{\gamma}$ and c.d.f. $F(\gamma) = f\gamma$ for $0 \leq \gamma \leq \overline{\gamma}$. The performance of a student at the examination is γe, and the standard set by the school is denoted by $s \geq 0$. A student with learning productivity γ who expands effort e graduates if and only if $\gamma e \geq s$.

Employers only observe whether a student graduates or not, whereas the actual examination performance γe, the learning productivity γ, and the effort e are private information of the student. By consequence, wages for graduates and non-graduates may differ, but wages cannot be conditioned on γ, e, or γe. In such a situation there is no reward to a student for exceeding the standard required for graduation. By consequence, a student with learning productivity γ will either expand just enough effort to satisfy the standard, $e = s/\gamma$, or she will dispense no effort at all, $e = 0$, and fail at the examination.

The wage w_o received by non-graduates is normalised to zero. Denoting by $\tilde{w}$ the wage which a graduate from a given school may expect in the labour market, for a student of this school graduation is worthwhile if $\tilde{w} - c\,(s/\gamma) \geq w_o - c(0) = 0$. For any standard s and expected wage $\tilde{w}$, the graduation threshold $\gamma(s, \tilde{w})$ is defined to be the solution γ to the equation

$$\tilde{w} - c\left(\frac{s}{\gamma}\right) = 0\,. \tag{1}$$

All students whose learning productivity is at least as high as the graduation threshold, $\gamma \geq \gamma(s, \tilde{w})$, will graduate, and all those with $\gamma < \gamma(s, \tilde{w})$ will not. For $0 \leq \gamma(s, \tilde{w}) \leq \overline{\gamma}$, the number of graduates from this school is then $1 - F\big(\gamma(s, \tilde{w})\big) = 1 - f\gamma(s, \tilde{w})$. From (1) and the identity $\tilde{w} = c\big(s/\gamma(s, \tilde{w})\big)$, we find the elasticities

$$\frac{\partial\gamma(s, \tilde{w})}{\partial s}\frac{s}{\gamma(s, \tilde{w})} \quad = \quad 1\,, \tag{2}$$

$$\frac{\partial\gamma(s, \tilde{w})}{\partial \tilde{w}}\frac{\tilde{w}}{\gamma(s, \tilde{w})} \quad = \quad -\frac{\gamma(s, \tilde{w})\, c\big(s/\gamma(s, \tilde{w})\big)}{sc'\big(s/\gamma(s, \tilde{w})\big)} \quad = \quad -\frac{1}{\eta} \quad > \quad -1\,. \tag{3}$$

That is, the graduation threshold rises proportionately with the standard, and decreases less than proportionately if the expected graduate wage increases.

Each school has an equal number of students, normalised to unity. There are two sets of schools $C = L, H$, where we denote also the numbers of the schools in both sets by L and H. The set L (H) contains schools with a student body originating from a disadvantaged (favoured) social background. As a convenient, if over-simplifying, label we call the former the "lower-class schools" and the latter the "higher-class schools". For example, such social segregation in schools may be the result of Tiebout sorting in the local property market combined with substantial costs of commuting to schools located far away from the student's residence. The sets L and H are interpreted as containing all schools with a given social background which supply graduates to the same regional labour market. As an interesting and plausible case, we specifically consider the possibility that lower class workers are less mobile than higher class workers. Then the relevant labour market is smaller for lower class schools than for higher class schools, i.e., $L < H$.

Conditional on the standard s_i required by a school $i \in C, C = L, H$, employers' willingness to pay for a graduate from school i is $\lambda_C s_i$. This formulation expresses the idea that the examination performance s_i determines productivity at work, which for simplicity is measured in the same units. Moreover, social origin affects the wages according to the parameters λ_C, where we assume $0 < \lambda_L \leq \lambda_H = 1$. That is, the wage paid to graduates from lower class schools is lower by the exogenous factor $\lambda_L \leq 1$. This parameter may reflect properties of disadvantaged students which are relevant for their productivity at the workplace but not tested in the examination, for example good manners, rhetorical abilities, stable families, belonging to social networks, or all sorts of "soft skills". As an alternative interpretation, λ_L might be identified with outright discrimination against disadvantaged workers in the sense that they are being paid less than workers from favourable origins in spite of identical productivity.[2]

There might be other reasons why social origin could be relevant for schooling outcomes. Specifically, disadvantaged students might enter school with an inherently lower ability, or might be less willing to exert effort so as to succeed in school. While it would be easy to integrate such differences in the model[3] we focus on labour market prospects so as to emphasise that class specific standards need not be the consequence of lower ability or a lack of willingness to learn on the part of the lower class students.

Employers do not observe the standard s_i required by an individual school but they observe the social origin of the school's students. Such an informational scenario will occur, for example, if the residences of disadvantaged students are clustered in space so that the location of a school contains information about the social background of the school's students. In the same time, it may be too costly for firms to monitor the grading standards of individual schools. Consequently, wages may differ between higher-class and lower-class schools but not according to the graduation standards of the individuals schools.

Denoting, for the schools $i \in C$, by s_i the standards set by these schools and by $\tilde{w}_i$ the wages expected for their graduates, the wage paid to graduates from any school $i \in C, C = L, H$, is given by

$$w_C = \sum_{i \in C} \lambda_C s_i \frac{1 - F\big(\gamma(s_i, \tilde{w}_i)\big)}{\sum_{j \in C} \big[1 - F\big(\gamma(s_j, \tilde{w}_j)\big)\big]} \,. \tag{4}$$

[2]Discrimination is difficult to rationalise when firms maximise profits. Since we do not explicitly describe firms' hiring choices, our model does not rule it out, however.

[3]These variants could be modelled by compressing the ability distribution, and by inflating the effort cost function, for lower class schools by factors analogous to λ_L. From (1), it is clear that these modifications affect the learning decision of lower class students essentially in the same way as a depressed wage.

That is, the wage is given by a weighted average of the wages which would, under full information, be paid to the graduates from the schools in the relevant labour market, where the weights are given by the shares of the individual schools in the total number of graduates. In an equilibrium the wage is correctly anticipated by students when they choose their effort levels. Thus, for any vector of standards $(s_i)_{i\in C}$, an equilibrium wage is a fixed point of (4) satisfying $w_C = \tilde{w}_i$, for all $i \in C$. Since for all expected wages $(\tilde{w}_i)_{i\in C}$, the right hand side of (4) is just a weighted average of the values $\lambda_C s_i$ for all schools, for all vectors of standards such a fixed point exists in the interval $\Big[\lambda_C \min_{i\in C}\{s_i\}, \lambda_C \max_{i\in C}\{s_i\}\Big]$.

To see how a school's choice of standard affects the equilibrium wage in class $C = L, H$, insert $w_C = \tilde{w}_i$ for all $i \in C$ in (4) so as to find

$$\sum_{i\in C}\Big[w_C - \lambda_C s_i\Big]\Big[1 - F\big(\gamma(s_i, w_C)\big)\Big] = 0\,. \tag{5}$$

Differentiating the equilibrium condition (5) implicitly, one obtains

$$\frac{\mathrm{d}w_C}{\mathrm{d}s_i} = -\frac{f\Big[\lambda_C s_i - w_C\Big]\dfrac{\partial\gamma(s_i, w_C)}{\partial s} - \lambda_C\Big[1 - F(\gamma(s_i, w_C))\Big]}{\sum_{j\in C}\Big[1 - F(\gamma(s_j, w_C))\Big] + f\sum_{j\in C}\Big[\lambda_C s_j - w_C\Big]\dfrac{\partial\gamma(s_j, w_C)}{\partial\tilde{w}}}\,. \tag{6}$$

In the following, we focus specifically on symmetric situations where all schools $i, j \in C$ of one class choose identical standards $s_i = s_j = s_C$, implying an identical graduation threshold $\gamma_C = \gamma(s_C, w_C)$. Then the equilibrium wage is uniquely determined and equal to $w_C = \lambda_C s_C$ for all schools in C. Moreover, starting from symmetric standards, the comparative static equation (6) reduces to $\mathrm{d}w_C/\mathrm{d}s_i = \lambda_C/C$. This equation will be used in the analysis of the standards set by schools, to which we now turn.

3 Optimal Standards

We assume that each school maximises the sum of the wages earned by its students. Schools thus care for their students, without however taking effort costs into account.[4] When deciding about the standards they require for graduation, schools anticipate the optimal choices by students and the equilibrium wage. If school $i \in C$ sets standard s_i, it thus expects that the wage for graduates from class C will be w_C according to (5), taking the standards $(s_j)_{j\in C, j\neq i}$ chosen by all other schools in the market as given. School i's maximisation problem can thus be stated as

$$\max_{s_i\geq 0} \quad W_i(s_i) = F\big(\gamma(s_i, w_C)\big)w_o + \big[1 - F\big(\gamma(s_i, w_C)\big)\big]\, w_C\,.$$

[4]This omission reflects current debates in education policy which do not seem to be very concerned about students enjoying insufficient leisure.

With $w_o = 0$ we obtain the necessary condition for an interiour solution:

$$\frac{\partial W_i}{\partial s_i} = -f w_C \left[\frac{\partial \gamma(s_i, w_C)}{\partial s} + \frac{\partial \gamma(s_i, w_C)}{\partial \tilde{w}} \frac{\mathrm{d} w_C}{\mathrm{d} s_i} \right] + \left[1 - F\big(\gamma(s_i, w_C)\big) \right] \frac{\mathrm{d} w_C}{\mathrm{d} s_i} = 0\,. \tag{7}$$

Condition (7) shows the trade-off faced by a school. On the one hand, as expressed by the first term in square brackets in (7), a more demanding standard decreases welfare by reducing the number of graduates. On the other hand, a higher standard raises the wage for graduates. This enhances welfare both directly, as measured by the last term in (7), and indirectly by increasing the number of graduates. This effect, which is formalised by the second term in the square brackets in (7), counteracts the decline in the graduation rate triggered by the higher standard.

In order to characterise the equilibrium, we write $\gamma_i = \gamma(s_i, w_C)$ and use (2), (3), and $F(\gamma_i) = \gamma_i/\overline{\gamma}$ so as to restate (7) as

$$\frac{\partial W_i}{\partial s_i} \quad = \quad -\frac{\gamma_i w_C}{\overline{\gamma} s_i} + \frac{\mathrm{d} w_C}{\mathrm{d} s_i} \left[1 - \frac{\gamma_i}{\overline{\gamma}} \left(1 - \frac{1}{\eta} \right) \right] \quad = \quad 0\,. \tag{8}$$

By inserting $s_i = s_C, \gamma_i = \gamma_C, w_C = \lambda_C s_C$ and $\mathrm{d} w_C / \mathrm{d} s_i = \lambda_C / C$ in (8),[5] we find the graduation threshold and, implicitly, the standard in a symmetric equilibrium:

$$\gamma_C^* = \frac{\overline{\gamma}}{C + 1 - (1/\eta)} \qquad \text{and} \qquad \lambda_C s_C^* - c\left(\frac{s_C^*}{\gamma_C^*} \right) = 0 \qquad \text{for } C = L, H. \tag{9}$$

In (9), the limiting case $C = 1$ represents a market consisting of only one school. Since in this special case each school effectively determines its own graduate wage, this is equivalent to a scenario where the employers have full information about the standards applied by each individual school.

From the first equation in (9) we note that $H > L$ implies $\gamma_H^* < \gamma_L^*$. Thus, if the market for graduates from higher class schools is larger than the market for disadvantaged graduates, then the graduation rate will be larger among the higher class students than among the lower class students. Higher graduation rates among better off students therefore need not be the result of superior abilities. Quite the contrary, a higher learning productivity $\overline{\gamma}$ will raise the graduation threshold γ_H^* and hence reduce graduation rates.

In order to obtain comparative static results for the equilibrium standard, we differentiate the second equation in (9), observing that γ_C^* depends on $\overline{\gamma}$ as given in the first equation in (9). Using $c(s_C^*/\gamma_C^*) = (s_C^*/\gamma_C^*)^\eta = \lambda_C s_C^*$ we arrive at:

$$\frac{\mathrm{d} s_C^*}{\mathrm{d} \lambda_C} = \frac{s_C^*}{\lambda_C(\eta - 1)} > 0 \qquad \text{and} \qquad \frac{\mathrm{d} s_C^*}{\mathrm{d} \overline{\gamma}} = \frac{\eta s_C^*}{\overline{\gamma}(\eta - 1)} > 0 \qquad \text{for } C = L, H. \tag{10}$$

[5] In the appendix it is shown that at a symmetric solution to (8), the second order condition for a maximum is satisfied.

Specifically, for the grading policy of lower class schools this implies:

Proposition 1 *The larger the wage discount for graduates from disadvantaged social backgrounds, and the lower the learning productivity of such students, the lower is the standard chosen by a school with students from lower social classes.*

This result shows that a school which cares about the incomes of their students will grade more leniently if its students are socially disadvantaged. As one may expect, such behaviour may simply be the consequence of lower abilities on the part of students from lower social classes. Proposition 1, however, shows that more lenient grading may just as well be the rational reaction of a school to the unfavourable job prospects of its graduates.

Figure 1 illustrates the students' choices of learning effort in the full information case $L = H = 1$. The steeper (flatter) straight line gives the wage obtained by graduates of the higher (lower) class school as a function of the standard. The convex curves describe the effort cost for students of various learning productivities. As the learning productivity increases from γ_o towards γ_1, these curves bend downwards. The maximal standard a student is willing to satisfy is determined by the intersection of the class-specific wage line with the effort cost curve corresponding to the student's learning productivity. A student with learning productivity γ^* will graduate[6] if the standard does not exceed the value s^*_C derived by the intersection of $c(s, \gamma^*)$ and the wage line corresponding to her social origin, w_L or w_H. As a consequence of the lower wage, this standard must be lower for a lower class student.

As a next result, we find from (9):

$$\frac{\mathrm{d}s^*_C}{\mathrm{d}C} = \frac{\eta s^*_C}{(1-\eta)[C + 1 - (1/\eta)]} < 0 \,. \tag{11}$$

Proposition 2 *A smaller market size C raises the standard s^*_C.*

This result illustrates the well-known grading externality among schools sharing a common labour market. If a lower class school $i \in L$ marginally lowers its standard the willingness to pay for a graduate from this school decreases by λ_L. Since school i has only weight $1/L$ in the group of lower class schools this translates only into a wage decrease of λ_L/L. Schools therefore have an incentive to free ride on the high wages brought about by the tough standards of other schools, by grading leniently themselves. The result is a general devaluation of standards which is the more pronounced the larger the market is.

As a consequence of Proposition 2, one conjectures that the equilibrium standard of lower class schools may be tougher than the one required by higher class schools if the market size for lower class graduates is sufficiently small. Making this intuition precise, we observe that

[6] Note that from (9), the graduation thresholds are equal, $\gamma^*_L = \gamma^*_H = \gamma^*$, if $L = H$.

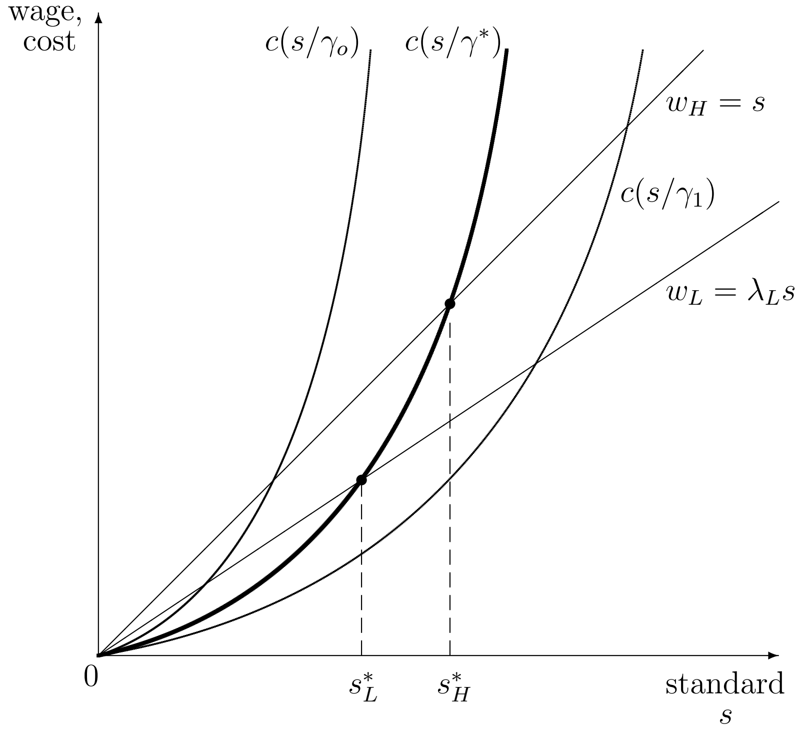

Figure 1: Graduation standards and graduation threshold of schools with students from different social backgrounds.

$s_L^* \geq s_H^*$, from $\eta > 1$, is equivalent to $s_L^{*(\eta-1)} \geq s_H^{*(\eta-1)}$. Using the second equation from (9), $\lambda_H = 1$, and the definion of the cost function $c(s_C^*/\gamma_C^*) = (s_C^*/\gamma_C^*)^\eta$, this inequality can be transformed into $\lambda_L \gamma_L^* \geq \gamma_H^*$. Replacing the graduation thresholds with the help of the first equation in (9), we find

Proposition 3 *The standard chosen by lower class schools is higher than or equal to the standard chosen by higher class schools if and only if*

$$\lambda_L \geq \left(\frac{L + 1 - \frac{1}{\eta}}{H + 1 - \frac{1}{\eta}} \right)^\eta . \tag{12}$$

According to Proposition 3, lower class schools demand a tougher standard than higher class schools if the wage discount for disadvantaged students is moderate compared to the difference in market sizes between the two kinds of schools.

We now turn to the empirical analysis of the interaction of standards and the social composition of schools. As an introduction, we give a brief account of the education system in the Netherlands.

4 Data and estimation approach

4.1 The Dutch education system

Dutch compulsory education encompasses twelve school years. At age five it starts with primary education (*Basisschool*) which lasts eight years. Today, parents may choose among three types (*opleidingen*) of institutions in secondary education (*Voortgezet Onderwijs, VO*):[7]

(i) Pre-vocational or middle level secondary education (*Voorbereidend middelbaar beroepsonderwijs, VMBO*) lasts four school years and comprises four main branches: The *Kaderberoepsgerichte leerweg (KB)*, the *Basisberoepsgerichte leerweg (BB)*, the *Theoretische leerweg (TL)* and the *Gemengde leerweg (GL)*. The latter two *(VMBO-GT)* are focussed on a more theoretical approach, whereas students in the former two branches *(VMBO-BK)* must partake in a practical central examination that replaces one theoretical subject. Most *VMBO* students move on to vocational training after graduation, but it is also possible to proceed to 4th grade of *HAVO* education (see below). The *VMBO* branch was established in 1999, replacing the *Middelbaar algemeen voortgezet onderwijs (MAVO)* (equivalent to *VMBO-GT*) and *Voorbereidend beroepsonderwijs, VBO* (equivalent to *VMBO-BK*) branches. As we will be using data from the classes of 2002 and 2003, in our analysis the branches are still referred to as *MAVO* and *VBO*, where the 2003 *VMBO-GT* graduates are counted as *MAVO* graduates.

(ii) Senior or higher general secondary education (*Hoger Algemeen Voortgezet Onderwijs, HAVO*) amounts to five years of schooling and is aimed at providing students with a general education and preparing them for entry into higher professional education (*HBO*), which leads to a bachelor's degree. Here as well, graduates can enroll in fifth grade *VWO* (see below) upon graduation rather than proceeding to *HBO*. Alternatively, they might opt for vocational training.

(iii) Pre-university education (*Voorbereidend Wetenschappelijk Onderwijs, VWO*) encompasses six years of schooling. Its goal is to enable students to take up a university education; it is thus the highest form of secondary education in the Netherlands.

At first glance, Dutch secondary education appears to be characterised by central standards, as students end their scholastic careers with central examinations. However, central exams (*centraal examen*) account for only half the final grade. The other half is determined via decentralised testing (*schoolexamen*),[8] leaving grading and standard-setting to a large extent at the individual school's discretion.

[7]Most Dutch schools offer more than one of these *opleidingen*, and often one school will provide access to all three branches of secondary education.

[8]In the basic vocational programme *VMBO-BB* the school exam accounts for two thirds of the final grade.

The *centraal examen* are centrally arranged and graded by the testing agency *CITO* (*Central Institute for Test Development*).[9] All students of the same *opleiding* are faced with identical questions and grading is done by *CITO* within 4 weeks' time. An official body, appointed by the Ministry of Education, *CEVO (Centrale examencommissie vaststelling opgaven)* establishes the norms for the central exams.[10] The school exams on the other hand are conducted and –more importantly– devised and graded by the local schools. There are, however, guidelines set by the department of education (*Ministerie van OCW*) concerning the subject matter covered in school exams, to which schools must abide. To this end, the local schools set up "exam rules" (*examenreglement*), which establish the curriculum and required reading for the local exams. The *examenreglement* needs to be accredited by the central authorities and is accessible to the respective school's students. Nonetheless, it is obvious that in essence it is the individual school which sets the standard, at least within a certain range.

4.2 Estimation approach

Our empirical approach uses the co-existence of central and school-specific grades in order to detect differences in local standards. Let G_i^c denote the average *centraal examen* grade and G_i^s the average *schoolexamen* grade in school i. Under coinciding central and school specific standards, we would expect $G_i^s = G_i^c$. An upward deviation of G_i^s from G_i^c then constitutes a local standard that falls short of the central standard and vice versa. Our (inverse) operationalisation for the standard s_i applied by school i is therefore the difference $\Delta G_i = G_i^s - G_i^c$ between the average grades obtained at this school in the school specific and in the central examination.

On a formal level, the continous variable ΔG_i departs from the binary pass-fail standard s_i featuring in the theoretical model. Since it is likely that a school which grades leniently also awards degrees more easily, it is, however, plausible that the factors determining graduation standards affect average grades in a similar way. Moreover, it is plausible that many employers and universities require a certain minimum grade from applicants whom they are willing to consider seriously. In such a case, this is the standard a student must meet, and the grading scale effectively determines a binary standard.

The difference ΔG_i cannot in itself be interpreted normatively. It does not say whether the school specific or the central standard is correct in the sense of measuring the "true" skill level of students. A positive ΔG_i might be a correction for an overambitious central standard rather than grade inflation by school i.[11] In this paper we will not, however,

[9] For further information, go to: http://www.cito.nl/com_assess_ex/nat_final_ex/eind_fr.html

[10] In some subjects, no centralised testing occurs (e.g. physical education and arts). The analysis in this paper is limited to subjects where both types of testing are employed and *schoolexamen* grades can thus be compared to *centraal examen* grades.

[11] It can be ruled out, however, that school grading is conditional on central grades, as the *centraal examen* is the last exam of the entire school career.

question the appropriateness of the central grading scheme and therefore accept it as the benchmark against which double standards are to be detected. This is justified by the main focus of our investigation. We are not primarily interested in grade inflation in itself. It may well be that on a local level teachers tend to award higher or lower grades in general, say because school exams tend to be standardised in a different manner than central exams. We would then expect ΔG_i to be different from zero but constant across schools. Our focus, instead, is on double standards depending on social status. If standards are socially differentiated, ΔG_i will be systematically affected by the social composition of the schools' student body, whatever the average deviation between school specific and central grades.

Such a systematic effect, if it is found, may be the consequence of diverging grading schemes or of a different choice of examination topics. If the first is the case, all schools ask essentially the same questions at the school specific examination but those with disadvantaged students then grant higher marks for any given answer. Alternatively, ΔG_i may be higher in schools with disadvantaged students because teachers ask questions which are tailored to the students' knowledge. Also in this case, however, we consider the label "double standards" to be appropriate since it does not really matter whether grades are better because expectations are lower, or because difficult topics are avoided.

Another property of ΔG_i is that it is unaffected by peer and sorting effects that may emerge in a system of school choice.[12] While sorting by ability will evidently have a massive influence on average central grades in a given school, the grade difference should not be affected, as we would expect school grades to change at the same rate. The same is true for possible peer effects that arise from sorting. Peer effects should impact both grades and thus leave ΔG_i unchanged.

In order to explain the grade difference ΔG_i, we use the estimation equation:

$$\Delta G_i = \beta_0 + \beta_1 \cdot y_i + \beta_2 \cdot x_i + \epsilon_i, \tag{13}$$

where i denotes the individual school, y_i are variables describing the student body's social composition, x_i is a vector of control variables, and ϵ_i is the error term. We will focus on two variables which capture the school-level social composition y_i:

(i) The percentage of students considered cultural minority students.

(ii) The percentage of students receiving federal study cost allowance *(Tegemoetkoming studiekosten)*, eligibility for which implies that parents have a low disposable income.

In accordance with Proposition 1, we expect to find decreasing local standards with increasing school-level percentages (i) and (ii). That is, if double standards are employed,

[12] There is a vast literature both theoretical and empirical on peer effects, see e.g. Epple and Romano (1998). Hsieh and Urquiola (2006) discuss the problems that arise in disentangling sorting, peer and competition effects when measuring changes in school quality.

we expect β_1 to be positive. According to Proposition 2, an increase in market size leads schools to set lower educational standards. While we cannot provide a direct test of this hypothesis, studying the different branches of secondary education will shed some light on this effect. It is reasonable to think of the market size to be increasing in the level of education. That is, the relevant labour market is smallest in geographical terms for students who have earned a diploma in *VBO* and largest for those who hold a *VWO*-diploma. Hence, the above estimation will be conducted for all school branches separately. The cutting of standards as measured by β_1 is suspected to be largest in *VWO*, smallest for those with a diploma in *VBO*, and in between for the other two branches.

4.3 Data sources

The data employed in this analysis stems from four different sources. School-level data concerning students' performance and social affiliation as well as schools' characteristics is taken from the *Kwaliteitskaart Voortgezet Onderwijs* (Quality Cards for Secondary Education), issues 1998 − 2004. The *Kwaliteitskaarten* are published on a yearly basis by the Netherlands Inspectorate of Education for all Dutch secondary schools. The dataset provides information on number of students, administrative form of the school (private/denominational/public), the school branches that can be attended at the school, average class sizes, subject-level average grades attained in school and central exams, the recommended type of secondary school based on students' performance in primary education (i.e. students' entrance levels of performance), the percentage of ethnic minority students, the percentage of students with a study cost allowance etc.

The dependent variable ΔG_i is constructed from the performance data of the 2003 and 2004 *Kwaliteitskaarten*. As the original *Kwaliteitskaarten* file contains interdisciplinary average grades only as a mean of school and central grades, we calculated the average school specific (G_i^s) and central (G_i^c) grades covering all subjects by weighing the average school and central grades in each subject with the number of students that had actually taken part in the exams in that particular subject.

The percentage of cultural minority (*CUMI*) students is defined as the share of students in a given school branch who have a non-Dutch background.[13] Along with the percentage of students receiving study cost allowance we use this variable as a proxy for low social status, as neither the Inspectorate nor the individual schools collect detailed data on parents' socioeconomic status. At the end of *basisschool* each student is given a non-binding advice by her teachers as to which school branch is deemed appropriate in secondary education.

[13] A student is considered part of a cultural minority if she satisfies one of the following criteria: both parents were born in (or have nationality of): one of the republics of former Yugoslavia, Greece; Italy; Cape Verde, Morocco, Portugal, Spain, Tunisia or Turkey; Moluccan background; Surinamese, Antillean or Aruban background; Roma background; caravan dwellers; other non-European background and not having completed full primary education in the Netherlands; Eastern European background and not having completed two years of Dutch schooling.

We add this advice as control variable for the incoming students' skill level. "Above advice" ("below advice") denotes a student attending a more (less) demanding branch than the one recommended. We also use the percentage of students in ability-tracked classes in the second year of secondary education and the average class size from the *Kwaliteitskaarten*.[14]

In addition to school level data, we use some variables which are available on a ZIP-code level only. Specifically, the *Statusscores postcodegebieden* are ZIP-code level data proxying for the students' social background. They are supplied by The Social and Cultural Planning Office of the Netherlands (SCP), a Dutch government agency. The status scores are calculated in 4 year intervals, taking into account variables such as mean education, mean income, average rents etc. Postcode areas that have a low social status are denoted with values greater than zero, areas of higher status receive negative values. We match these scores with the schools' 4-digit postcodes taken from the *Kwaliteitskaarten*. More data on a postcode level comes from the *Kerncijfers postcodegebieden 2003* as well as the *Kerncijfers wijken en Buurten 2001-2005*, published by the Dutch Office of Statistics (*CBS*). The percentage of school-aged children is calculated from the dataset *Bevolking per 4-cijferige postcode 2004*, published by the *CBS*.

Since school-level financial endowment as well as characteristics of the teaching staff might influence average grades and standard setting, data from the series *Onderwijs in Cijfers (OIC)* is used in the estimation, too. *Onderwijs in Cijfers* is published annually by the Dutch Ministry of Education and is intended to provide school managers with information on the above mentioned matters for all Dutch secondary schools.

Table 1 displays descriptive statistics for the exogenous variables used in the estimation.

5 Empirical Results

5.1 Determinants of standards in Dutch schools

Descriptive statistics for the difference between school specific and central grades, ΔG_i, are shown in Table 2 for the pooled classes of 2002 and 2003. On average, grades awarded in school exams are higher than those awarded in central exams in all branches but VBO. Thus, the local school standards in these branches seem to be - on average - lower than the centrally devised standard. The difference is highest for pre-university education and lowest for *VBO* schools. The minimum values also indicate that on a school-level, *VWO* and *HAVO schoolexamen* grades drop only slightly below the central grades, at most. In *MAVO* and *VBO* education, however, some schools underscore more heavily in the school exams. Altogether, it seems that the schools in the higher branches of secondary education

[14]Ability-tracked in this context means that students attend classes with students from their chosen branch only, whereas non-tracked students attend classes together with students from other branches. After the second year of secondary education there are no mixed classes.

Table 1: Descriptive statistics

Variable	N	Mean	Std. Dev.	Min	Max
Above advice % VWO	856	22.380	16.983	0	100
Above advice % HAVO	852	12.756	12.311	0	88
Above advice % MAVO	1013	6.537	12.806	0	96
Above advice % VBO	619	12.753	16.036	0	100
Below advice % HAVO	852	4.579	4.950	0	43
Below advice % MAVO	1013	14.236	11.006	0	82
Below advice % VBO	619	16.436	15.099	0	100
Minority students % VWO	1026	4.320	8.371	0	78.947
Minority students % HAVO	979	6.141	10.104	0	74.419
Minority students % MAVO	1176	8.089	12.673	0	100
Minority students % VBO	806	14.037	16.638	0	100
Study cost % VWO	867	27.003	10.611	0	100
Study cost % HAVO	834	37.712	13.852	0	100
Study cost % MAVO	1110	41.242	20.248	0	100
Study cost % VBO	648	65.439	22.758	0	100
Tracked %	1479	63.204	37.208	0	100
Class size	1497	23.327	4.179	6	31.667
No. Students (x1000)	1553	0.871	0.503	.027	2.945
Short term debt (share of balance)	1450	0.306	0.112	0	0.868
Long term debt (share of balance)	1450	0.036	0.067	0	0.684
Staff growth	1422	0.035	0.052	-.255	0.272
No. students growth	1438	0.016	0.053	-.204	0.369
Part time staff (share)	1422	0.361	0.092	.091	0.852
Status ZIP	1548	0.175	1.037	-2.745	3.889
Avg. income ZIP (x1000)	1520	13.417	2.374	8.777	27.677
School aged ZIP %	1552	17.131	3.643	2.564	37.5
Population ZIP (x1000)	1532	8.434	4.226	0.040	42.78

reduce the standards by more than those at the lower end. This is consistent with the idea of geographically larger job markets causing lenient grading.

We carried out OLS regressions for all four branches of secondary education separately, with the branch-specific difference ΔG_i as the dependent variable. We will first describe the results for pre-university education *VWO* (Table 3). Specification (1) includes typical school-level variables only, (2) adds *Onderwijs in Cijfers* variables and (3) includes postcode-level data as well. In accordance with the theoretical predictions, in specifications (1) to (3) we find that an increasing percentage of cultural minority students leads to lower standards in local exams. The share of students eligible for study cost allowance does not seem to be linked to lenient grading. One reason is that a large percentage of minority students also qualify for study cost allowance, resulting in overlapping effects for the two variables. This becomes obvious if we interchangeably employ only one of these two proxies for social status. Omitting the percentage of minority students from the estimation (Specification (4)) results in a considerably larger and significant effect of the

Table 2: Summary statistics ΔG_i 2002/2003

Variable	**N**	**Mean**	**Std. Dev.**	**Min.**	**Max.**
VWO	835	4.667	2.535	-2.093	16.446
HAVO	786	2.549	2.309	-4.985	12.422
MAVO	1101	1.602	2.768	-6.737	17.079
VBO	646	-0.115	3.396	-9.222	14.5

share of study cost recipients on standard depreciation. Dropping the study cost allowance recipients yields similar effects on the minority share coefficient, strengthening the idea, that these variables are to some extent congruent (Specification (5)).

The third variable accounting for social status (Status ZIP) further supports our hypothesis. Lower status on a ZIP code level also leads to higher grading differences (remember, the status variable is coded inversely). Somewhat surprising is the fact that higher ZIP-level incomes also lead to larger gaps in grading standards. This is difficult to explain, but it suggests that status and income do not measure the same thing.

We also find that public schools tend to inflate grades more than private schools, possibly due to competitive pressure.[15] In addition, we observe that indebted schools tend to grade more leniently. Setting lower standards may be an attempt of these schools at attracting more students and exploiting economies of scale.[16] A higher percentage of students attending a school branch deemed too demanding for them in their advice after *basisschool* also leads to a decline in standards. Here, a high share may be an indicator of competitive pressure leading schools to admit students above their initial advice.[17] A low percentage of ability-tracked students is also linked with the deterioration of standards.

Table 4 extends the analysis to the other three school branches, where column (1) describes *HAVO*, column (2) *MAVO*, and column (3) *VBO* education. The main result holds for *HAVO* as well as *MAVO*: A lower class student body, if measured by the share of cultural minority students, causes a depreciation in standards. The coefficient for study cost allowance recipients, however, is not significant. It becomes so upon omission of the cultural minority variable for *HAVO* and MAVO.[18] Interestingly, the coefficients of the share of cultural minority students in *HAVO* and *MAVO* are only about half as large as the one found for *VWO* (see specification (3) in Table 3). Moreover, the grade gap in *VBO* education does not at all seem to be related to social composition of the student body. In fact, the explanatory power of the *VBO* model is completely driven by the year dummy. Together, we take these results as tentative evidence in favour of the hypothesis,

[15]School choice leads to approximately 70% of Dutch students actually attending private schools, many of which are denominational.

[16]Dutch schools receive a fixed federal transfer for every student.

[17]If there were only the central examination, such admission practices would not make much sense, as the students might just fail to attain the diploma.

[18]We do not report the regressions for these specifications. They are available upon request.

Table 3: Estimation results for ΔG_i in pre-university education VWO

	(1)		**(2)**		**(3)**		**(4)**		**(5)**	
Above advice %	0.031***	(0.005)	0.032***	(0.006)	0.032***	(0.006)	0.043***	(0.006)	0.031***	(0.007)
Minority students %	0.089***	(0.015)	0.114***	(0.016)	0.112***	(0.016)			0.117***	(0.014)
Study cost %	0.013	(0.010)	0.001	(0.011)	0.007	(0.011)	0.046***	(0.010)		
Tracked %	-0.009***	(0.002)	-0.010***	(0.002)	-0.011***	(0.003)	-0.011***	(0.003)	-0.010***	(0.002)
Avg. class size	-0.106***	(0.036)	-0.084**	(0.040)	-0.085**	(0.040)	-0.121***	(0.041)	-0.089**	(0.040)
Public dummy	0.390**	(0.180)	0.286	(0.191)	0.362*	(0.194)	0.407*	(0.201)	0.363*	(0.194)
No. of students	0.216	(0.195)	0.262	(0.220)	0.289	(0.224)	-0.006	(0.228)	0.297	(0.224)
Short term debt			2.515***	(0.808)	2.165***	(0.812)	2.423***	(0.841)	2.120***	(0.809)
Long term debt			2.880**	(1.125)	2.201*	(1.134)	1.592	(1.171)	2.223**	(1.133)
Staff growth			0.359	(1.896)	0.222	(1.903)	0.083	(1.971)	0.251	(1.901)
No. students growth			-1.925	(2.167)	-1.913	(2.177)	-1.359	(2.254)	-1.867	(2.175)
Part time staff %			-0.034	(0.984)	-0.306	(1.057)	-0.408	(1.095)	-0.355	(1.054)
Status ZIP					0.305**	(0.152)	0.398**	(0.157)	0.312**	(0.151)
Avg. income ZIP					0.135**	(0.062)	0.176***	(0.064)	0.133**	(0.062)
Share school aged ZIP					0.066**	(0.030)	0.056*	(0.031)	0.065**	(0.030)
Population ZIP					-0.068***	(0.021)	-0.066***	(0.022)	-0.067***	(0.021)
Year 2003	-0.170	(0.165)	-0.131	(0.176)	-0.162	(0.177)	-0.312*	(0.182)	-0.140	(0.174)
Intercept	6.306**	(1.060)	5.106***	(1.279)	2.791*	(1.693)	2.829*	(1.754)	3.108**	(1.625)
N	759		649		635		635		635	
Adj. R^2	0.208		0.237		0.257		0.203		0.258	

Standard errors in parentheses.

*, **, *** designate significance at the ten, five, and one percent level respectively.

Table 4: Estimation results for ΔG_i: HAVO (1), MAVO (2), and VBO (3)

	(1)		**(2)**		(3)	
Above advice %	0.045***	(0.008)	0.045***	(0.007)	0.010	(0.009)
Below advice %	-0.003	(0.017)	-0.034***	(0.008)	-0.020*	(0.010)
Minority students %	0.058***	(0.011)	0.043***	(0.007)	0.012	(0.009)
Study cost %	0.007	(0.008)	0.002	(0.005)	0.008	(0.006)
Tracked	-0.004*	(0.002)	-0.005**	(0.002)	-0.004	(0.004)
Avg. class size	0.032	(0.037)	-0.119***	(0.026)	-0.075	(0.051)
Public dummy	0.374**	(0.189)	-0.132	(0.170)	-0.435	(0.291)
No. of students	0.286	(0.215)	0.030	(0.154)	0.234	(0.264)
Short term debt	0.484	(0.775)	0.262	(0.707)	-2.503**	(1.211)
Long term debt	1.303	(1.263)	-2.133*	(1.171)	-2.805	(1.941)
Staff growth	-1.729	(1.836)	1.366	(1.628)	-1.796	(2.859)
No. students growth	-0.776*	(2.076)	-0.047	(1.784)	-3.654	(2.824)
Part time staff %	0.635	(1.097)	-0.887	(1.029)	2.386	(2.113)
Status ZIP	0.573***	(0.144)	0.292**	(0.135)	0.057	(0.258)
Avg. income ZIP	0.276***	(0.061)	0.236***	(0.063)	0.251	(0.175)
Share school aged ZIP	0.078**	(0.030)	0.010	(0.031)	0.022	(0.064)
Population ZIP	-0.023	(0.019)	0.010	(0.018)	0.050	(0.031)
Year 2003	0.130	(0.171)	-1.010***	(0.155)	3.527***	(0.264)
Intercept	-5.004***	(1.626)	1.740	(1.301)	-4.505	(3.525)
N	597		802		454	
Adj. R^2	0.215		0.280		0.315	

Standard errors in parentheses.
*, **, *** designate significance at the ten, five, and one percent level respectively.

stated in Proposition 2, that schools supplying graduates to smaller job markets tend to inflate grades by less.

Contrary to pre-university education *VWO*, financial variables do not appear to drive the cutting of standards in either *HAVO* or *MAVO*. An explanation for this difference may be that competition among schools for *VWO* students is more intense than in the case of *HAVO* or *MAVO*, because prospective *VWO* students make more use of their right to choose, or choose among a larger set of schools. Furthermore, if educating *VWO* students is less expensive, schools that offer more than one branch may resort to attracting *VWO* students rather than *HAVO* or *MAVO* students. Most of the other control variables do not differ much across school branches.

In essence, even though some of the control variables' coefficient signs are not as expected, the empirical results for *VWO*, *HAVO* and *MAVO* clearly reject the hypothesis that social composition does not influence the magnitude of standard cutting. Not only do we find significant effects of social composition on grading standards, we also can dismiss concerns that students of low status might, on top of their low status, be discriminated when it comes to grading. The opposite is true. Moreover, the size of the coefficients for the different branches and the estimation's lacking explanatory power for *VBO* seem consistent with the hypothesis that standard cutting increases in relevant labour market size.

5.2 Endogeneity issues

In this section we address the possibility that the share of minority students itself may be a function of the grade difference ΔG_i. Such an endogeneity problem would arise if minority students were to choose schools with more lenient grading while non-minority students do not behave in this manner. Given the absence of catchment areas in the Netherlands, this problem could even be aggravated since students and their parents do not have to move to the vicinity of the desired school.

Intuitively, we have no reason to believe that parents of lower social status care more about their children's grades than their well-off counterparts. Quite the contrary, one would probably expect parents of higher social status to be rather more career-oriented. On top of that, it should be noted that even in the absence of catchment areas, sending an offspring to farther away schools which award better grades entails travel costs and is thus more easily feasible for well-off families.

On a technical level, we carried out instrumental variable (IV) regressions in order to account for possible endogeneity of social status. As a measure of status in these regressions we restrict attention to the minority share as the stronger predictor of grade differences, dropping the study cost variable. We do not report full regression results for the IV estimations. Rather, Table 5 shows the coefficients of the instrumented explanatory variable "minority share" for all four school branches and all instruments.

The first instrument we employ is the population density in the schools' respective postcode areas, the reasoning being that immigrants and subsequent minority generations tend to live in the larger cities and thus in densely populated areas. On the other hand, we do not expect population density to be linked to grade difference other than through the higher minority share.[19] First stage regressions, which are not reported, insinuate that population density is a relevant instrument for all branches. In all four branches, the results are in line with the OLS results and suggest that OLS underestimates the effect of the minority share.

[19]Higher grade differences do not cause a change in population density in the absence of catchment areas as there is no need for relocation. See Karsten (2006) for evidence that place of residence and location of attended school often differ.

Table 5: Coefficients for minority share by school branch and instrument

expl. Variable \ Instrument	OLS	Population density	Postcode	Minority % VWO	Minority % HAVO	Minority % MAVO	Minority % VBO
Minority %	0.118***	0.186***	0.258***		0.124***	0.105***	0.197***
VWO	(0.014)	(0.031)	(0.071)		(0.016)	(0.019)	(0.045)
Minority %	0.062***	0.078***	0.123***	0.061***		0.060***	0.047
HAVO	(0.010)	(0.022)	(0.046)	(0.012)		(0.013)	(0.031)
Minority %	0.045***	0.085***	0.130***	0.045***	0.042***		0.073***
MAVO	(0.007)	(0.015)	(0.028)	(0.011)	(0.010)		(0.016)
Minority %	0.014	0.034	-0.067	0.029	0.020	0.021	
VBO	(0.009)	(0.022)	(0.049)	(0.028)	(0.023)	(0.016)	

Standard errors in parentheses.
*, **, *** designate significance at the ten, five, and one percent level respectively.

The next instrument is also related to geography: the Dutch postcodes. These are arranged so that the west of the country, where the larger cities are located, is assigned low postcode numbers which increase as one moves north-west. The instrument is relevant since minorities historically cluster in the regions with lower postal codes;[20] it is also exogenous as postal codes are not suspected to influence grade differences. Here, the coefficients are larger than those obtained from OLS. This suggests that by using OLS, we rather underestimate the true effect of the minority students' share on standards.

Our final instrument is the share of minority students in a different school branch at the same school. Most schools offer more than one branch, and we expect minority shares to be highly correlated across branches within the same school. Grade differences in one school branch should be independent of the minority share in another branch unless the school applies a school-specific grading policy, regardless of branch. If this was true, however, we would expect to find lenient grading also in *VBO* when a large share of minority students is present. As this is not the case, such common grading policies do not seem to be present. We employed as instruments the minority share in all school branches but the one under consideration, rendering us with three instruments per school branch (e.g. the minority share in *VWO* schools is instrumented by the minority shares in *HAVO*, *MAVO*, and *VBO*). As displayed in Table 5, the coefficients for all school branches are close to the OLS results, regardless of the instrument used.[21] Altogether, the instrumental variable regressions confirm the OLS results.

[20] On ethnical clustering in the Netherlands, see de Graaf, Gorter, Nijkamp (2001).

[21] Whereas population density and postcode are available for all schools, the shares of minority students in other school branches are obviously only available if a certain school offers more than one *opleiding*. This is not always the case, especially when the instrument is not from an adjacent school branch to the instrumented one. Since results are virtually invariant to the choice of school-branch instrument, there is no reason to suspect selection effects, though.

Table 6: Fixed and random effects for VWO branch

	Fixed effects		Random effects	
Minority students %	0.142*	(0.077)	0.138***	(0.019)
Study cost %	0.005	(0.020)	0.018	(0.015)
Study cost × post 2001	-0.006	(0.013)	-0.009	(0.012)
year fixed effects	Yes		Yes	
N	1171		1171	
R^2	$\underbrace{0.4402}_{within}$		$\underbrace{0.2538}_{overall}$	

Standard errors in parentheses.
*, **, *** designate significance at the ten, five, and one percent level respectively.

5.3 Longitudinal data

Lastly, we intend to back up our cross-sectional results with longitudinal evidence. As many control variables are unavailable for the years prior to 2002, we focus on a specification which includes the social status variables and year-fixed-effects only, thus retaining a much larger number of observations. Since we could not find any signs of endogeneity in the status variables in the previous subsection, we treat them as exogenous in the longitudinal analysis.

When the minority share is used as the key explanatory variable in fixed effects (FE) estimation, we cannot obtain any significant results. This is not very surprising because the minority share varies little over time, and we do not expect schools to react immediately to a change in the minority students' share. Rather, schools will establish or adjust a lenient grading policy over the course of time. To account for this fact, we calculate a weighted moving average of the minority shares at the year of graduation and the three years prior. The weights are the number of students in the respective years. This makes sense because the minority share is defined as the share of *CUMI*-students in *opleiding*-population. Hence, the current graduating class influences the *CUMI*-share in all of these four years. Our moving average then describes the school's average minority share while the graduating cohort attended school.

We cannot easily calculate moving averages for the study cost (*TS*) recipients share, as there is a structural break in the data, due to a change in the legal definition of eligibility. This lead to a huge increase in the percentage of students receiving *TS* in the year 2002. Moreover, the new group of recipients might be very different from the earlier one with respect to their social status. We account for this by including an interaction term of *TS*-share and a dummy for the post-2001 version of the law.

Table 6 reports the results for the fixed and random effects (RE) regressions in the university preparatory branch *VWO* when we use the weighted average *CUMI*-variable. The coefficient for minority share is similar in magnitude and significant in both FE and RE. The Hausman test, however, generated a negative test statistic. Sometimes, this is interpreted as being below the critical value and thus allowing for the use of random effects. To be sure, though, we calculated an unweighted moving average of minority shares, assuming constant student population, and ran the FE and RE regressions again. When this approach is employed, the results are not much different from before, the (positive) Hausman test now however clearly allows for the use of random effects.

Given the short time-series and the lack of control variables, one should obviously not overstress the results of the longitudinal analysis. Nevertheless, the results from this subsection suggest that there is a positive within-school influence of minority share on the difference between school exam and central exam grades, supporting the results from the cross-sectional estimation presented in the previous subsections.

6 Conclusion

In this paper, we analysed the impact of social class on the choice of grading standards by schools. We showed in a theoretical model that schools with a disadvantaged student body tend to apply less demanding standards if graduates from such backgrounds face less appealing job market conditions than others. The predictions of the model were then tested on data from the Netherlands since the Dutch educational setup provides the rare opportunity of measuring decentralised grades awarded by the individual schools against the benchmark of central test results. The empirical results show that schools with many students from cultural minorities, or receiving financial aid, award better grades.

This result is strikingly different from the result of the PISA study reported by Prenzel et al. (2005). Whereas this study suggests that students from lower classes get a rough deal from the grading and examination system which holds them back from enjoying more rewarding types of education, our results rather show that such students are held to less demanding standards than students from average backgrounds. While we refrain from drawing any more general conclusions from these results at this stage, it seems safe to point out that, in order to explain the treatment of lower class students by the schooling system, more than a simple appeal to discrimination is needed.

This observation suggests that much more research is required in order to enhance the understanding of how standards are set. For example, it will be fruitful to integrate other motives for the choice of standards. As some of our empirical results suggest, competition for students may be an important driver of grade inflation. This will be analysed theoretically and treated in more detail in future empirical work.

Appendix: Second order condition

Differentiating (8) once more with respect to s_i, we find, with $f = 1/\overline{\gamma}$

$$
\begin{aligned}
\frac{\partial^2 W_i}{\partial s_i^2} &= -f\frac{\gamma_i}{s_i}\frac{\mathrm{d}w_C}{\mathrm{d}s_i} - f w_C \left[\frac{\partial\left(\frac{\gamma_i}{s_i}\right)}{\partial s_i} + \frac{\partial\left(\frac{\gamma_i}{s_i}\right)}{\partial \tilde{w}}\frac{\mathrm{d}w_C}{\mathrm{d}s_i}\right] \\
&\quad -f\left(1-\frac{1}{\eta}\right)\left[\frac{\partial\gamma_i}{\partial s_i} + \frac{\partial\gamma_i}{\partial\tilde{w}}\frac{\mathrm{d}w_C}{\mathrm{d}s_i}\right]\frac{\mathrm{d}w_C}{\mathrm{d}s_i} + \left[1 - f\gamma_i\left(1-\frac{1}{\eta}\right)\right]\frac{\mathrm{d}^2 w_C}{\mathrm{d}s_i^2}\,.
\end{aligned}
$$

From (2) and (3) one has $\partial\gamma_i/\partial s_i = \gamma_i/s_i$, hence $\partial(\gamma_i/s_i)/\partial s_i = 0$, and $\partial\gamma_i/\partial\tilde{w} = -(1/\eta)\gamma_i/w_C$. Inserting this and rearranging, one arrives at

$$
\frac{\partial^2 W_i}{\partial s_i^2} = -f\left(1-\frac{1}{\eta}\right)\frac{\gamma_i}{s_i}\left(2 - \frac{1}{\eta}\frac{s_i}{w_C}\frac{\mathrm{d}w_C}{\mathrm{d}s_i}\right)\frac{\mathrm{d}w_C}{\mathrm{d}s_i} + \left[1 - f\gamma_i\left(1-\frac{1}{\eta}\right)\right]\frac{\mathrm{d}^2 w_C}{\mathrm{d}s_i^2}\,. \tag{A.1}
$$

At a symmetric situation with $w_C = \lambda_C s_C = \lambda_C s_i$, $\gamma_i = \gamma_C$, and $\mathrm{d}w_C/\mathrm{d}s_i = \lambda_C/C$, (A.1) reduces to

$$
\frac{\partial^2 W_i}{\partial s_i^2} = -f\left(1-\frac{1}{\eta}\right)\frac{\lambda_C}{C}\frac{\gamma_C}{s_C}\left(2 - \frac{1}{\eta C}\right) + \left[1 - f\gamma_C\left(1-\frac{1}{\eta}\right)\right]\frac{\mathrm{d}^2 w_C}{\mathrm{d}s_i^2}\,. \tag{A.2}
$$

From $\eta > 1$ and $C \geq 1$, it follows that the first term in (A.2) is strictly negative. Moreover, $\eta > 0$ and $f\gamma_C < 1$ imply $1 - f\gamma_C[1-(1/\eta)] > 0$. Thus, $\mathrm{d}^2 w_C/\mathrm{d}s_i^2 \leq 0$ is sufficient for $\partial^2 W_i/\partial s_i^2 < 0$ at a symmetric situation.

Defining $A(s_i) = f(\lambda_C s_i - w_C)(\partial\gamma_i/\partial s_i) - \lambda_C[1 - F(\gamma_i)]$ and $B(s_i) = \sum_{j\in C}[1 - F(\gamma_j)] + f\sum_{j\in C}(\lambda_C s_j - w_C)(\partial\gamma_j/\partial\tilde{w})$, we can write (6) as $\mathrm{d}w_C/\mathrm{d}s_i = -A(s_i)/B(s_i)$. We have

$$
\begin{aligned}
\frac{\mathrm{d}A}{\mathrm{d}s_i} &= f\left(\lambda_C - \frac{\mathrm{d}w_C}{\mathrm{d}s_i}\right)\frac{\partial\gamma_i}{\partial s_i} \\
&\quad + f(\lambda_C s_i - w_C)\left(\frac{\partial^2\gamma_i}{\partial s_i^2} + \frac{\partial^2\gamma_i}{\partial s_i\partial\tilde{w}}\frac{\mathrm{d}w_C}{\mathrm{d}s_i}\right) + f\lambda_C\left(\frac{\partial\gamma_i}{\partial s_i} + \frac{\partial\gamma_i}{\partial\tilde{w}}\frac{\mathrm{d}w_C}{\mathrm{d}s_i}\right)\,.
\end{aligned} \tag{A.3}
$$

From symmetry, $\lambda_C s_i - w_C = 0$, and hence the second term in (A.3) drops out. Using, in addition, $s_i = s_C, \gamma_i = \gamma_C, \mathrm{d}w_C/\mathrm{d}s_i = \lambda_C/C$, (2), and (3), (A.3) reduces to

$$
\frac{\mathrm{d}A}{\mathrm{d}s_i} = f\lambda_C\frac{\gamma_C}{s_C}\left(2 - \frac{1}{C} - \frac{1}{\eta C}\right)\,. \tag{A.4}
$$

Similarly, one finds

$$
\frac{\mathrm{d}B}{\mathrm{d}s_i} = f\left(\lambda_C\frac{\partial\gamma_i}{\partial\tilde{w}} - \frac{\partial\gamma_i}{\partial s_i}\right) - 2\sum_{j\in C} f\frac{\partial\gamma_j}{\partial\tilde{w}}\frac{\mathrm{d}w_C}{\mathrm{d}s_i} + \sum_{j\in C} f(\lambda_C s_j - w_C)\left(\frac{\partial^2\gamma_j}{\partial\tilde{w}\partial s_i} + \frac{\partial^2\gamma_j}{\partial\tilde{w}^2}\frac{\mathrm{d}w_C}{\mathrm{d}s_i}\right)
$$

which, with symmetric standards and using (2) and (3), reduces to

$$\frac{\mathrm{d}B}{\mathrm{d}s_i} = -f\frac{\gamma_C}{s_C}\left(1 - \frac{1}{\eta}\right). \tag{A.5}$$

Moreover, symmetry yields $A = -\lambda_C(1 - f\gamma_C)$ and $B = C(1 - f\gamma_C)$. Together with (A.4) and (A.5), one so finds

$$\frac{\mathrm{d}^2 w_C}{\mathrm{d}s_i^2} = -\frac{\frac{\mathrm{d}A}{\mathrm{d}s_i}B - \frac{\mathrm{d}B}{\mathrm{d}s_i}A}{B^2} = -\frac{2f\lambda_C}{C(1 - f\gamma_C)}\frac{\gamma_C}{s_C}\left(1 - \frac{1}{C}\right) \leq 0\,,$$

where the sign follows on $C \geq 1$. □

References

Betts, J. (1998a), The impact of educational standards on the level and distribution of earnings, *American Economic Review* 88, 266-275.

Betts, J. (1998b), The two-legged stool: The neglected role of educational standards in improving America's public schools, *Economic Policy Review* 4, 97-116.

Betts, J. and R. Costrell (2001), Incentives and equity under standards-based reform, *Brookings Papers on Education Policy*, 9-74.

Bishop, J. (1997), The effect of national standards and curriculum-based exams on achievement, *American Economic Review* 87, Papers and Proceedings, 260-264.

Bishop, J. (1999), Are national exit examinations important for educational efficiency?, *Swedish Economic Policy Review* 6, 349-398.

Bishop, J., F. Mane, M. Bishop and J. Moriarty (2001), The role of end-of-course exams and minimum competency exams in standards-based reforms, *Brookings Papers on Education Policy*, 267-345.

Centraal Bureau voor de Statistiek, Kerncijfers postcodegebieden 2003.

Centraal Bureau voor de Statistiek, Kerncijfers wijken en buurten 2001-2005.

Chan, William, Hao Li, and Wing Suen (2005), A signalling theory of grade inflation, *mimeo*, University of Hong Kong.

Costrell, R. (1994), A simple model of educational standards, *American Economic Review* 84, 956-971.

Costrell, R. (1997), Can centralized educational standards raise welfare?, *Journal of Public Economics* 65, 271-293.

Epple, D., E. Newlon, and R. Romano (2002), Ability tracking, school competition, and the distribution of educational benefits, *Journal of Public Economics* 83, 1-48.

EPPLE, D. and R. ROMANO (1998), Competition between private and public schools, vouchers, and peer group effects, *American Economic Review* 88, 33-62.

GRAAFF, T. DE, C. GORTER, and P. NIJKAMP, (2001), Effects of ethnic geographical clustering on educational attainment in the Netherlands, *Tinbergen Institute Discussion Papers* 01-028/3, Tinbergen Institute.

HANUSHEK, E. (2002), Publicly provided education, in: A. Auerbach and M. Feldstein (eds.), *Handbook of Public Economics* 4, Amsterdam: Elsevier, 2047-2141.

HSIEH, C.-T. and M. URQUIOLA (2006), The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program, *Journal of Public Economics* 90, 1477-1503.

INSPECTIE VAN HET ONDERWIJS, Kwaliteitskaart Voortgezet Onderwijs, uitgaven 1998-2004, Steinmetz Archief, Amsterdam.

JACOB, B. (2005), Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago public schools, *Journal of Public Economics* 89, 761-796.

JÜRGES, H., W. RICHTER, and K. SCHNEIDER (2005), Teacher quality and incentives: Theoretical and empirical effects of standards on teacher quality, *Finanzarchiv N.F.* 61, 298-326.

JÜRGES, H., K. SCHNEIDER, and F. BÜCHEL (2005), The effect of central exit examinations on student achievement: Quasi-experimental evidence from TIMSS Germany, *Journal of the European Economic Association* 3, 1134-1155.

KARSTEN, S., G. LEDOUX, W. MEIJNEN, J. ROELEVELD and E. VAN SCHOOTEN (2006), Choosing segregation or integration?: The extent and effects of ethnic segregation in Dutch cities, *Education and Urban Society* 38(2), 228-247.

MINISTERIE VAN ONDERWIJS CULTUUR EN WETENSCHAPPEN, Onderwijs in Cijfers 2003.

NECHYBA, T. (1999), School finance induced migration and stratification patterns: The impact of private school vouchers, *Journal of Public Economic Theory* 1, 5-50.

PRENZEL, M., J. BAUMERT, W. BLUM, R. LEHMANN, D. LEUTNER, M. NEUBRAND, R. PEKRUN, H.-G. ROLFF, J. ROST, and U. SCHIEFELE (eds.) (2005), *PISA 2003: Ergebnisse des zweiten internationalen Vergleichs – Zusammenfassung*, IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften an der Universität Kiel, available under http://pisa.ipn.uni-kiel.de/Ergebnisse_PISA_2003.pdf.

SOCIAAL EN CULTUREEL PLANBUREAU, Statusscores postcodegebieden 2002.

WIKSTRÖM, C. and M. WIKSTRÖM (2005), Grade inflation and school competition: An empirical analysis based on the Swedish upper secondary schools, *Economics of Education Review* 24, 309-322.

WÖSSMANN, L. (2003), Schooling resources, educational institutions and student performance, *Oxford Bulletin of Economics and Statistics* 65, 117-170.