EVERYDAY DECISION MAKING: A THEORETICAL AND EMPIRICAL STUDY.

Dissertation
zur Erlangung des wirtschaftswissenschaftlichen Doktorgrades der
Wirtschaftswissenschaftlichen Fakultät der Universität Göttingen

vorgelegt von
Kamila Danilowicz-Gösele
aus Kattowitz

Göttingen, 2016

Erstgutachter: Prof. Dr. Robert Schwager
Zweitgutachter: Prof. Dr. Thushyanthan Baskaran
Drittprüfer: Prof. Dr. Holger Strulik
Tag der mündlichen Prüfung: 19. Dezember 2016

Acknowledgements

First of all, I would like to thank my advisor, Robert Schwager, for his trust in my skills and the opportunity to work on something I really wanted. His insightful ideas and knowledge have always been a source of inspiration for me and it would not be an exaggeration to say that without him my academic life as a doctoral student would not have even begun. In addition, I would like to thank my second supervisor, Thushyanthan Baskaran. I appreciated his immediate replies with valuable comments and useful advices very much. I would also like to express my gratitude to Holger Strulik for accepting to be a member of my examination committee so promptly. Moreover, I gratefully acknowledge the financial support from the German Federal Ministry of Education and Research under grant number 01PW11004.

My colleagues in Göttingen have been exceptional. I am particularly grateful to my co-authors Katharina Lerche, Johannes Meya and Robert Schwager for a very fruitful and enjoyable collaboration and many great discussions. Futhermore, I would like to thank the current and former members at the Chair of Public Economics for the wonderful atmosphere with inspiring coffee breaks in between and everything else that made the days go by so fast.

My heartfelt appreciation goes to my family for giving me strength to reach for the stars and chase my dreams. First and foremost my dear parents, Teresa and Wojciech, for all they have done for me to give me the best life they could possibly give. I am deeply grateful for your endless love and for letting me pursue my dreams so far from home. Finally, I would like to thank my best friend and husband David and my beloved daughters Klara and Lila for their love, encouragement, incredible patience and just being there. I love you.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Every day, we are challenged with decisions, big and small, rational and irrational, that are made in our personal lives, workplaces, schools and beyond. Some of them are strong and self-confident, others need support and guidance. Some are perceived to be more pleasant in the very moment, but might not consider our best interest in the future. All of them, however, have one thing in common: the impact they create, short or long-term, individual or collective. This thesis analyzes the individuals' decision-making and the role of institutions in creating incentives and influencing individuals' choices.

Although the decision-making happens on a daily basis, it still represents one of the most complex cognitive processes of human performance. It involves more than just synthesizing and applying all the information and experiences that individuals might have accumulated throughout their life span. Individual characteristics and preferences, as well as the environment, in which the decision maker is behaving, also have a considerable impact on the process itself and the corresponding outcome.

The research on decision-making is, in fact, an integral part of many fields of study and thus methodologically different. However, in all of them the research was hampered for years by the standard assumptions of rationality, self-interest and perfect information. It took quite a while to pick up on this issue and empirically show that these

assumptions still hold in some situations, but not in general. Especially decisions with long-term consequences are hard to make and are often based only on short-time considerations. Also behavioral biases such as self-control problems or inertia often lead to bounded rationality where individuals fail to act in their own best interests, thus opening up space for interventions and strengthening the role of public policy. Hence, on the one hand, individuals influence public policy and institutions through their preferences and choices. On the other hand, public policy and institutions provide a framework of *rights* and *wrongs* within which individuals act and interact with each other, shape their expectations and limit their choices. So above all, this thesis is about incentives and the role of institutions that create them.

This thesis is divided into two major parts. The first part is descriptive and theoretical and deeply rooted in behavioral public economics. It aims at understanding how people make decisions and the factors that affect their decision-making process. The second part provides an empirical analysis of higher education and is at the heart of research in economics of education. It focuses on the interaction between individuals' characteristics and choices that determine their performance within a given institutional framework, namely university. Both parts are relatively independent and consider individuals in different roles and situations. The following paragraphs provide a brief summary of the remaining chapters contributing to this thesis.

Chapter 2 consists of a short essay that provides a critical discussion on the validity of the neoclassical rationality paradigm as a normative standard. Furthermore, it addresses the alternative, behavioral approach to bounded irrationality and discusses the need and importance of paternalistic interventions.

Chapter 3 examines individuals' behavior regarding healthy good consumption by conducting a theoretical analysis of government's intervention in the decision-making of individuals with self-control problems. It investigates the effectiveness of a particular paternalistic intervention, namely subsidies for health care, aimed at changing consumption behavior of present-biased individuals, who underestimate the effect of today's consumption on future health. Such preference for immediate gratification allows paternalistic policymakers to indirectly regulate behavior by exploiting the time-inconsistency of individuals.

In order to analyze this effect, immediate subsidies paid for health-conscious consumption are compared to the future subsidies rewarding a good health outcome. The main finding of this chapter is that, while both policies can implement the first-best choice, doing so by future subsidies results in higher costs for the government. This arises since the individual anticipates that, from today's perspective, she will make biased use of future subsidies. Hence, in order to create the same incentive effect, a future subsidy must be higher in present value terms.

Chapter 4 presents an econometric analysis of the determinants of students' academic performance at university. It is based on a unique administrative data set on more than 12,000 student careers. The obtained microanalysis confirms previous findings from the literature on higher education, that the high school grade is a strong predictor of the students' academic success, measured both as the graduation probability and as the final university grade. The socio-economic variables have, in contrast, a relatively small impact.

When looking at the different faculties separately, the link between high school performance and university success is shown to vary substantially, suggesting differences in the teaching and examination cultures. At some faculties, for instance, social sciences or humanities, the probability of graduating is rather low, while grades are quite good conditional on high school performance. At others, like economic sciences or forest sciences, weaker students have a greater chance of graduating, but grades are more differentiated. However, at faculties like mathematics and physics, weaker students have neither a good chance of graduating nor of obtaining a good grade.

Chapter 5 analyzes professors' effect from a fundamental first-year course in Economics on students' later performance in follow-on courses. It focuses on the differences in grading policies between five different professors assigned to the same mandatory course. The econometric analysis is based on an extensive dataset consisting of administrative data on more than 2,900 students from the University of Göttingen. By applying an instrumental variable (IV) strategy, special attention is given to the problem of self-selection bias of students toward certain professors.

The obtained results indicate that professors have powerful effects on students' achievement: Besides huge differences in grading policies, we find a significant effect of

having a certain professor in the mandatory first-year course on students' later performance. However, the sign of this effect is ambiguous, and depends on the mathematical rigor of the course and the examination style.[1]

When talking about decisions, ultimately, it is up to the reader to decide whether or not this thesis is valuable and thus worth reading. For me it was definitely worth writing.

[1]The summaries of Chapters 3-5 heavily borrow from the abstracts of the discussion paper versions of these chapters Danilowicz-Gösele and Schwager (2016), Danilowicz-Gösele et al. (2014) and Danilowicz-Gösele (2016) respectively.

Chapter 2

Bounded Rationality

> "Bounded rationality is not irrationality. ... On the contrary, I think there is plenty of evidence that people are generally quite rational; that is, they usually have reasons for what they do."
>
> — (Simon, 1985)

> "Dividing our field into *behavioral* and *neoclassical* economics is akin to distinguishing *time separable* economists from others."
>
> — (Chetty, 2015)

The standard economic theory of consumer behavior builds on the assumption that individuals act consistently and rationally so as to maximize a unique, well-defined, objective function. Over the last decades, the validity of this view has been strongly challenged by the empirical evidence from laboratory experiments in social sciences. Hence, people often make decisions that are not in their best interest, ranging from postponing the unpleasant to chasing the immediate satisfaction of wants. Even economists increasingly raise doubts about the rationality of the individual actions, supporting the need to provide a more suitable framework for capturing individuals' systematic deviations from the standard economics rules of decision-making. Bounds on rationality seem to be important often enough to include them in economic analysis, and that is how the advocates of paternalism come into play.

Indeed, many behavioral researchers have already abandoned the notion of individuals as fully rational and suggested a more complex picture in which individuals do not

exhibit rational expectations but instead a dynamic inconsistency (Thaler, 1981), use heuristics (Gabaix et al., 2006) and are subject to preference reversals due to framing effects (Kahneman and Tversky, 1979).

Although the research field of behavioral economics has grown rapidly, its validity as an alternative approach to the neoclassical model is still a matter of debate (List, 2004; Levitt and List, 2007; DellaVigna, 2009). The main questions raised are, on the one side, whether individuals are indeed plagued by systematic biases and errors, and on the other side, whether the authorities have the right to steer individuals' choices in the right direction.

Moreover, there is no agreement among behavioral researchers on the criteria for what would count as true preferences and how paternalistic decision-makers know what is best for us. Some researchers also question putting the blame for societal problems solely on the individuals' mind, thus ignoring the effect of industries nudging unhealthy behavior. A further important question is whether individuals can learn to deal with risk and uncertainty, such that educating people can be considered a more sustainable alternative to nudging.

As one can notice from the above, the existing theoretical literature provides neither a clear measure of irrationality nor formulates common criteria for the justification of paternalistic interventions. Instead, it presents multiple approaches often with widely convoluted and self-contradictory arguments. This paper aims at addressing these and related concerns by discussing different arguments and different lines of evidence in the paternalism-debate presented by some behavioral researchers.

We begin with probably the most famous anomaly hunter among economists - Richard Thaler, who unveiled his idea of paternalism in the paper co-authored with Cass Sunstein (Conlisk, 1996). Thaler and Sunstein (2003) argue in favor of libertarian paternalism and believe that the anti-paternalistic stance of many economists is based on false assumption and some misconceptions. The false assumption is that individuals know their best interests better than the third parties that try to impose paternalism on them. The two misconceptions mentioned in their paper relate to the claims that, firstly, paternalism always involves coercion, and, secondly, that there are viable alternatives to paternalism. The authors classify a policy as paternalistic if it

aims at influencing individuals' choices in order to make the individuals better off, with better off as an objective measure and not necessarily an equal sign between revealed preferences and welfare. For this reason, researchers should concentrate on identifying the most effective paternalistic options, rather than on raising the question of whether paternalistic policies should be implemented or not.

Assuming such a self-evident role and function of paternalistic policy-makers does not fully convince me. Let me make my concerns more concrete. The presumption that paternalism should interfere with individuals' freedom is usually based on the assumption that individuals do not know what is best for themselves. Hence, paternalism, no matter in which form, requires benevolent experts who are time-consistent and know what is best for us in order to be able to improve our decision making. Nevertheless, this requirement appears to be self-contradictory: On the one hand, the benevolent experts suffer from the same cognitive biases as other individuals; on the other hand, they are supposed to steer people towards more preferable behaviors. Policy-makers like all individuals often make the best decision given the situation. Therefore, unsurprisingly, opponents of paternalism commonly argue that paternalistic interventions may cause harm by interfering with individuals' autonomy by producing consequences that are not always in individuals' best interest.

Thaler and Sunstein (2003) give many examples of paternalistic interventions in their paper. One of them is the automatic enrollment in 401(k) employee saving plans. In most of the existing saving plans, employees have to actively opt-in to join the plan, such that the default is *no enrollment*. However, in the preferable case of automatic enrollment, employees are enrolled in the plan unless they actively opt-out - a paternalistic intervention introduced to promote employees' welfare. Another possible alternative for the enrollment decision would be to require the employees to make a choice, either to opt-in or opt-out.

Even though this alternative is not seriously considered by the authors, since employees are prone to procrastination and plagued by inertia and therefore should not be forced to make a choice, I cannot fully agree with this argument. The adoption of automatic enrollment should be considered paternalistic, because "most employees would prefer to join the saving plan if they took time to think about it and did not lose

the enrollment form". If this is true, why should then someone take the active choice away from the employees? If refraining from taking any action would not be an option, would the employees, according to the above argument, make a worse choice than in the paternalistic case?

Furthermore, I am not convinced with the argument, that in the case of automatic enrollment, since no one is forced to do anything, such intervention should not be considered unobjectionable. If an individual is automatically enrolled in a saving plan, she is committed to give up a certain amount of money, unless she opts-out. As a result of such commitment, an employee might be then forced to rethink other choices in her life, because the money left does not allow her to maintain the previous standard of living. Is this kind of forcing less problematic?

According to the authors' guess in the case of automatic enrollment, most employees stay in the plan eventually, and only very few opt out. Thus, cost-benefit analysis would appear to favor the selection of this default option, since the cost of insufficient savings for retirement should be greater than the cost of excessive savings. As I mentioned, however, the paternalism should make some people better off, without making others worse off. What if someone is willing to opt-out but simply forgets to do so? Would this be an example of a parallel to losing the enrollment form? Why are the costs of someone who forgot to opt-in higher than of someone who forgot to opt-out? Therefore, I am not convinced that paternalism becomes, not only permissible, but also generally unavoidable to correct for the status quo bias of individuals.

Gigerenzer (2015) also partly disagrees with Thaler's and Sunstein's attitude toward the justification of paternalism. He concludes in his paper that the claim that individuals are hardly educable lacks evidence and that teaching individuals to become more aware of their choices offers a true alternative to nudging. In his opinion libertarian paternalism forgoes using incentives and enforcing behavior in favor of tricks called *nudges* to steer people's choices. Hence, *libertarian* stands for freedom of choices and *paternalism* for the paternal protection from ourselves, from our systematic biases and errors. Gigerenzer (2015) does not criticize *nudging* per se, but the justification of *nudges* on the basis of a latent irrationality by libertarian paternalists such as Thaler and Sunstein. His criticism is twofold: Firstly, libertarian paternalism should not focus

on the individuals' mind as the real cause of societal problems and use *nudging* as an excuse for not protecting consumers. Secondly, the suggestion that releasing individuals from their biases through education is nearly impossible, is a bias itself and therefore should be reconsidered in light of experimental evidence that even children can become risk savvy and that less-educated people tend to have worse health outcomes. Libertarian paternalism turns out to be even more severe than some forms of hard paternalism, since its justification of governmental interventions is based not only on struggles with achieving goals but also on problems with setting goals worth to achieve.

I certainly agree with the author, that educating people offers a more sustainable solution than *nudging*. At the same time I can think of many reasons for politicians to prefer *nudging* over educating people. *Nudging* comes with governmental and industrial profits and it is a cost-saving alternative to educating people. *Nudging* does not aim at understanding but very often at achieving a certain goal such as increasing participation rates or changing lifestyles. Last but not least, low-educated people are more obedient, do not question things and are easier to manipulate than educated individuals. *Nudging* individuals into proper behaviors has a chance to succeed only if, at the same time, industries stop nudging people into the opposite direction, namely unhealthy behaviors. Nonetheless, educating people to risk-savvy consumers is still the most promising solution.

As already noticed by Simon (1985), the fact that people are bounded when making decisions, however, does not mean that they are irrational, and I strongly endorse this view. Bounded rationality, unlike irrationality, is not based on inconsistent preferences but rather on a goal-oriented behavior, which implies the presence of intention. The intended rational behavior, that is making decisions with the intent of reaching some short-term other long-term goals, is, however, constrained by the uncertainty of environmental courses of action and by the limited computational and information-processing capacities of human beings, and thus inherently bounded. According to Simon (1947), individuals reduce the complexity of the decision-making process to a simplified individual model that consists only of a limited number of factors and alternatives. Subsequently they choose the most satisfactory rather than the most optimal option, where *satisfactory* itself may refer to reasonable, adequate or acceptable. In other words,

individuals make *good enough* decisions given the situation. Because the first satisfactory alternative is selected, the decision-making process is strongly driven by the order in which alternatives are presented and by the effort of the involved decision-maker. Therefore, Simon (1947) labeled such decision-making behavior as *satisficing*.

Regarding the future of behavioral economics, Conlisk (1996) and Chetty (2015) suggested that the research on bounded rationality should not be seen as a "departure from economic reasoning, but a needed extension of it", which I more than agree with. It neither should be considered as a separate sub-field and a challenge to neoclassical economic model but rather as a natural progression of it. Therefore, instead of discussing the validity of the traditional rationality assumption, we should incorporate behavioral factors into regular policy by: (1) using new policy tools in order to steer individuals' behavior in a certain direction, (2) improving empirical predictions about the effects of existing policies on individual's behavior, (3) accounting for the impact of differences between decision and experienced utilities and related new welfare implications. Such a pragmatic rather than philosophical approach to behavioral economics may be more useful than framing the debate around the assumptions about the rationality/irrationality of individuals. A universal assumption of appropriate rationality is not something the researchers should try to agree on.

Chapter 3

Subsidizing Health-Conscious Behavior Now or Later*

3.1 Introduction

Life consists of three cycles – past, present and future. We regret our past, fully appreciate the present pleasures and very often refuse to move forward. Sometimes there is not even an inch of movement, and although we consistently plan to direct our behavior towards the future, currently the most immediate moment still gets the highest weight. This so-called present-bias is reflected in the popular saying – *Eat, drink and be merry, for tomorrow we diet.* According to that saying, people either underestimate the effect of today's consumption on future health or postpone health investments to a later date, since unpleasant activities seem to be even more unpleasant the closer they are to the present. As a consequence, a well-meaning paternalistic government may intervene to counterbalance the intertemporal distortion of consumption toward the present and hence improve the health status.

In this paper we analyze a specific instrument for this intervention, health-related subsidies. These can be designed in two different ways, depending on the timing and the target of the subsidy. They can either immediately reward an individual's health-conscious consumption, or reward the individual's health outcome in the future. We show that this distinction is relevant both for the determination of the optimal subsidy rate and for the subsidy's effectiveness, measured by the tax revenues required to overcome the present-bias. This is due to the fact that an immediate and a future subsidy are paid to different 'selves' of the individual.

Apart from subsidies, taxes on unhealthy behavior, so-called 'sin taxes', come to

*This chapter originates from joint work with Robert Schwager (see Danilowicz-Gösele and Schwager, 2016).

mind as an alternative instrument to overcome present-biased behavior. Indeed, from a public finance point of view, taxes may appear as the superior instrument since they have the obvious advantage of raising revenues. Consequently, research on paternalistic interventions has so far mostly concentrated on taxes. Nevertheless, subsidies are also widespread, as the examples below illustrate. The reason for this fact may be that rewarding (with subsidies) has more positive connotation than punishing (with taxes): giving instead of taking away. In particular, a government which would create incentives for healthy consumption by taxing the sick would be considered nasty and inhuman, even if the individuals themselves have triggered the illness by their own behavior.

For these reasons, we study health-related subsidies. We begin, in Section 3.2, with some examples of immediate and future health subsidies, followed by a brief review of the literature related to paternalistic policies. Our analysis is based on a model, presented in Section 3.3, of an infinite-horizon consumer choice problem with two goods per period. In addition to a numeraire good, there is a healthy good that creates positive health consequences in the future in addition to current utility from consumption. Present-bias is modeled in the form of quasi-hyperbolic discounting as in Harris and Laibson (2001).

We consider a government which treats such short-time desire as an error and hence intervenes to correct the individual's choices. In Section 3.4, we analyze an immediate subsidy paid for health-conscious consumption in the same period and a future subsidy rewarding a good health outcome one period later. For both forms of subsidy, we characterize the rate which, despite present-biased preferences, induces the consumption pattern which an unbiased individual would choose. While the optimal rate of immediate subsidy simply bridges the gap between the biased and the unbiased evaluation of health benefits, the future subsidy must take into account two behavioral responses of the individual which are specific to present-biased preferences. On the one hand, future transfer income, just like future health, is valued less by the individual – the *discounting effect* of the future subsidy. On the other hand, the individual can change the behavior of her future self by increasing future income – the *instrumental effect* of the future subsidy.

In a next step, in Section 3.5, we investigate how the balance of these effects de-

termines the present value of taxes required to finance the optimal subsidy. From the point of view of a paternalistic decision maker, who discounts future payments in an unbiased fashion, the immediate subsidy entails lower cost since the discounting effect dominates. This result suggests that policies of the second kind, where the reward is delayed, are less effective in the presence of present-biased preferences than policies of the first kind, where health conscious behavior is subsidized immediately.

Finally, in Section 3.6, we abandon the assumption of a paternalistic government, and instead consider a government which represents the individual. This government follows the same present-biased preferences as the consumer, but has the same fiscal instruments at its disposal as the paternalistic government. We show that such a government implements the optimal subsidy only if it puts sufficiently high weight on future payoffs. Otherwise, the gain from committing future selves to increased healthy consumption does not outweigh the immediate loss from forcing an unwanted change in behavior upon one's own current self. Furthermore, since a present-biased government discounts future tax payments more heavily, the cost of future subsidies counts for less in its evaluation of present values than in the computation of a paternalistic government. Therefore, in contrast to the paternalistic government, the government which represents the consumer will in some cases favor the future over the immediate subsidy scheme. We conclude in Section 3.7 by summarizing our findings and suggesting possible extensions of our analysis.

3.2 Examples and Literature

An example for an immediate subsidy is provided by a new nutrition program, Healthy Incentives Pilot (HIP), developed by the United States Department of Agriculture, which has been tested for 12 months in Hampden County (Massachusetts).[3] HIP is an incentive-based program to empower low-income people, Supplemental Nutrition Assistance Program (SNAP) recipients, to consume more fruits and vegetables. For every dollar spent with the SNAP Electronic Benefit Transfer cards on fruit and vegetables, participants earned an incentive of 30 cents. The incentive was immediately added to their SNAP account, thus cutting the costs for fruit and vegetables by almost

[3]See Bartlett et al. (2014).

one third. According to the final report, HIP participants increased their consumption of targeted fruits and vegetables by 26%, driving the Healthy Eating Index by 5%. The self-reported total spending on fruits and vegetables of HIP households was 8.5% higher than spending reported by non-HIP households, and stayed stable across the pilot period.

Another example is the School Fruit Scheme (SFS), an EU-wide voluntary program designated to encourage young people to consume more fruit and vegetables.[4] According to the final report, in 2010/11 SFS reached 8,146,290 children in 54,267 schools (European Commission, 2012). The program thus proved successful in increasing the fruit and vegetables consumption of children in the short-run. Further examples for immediate subsidies are financial rewards for abandoning unhealthy behaviors (Volpp et al., 2008, 2009) and pricing and promotion strategies related to healthy items (French et al., 2001).

There are also examples for programs where the reward is given after a certain period of time. In the United States, many companies offer *Healthy Rewards Cards* and other rebate programs as an incentive to promote healthy lifestyles. Participants of health reward programs earn points and gift cards for engaging in healthy behaviors such as quitting tobacco (IBM: "Healthy Living Rebate" program), participating in disease management programs (FedEx: diabetes-management program; IBM: childhood obesity management program "Children's Health Rebate"), taking the general health assessments or exercising regularly (Scotts Miracle-Gro: "Health Quotient" and "Wellness Center").[5]

Also in Germany future subsidies in form of bonus programs offered by statutory health insurance companies are a part of modern life. Finally, an example for a subsidy on future health outcomes is provided by the widespread use, in health insurance contracts, of rebates granted to clients who do not claim any expenses during some period. These examples show that health-related subsidies are a widely used and effective policy instrument.

In public finance research, present-biased preferences and paternalistic policies to

[4]Participating Member States are in addition required to implement strategies including educational and awareness-raising initiatives.

[5]See Business Roundtable (2007).

improve individual decision making have received increasing attention. O'Donoghue and Rabin (2003, 2006) study optimal sin taxes in an economy with heterogeneity in preferences for the sin good and in the degree of time-inconsistency. They show that when there is some degree of self-control problems in the population such taxes not only counteract overconsumption by consumers with self-control problems, but can even create Pareto improvements.

A similar problem has been addressed by Aronsson and Thunström (2008), who consider the policy implications related to unhealthy food consumption in an economy with time-inconsistent individuals. Their result shows that a combination of subsidies for wealth and health capital makes the individual choose the same resource allocation as the social planner, thus internalizing the externality imposed by the individual's current self on her future selves.

Cremer et al. (2012) study the interaction between sin goods and health spending within two settings. In the first one, an individual is subject to persistent error and hence continues making biased choices. In the second one, an individual is modeled as a dual self, meaning that she later acknowledges the mistakes made by her previous biased self. Cremer et al. (2012) show that the first-best optimum can be decentralized by individualized taxes and subsidies. In the first setting, sin goods should be taxed and health-care expenditures subsidized. In the second, there is no need for subsidizing health-care expenditures, but a subsidy on saving is desirable.

Self-control problems related to smoking have been studied by Gruber and Kőszegi (2001). These authors extend the model of Becker and Murphy (1988), where consumers are "rational addicts", by introducing time-inconsistent preferences. According to their results, individuals decrease their cigarette consumption already when future increases in tobacco taxes have been legislatively enacted but are not yet effective. This strongly refers to forward-looking behavior in consumption decisions. Moreover, the authors find that in the presence of time-inconsistency the optimal tax on cigarettes should not only depend on externalities, but also on the internal costs of smoking. Thus, excise taxes on cigarettes have a self-control function that is of high value to smokers who suffer from lack of commitment. In Gruber and Kőszegi (2004), the authors introduce a self-control adjustment to standard tax incidence measures. This firstly lowers the

overall incidence of tobacco taxes and secondly reduces their regressivity.

A systematic review from recent U.S. studies on the effectiveness of food and beverage taxes and subsides in improving public health was done by Powell et al. (2013). These studies analyze the relationship between prices/taxes on the demand for sugar-sweetened beverages, fast-food, fruits and vegetables and on body weight outcomes. Soda taxes imposed on sugar-sweetened beverages, with mean price elasticity of -1.21, do not seem to have much impact on weight. On the contrary, reducing prices for fruits and vegetables, with price elasticity of demand of about -0.5, was found to be associated with lower weight outcomes. As the authors emphasize, this shows the effectiveness of subsidizing fruits and vegetables consumption. Hence, this review proves that relative price changes induced by taxes and subsidies have a significant impact on weight outcomes through consumption patterns.

However, not only governments, researchers and companies pay more attention to health outcomes and health-related interventions. Also the general public would support reducing health insurance taxes for individuals with healthy habits such as exercise and abstention from smoking, as shown by a cross-sectional telephone survey in Israel done by Brezis and Marans (2010). Support for a policy of differential taxation according to lifestyle was high across all sectors of society, even among smokers.

From all this we may conclude that, firstly, paternalistic instruments to promote long-term changes in an individual's behavior are welcomed by different social groups and, secondly, subsidies related to health-conscious behavior can improve health outcomes. By studying such subsidies, our paper adds to previous research which has put more emphasis on taxes. In particular, to the best of our knowledge, the timing of subsidies which is at the heart of our approach has not yet been analyzed.

We now turn to presenting the model which we use to address this issue.

3.3 The Model

An individual consists of a sequence of autonomous temporal selves, which are indexed by the corresponding periods, $t = 0, 1, 2, \ldots$. In each period t, a self with the exogenous and constant per-period income y and the cash-on-hand x_t, which may differ from income because of taxes and subsidies, consumes a healthy good c_t and a numeraire

good d_t. We assume throughout that individuals cannot borrow or save.[6] We will restrict our attention to steady state equilibria where choices and state variables are constant over time.

Normalizing current commodity prices to 1, in the absence of government intervention, self t's budget constraint is given by $x_t = y = c_t + d_t$. Her instantaneous utility in period t is

$$u_t \equiv w(c_t) + v(d_t) + h_t, \tag{3.1}$$

where $w(c_t)$ and $v(d_t)$ denote a self's period-t utility from consumption of healthy good and the numeraire, respectively. The function $h_t = h(c_{t-1})$ represents the positive health consequences from past healthy good consumption, with $h_0 \geq 0$ as the individual's initial (previous) health status. For this function, we use the normalization $h(0) = 0$.

We assume that $v'(d_t) > 0$, $w'(c_t) > 0$ and $w''(c_t) \leq 0$, $v''(d_t) \leq 0$, so that there are positive and weakly decreasing marginal benefits of consumption. Similarly, consumption of the healthy good has positive but non-increasing marginal benefit for health, $h'(c_{t-1}) > 0$ and $h''(c_{t-1}) \leq 0$. To rule out corner solutions, we assume that at least one of the two second derivatives w'' or v'' is strictly negative. Moreover, we impose $w'(0) + \delta h'(0) > v'(y)$ and $w'(y) + \delta h'(y) < v'(0)$, where $\delta \in (0, 1)$ represents time-consistent discounting.

Following Laibson (1994, 1997) and O'Donoghue and Rabin (1999), we adopt (β, δ)-preferences in our model. A self's intertemporal preferences at time t are thus given by

$$U_t = u_t + \beta \sum_{i=1}^{\infty} \delta^i u_{t+i}, \tag{3.2}$$

where $\beta \in (0, 1]$ measures the bias for the present. Within this preference structure we can distinguish two cases: for $\beta = 1$, the preferences are time-consistent and reduced to exponential discounting, and for $\beta < 1$, the preferences are present-biased and the discount rates decline over time.

[6] This paper does not aim at analyzing the interaction of savings with diverse paternalistic policies. Several papers have already demonstrated that hyperbolic consumers save less than exponential consumers (Laibson, 1997, 1998; Angeletos et al., 2001; Diamond and Kőszegi, 2003). This result is also expected in our framework.

Since a time-inconsistent individual consists of multiple selves, she is not able to commit to a particular future consumption behavior. Every self has a tendency to pursue immediate gratification in a way that their future selves do not appreciate. She will therefore choose a consumption level (c_t, d_t) that maximizes her current utility u_t plus a biased version of future utilities, as in (3.2), and not the individual's long-run utility as expressed by U when $\beta = 1$.

We begin our analysis by solving, as a reference case, the individual's optimization for $\beta = 1$, where there is no need for government intervention. In this case, in each period t, a self with cash-on-hand x_t chooses consumption c_t and d_t. In the absence of taxes or subsidies, the first state variable x_t is given exogenously and evolves according to the equation $x_{t+1} = y$. The second state variable h_t is influenced by the past healthy good consumption and evolves according to the equation $h_{t+1} = h(c_t)$.

The unbiased choice is derived from the value function

$$V(h_t, x_t) = \max_{c_t, d_t} \{w(c_t) + v(d_t) + h_t + \delta V(h_{t+1}, x_{t+1}) | x_t - c_t - d_t = 0\}. \tag{3.3}$$

Denoting optimal choices as functions of cash-in-hand by $c(x_t)$ and $d(x_t)$, one finds after inserting $h_{t+1} = h(c_t)$ the first-order conditions

$$w'(c(x_t)) + \delta \left[\frac{\partial V(h_{t+1}, x_{t+1})}{\partial h_{t+1}} \cdot h'(c(x_t)) \right] = v'(d(x_t)) = \lambda_t, \tag{3.4}$$

where λ_t is the Lagrangian multiplier associated with the budget constraint. From the envelope theorem, one has $\partial V(h_{t+1}, x_{t+1})/\partial h_{t+1} = 1$. Using this in (3.4) shows that, for any time-constant exogenous per-period income $x_t = y$, the optimal solution of the above optimization problem $c^* = c^*(y)$ and $d^* = d^*(y)$ is stationary and given by the simultaneous solution to the budget constraint $c^* + d^* = y$ and

$$w'(c^*) + \delta h'(c^*) = v'(d^*)\,. \tag{3.5}$$

Notice that from the assumptions on w'', v'' and h'' and the boundary assumptions on $w' + \delta h'$ and v', this solution is unique and satisfies $0 < c^*(y), d^*(y) < y$. Intuitively, in the absence of present-biased preferences, marginal utilities of the numeraire and the

healthy good are equalized, with the latter consisting of the immediate marginal benefit of consumption and of the delayed marginal impact on health.

3.4 Corrective Policy

We now consider the case where the individual has present-biased preferences. In the following Subsections 3.4.1 and 3.4.2, we analyze two measures to counterbalance the intertemporal distortion of consumption toward the present: an immediate subsidy related to health investment, and a future subsidy related to health outcome. Subsection 3.4.3 illustrates the results by means of two examples.

3.4.1 Immediate Subsidy

Suppose that the government introduces a per unit subsidy $\tilde{z}_t$ on the individual's healthy good consumption c_t. To finance the subsidy payments, the government imposes a lump-sum tax $\tilde{\tau}_t$. Thus, cash-in-hand is $x_t = y - \tilde{\tau}_t$, and the price for the healthy commodity is reduced to $1 - \tilde{z}_t$. The binding budget constraint is then given by $x_t = (1 - \tilde{z}_t)c_t + d_t$ or equivalently by $y = c_t + d_t + \tilde{\tau}_t - \tilde{z}_t c_t$. The state variable x_t evolves according to the equation $x_{t+1} = y - \tilde{\tau}_{t+1}$. The state variable h_t depends on the past healthy good consumption, so that it evolves according to $h_{t+1} = h(c_t)$. This is similar to the reference case, with the difference that consumption c_t is now influenced by the subsidy $\tilde{z}_t$.

At time t, the current self t uses the discount factor $\beta\delta$ and her current-value function can be written as

$$W(h_t, x_t) = \max_{c_t, d_t} \left\{ w(c_t) + v(d_t) + h_t + \beta\delta V(h_{t+1}, x_{t+1}) | x_t - (1 - \tilde{z}_t)c_t - d_t = 0 \right\}.$$

Denoting optimal choices again by $c(x_t)$ and $d(x_t)$, the first-order condition is

$$\frac{w'(c(x_t)) + \beta\delta \left[\frac{\partial V(h_{t+1}, x_{t+1})}{\partial h_{t+1}} \cdot h'(c(x_t)) \right]}{(1 - \tilde{z}_t)} = v'(d(x_t)). \tag{3.6}$$

Solving for $\tilde{z}_t$ yields

$$\tilde{z}_t = 1 - \frac{w'(c(x_t)) + \beta\delta\left[\frac{\partial V(h_{t+1},x_{t+1})}{\partial h_{t+1}} \cdot h'(c(x_t))\right]}{v'(d(x_t))}.$$

Using $\partial V(h_{t+1}, x_{t+1})/\partial h_{t+1} = 1$, substituting $c^* = c^*(y)$ for $c(x_t)$ and $d^* = d^*(y)$ for $d(x_t)$ and rearranging, we get

$$\tilde{z}_t = \tilde{z} = \frac{v'(d^*) - w'(c^*) - \beta\delta h'(c^*)}{v'(d^*)}.$$

With this subsidy rate, the unbiased choices satisfy the first-order condition (3.6) of every self t. When in addition, a lump sum tax is levied which covers subsidy payments, these choices also satisfy the budget constraint in every period. Par Rewriting the subsidy rate $\tilde{z}$ with the help of (3.5), we summarize this result in the following proposition.

Proposition 1. *The government can induce first-best consumption $c_t = c^*(y)$ and $d_t = d^*(y)$ in all periods $t = 0, 1, 2, ...$ by granting an immediate subsidy*

$$\tilde{z} = \frac{\delta(1-\beta)h'(c^*)}{v'(d^*)} \tag{3.7}$$

per unit of healthy consumption financed by lump-sum taxes $\tilde{\tau}_t = \tilde{\tau} = \tilde{z}c^*(y)$ in all periods $t = 0, 1, 2, ...$

For $\beta = 1$, the numerator in (3.7) is zero so that there is no need for subsidizing health-conscious consumption. For $\beta < 1$, the numerator gives the present value of the undervaluation of the marginal health benefit. The denominator is the marginal utility of income, so that the fraction $\tilde{z}$ describes by how much the marginal willingness to pay for the healthy good differs between the unbiased and the biased consumer. Thus, the optimal immediate subsidy $\tilde{z}$ balances the wedge between the biased and unbiased evaluation of health.

3.4.2 Future Subsidy

Now we examine the second form of subsidies, namely, future subsidies rewarding good health outcomes. Suppose, therefore, that the government grants a subsidy z_t on h_t and imposes a lump-sum tax τ_t in order to finance the subsidy payments. Cash in hand is then given by $x_t = y - \tau_t + z_t h_t$, and the binding budget constraint is given by $x_t = c_t + d_t$ or $y = c_t + d_t + \tau_t - z_t h_t$. In the case of future subsidies the state variable x_t evolves according to the equation of motion $x_{t+1} = y + z_{t+1} h_{t+1} - \tau_{t+1}$ and the state variable h_t according to the equation $h_{t+1} = h(c_t)$.

The maximization problem of the present self t is given by the current-value function

$$W(h_t, x_t) = \max_{c_t, d_t} \{w(c_t) + v(d_t) + h_t + \beta\delta V(h_{t+1}, x_{t+1}) | x_t - c_t - d_t = 0\}, \tag{3.8}$$

where $V(h_{t+1}, x_{t+1})$ is the continuation value function as in (3.3). The optimal choices solving this problem are again denoted by $c(x_t)$ and $d(x_t)$. In the Appendix 3.A.I it is shown that these functions satisfy the Euler equation

$$\begin{aligned} v'(d(x_t)) = w'(c(x_t)) + \beta\delta h'(c(x_t)) + \{v'(d(x_{t+1})) - (1-\beta)\,[w'(c(x_{t+1})) \cdot c'(x_{t+1}) \\ + v'(d(x_{t+1})) \cdot d'(x_{t+1})]\}\,\delta z_{t+1} h'(c(x_t)). \end{aligned} \tag{3.9}$$

For given subsidy rate z_{t+1}, a solution to this equation consists of a choice function $c(x_t)$ and its derivative $c'(x_{t+1})$. From the budget constraint, these two values then determine $d(x_t)$ and $d'(x_{t+1})$. As there are still two free variables $c(x_t)$ and $c'(x_{t+1})$, there are multiple solutions to the Euler equation, even if one restricts attention to steady states.[7] This arises since the decision of self t depends on her expectation of self $t+1$'s reaction to an increase in income $c'(x_{t+1})$. However, since this increase does not occur in equilibrium, the expectation is not determined in the model.[8]

In our analysis, we focus on a particularly appealing equilibrium, namely the one

[7] This is a common feature of models with quasi-hyperbolic discounting and infinite time horizon. See e.g. Karp (2005), p. 269-271.

[8] In contrast, when the subsidy is paid immediately as in Proposition 1, self t's choice does not affect self $t+1$'s behavior, and hence her behavior is uniquely determined.

where the first-best choice function $c^*(y)$ and its derivative $c^{*\prime}(y)$ solve the Euler equation in every period, that is $c(x_t) = c^*(y)$ and $c'(x_{t+1}) = c^{*\prime}(y)$ for all $t = 0, 1, 2, ...$ To find the subsidy rate which achieves this, solving (3.9) for z_{t+1} yields

$$z_{t+1} = \frac{v'(d(x_t)) - w'(c(x_t)) - \beta\delta h'(c(x_t))}{\delta h'(c(x_t))\{v'(d(x_{t+1})) - (1-\beta)[w'(c(x_{t+1})) \cdot c'(x_{t+1}) + v'(d(x_{t+1})) \cdot d'(x_{t+1})]\}}.$$

By substitution of $c^* = c^*(y)$, $d^* = d^*(y)$, $c^{*\prime}(y)$, and $d^{*\prime}(y)$ from the stationary first-best solution given in (3.5), and observing that $x_{t+1} = y$ if the government runs a balanced budget, we arrive at the following result:

Proposition 2. *If the government subsidizes health outcome in periods $t = 1, 2, ...$ at the rate*

$$z = \frac{(1-\beta)}{v'(d^*) - (1-\beta)[w'(c^*)c^{*\prime}(y) + v'(d^*)d^{*\prime}(y)]} \tag{3.10}$$

and imposes taxes $\tau_0 = 0$ and $\tau_t = \tau = zh(c^*)$ in all subsequent periods $t = 1, 2, ...$, the first-best behavior $c_t = c^*(y)$, $d_t = d^*(y)$ in all $t = 0, 1, 2, ...$ is an equilibrium despite present-biased preferences.

To interpret the rate z in (3.10), note first that for $\beta = 1$ the numerator is zero and there is no need for subsidizing health-conscious consumption. For $\beta < 1$, we multiply the numerator and the denominator on the right-hand-side of (3.10) by $\delta h'(c^*)$ and rearrange to obtain

$$z\delta h'(c^*)\big\{v'(d^*) - (1-\beta)[w'(c^*)c^{*\prime}(y) + v'(d^*)d^{*\prime}(y)]\big\} = \delta(1-\beta)h'(c^*). \tag{3.11}$$

The right-hand-side of (3.11) equals, as in (3.7), the marginal benefit of healthy consumption which self t does not take into account because of her present bias. Considering the left-hand-side of (3.11), we observe first that increasing healthy consumption by one unit in period t increases the subsidy in period $t+1$ by $zh'(c^*)$ units. Moreover, the curly bracket is $\beta\partial V(h_{t+1}, x_{t+1})/\partial x_{t+1}$, evaluated at first-best values.[9] Multiplied by δ, this is the utility gain accruing to self t if the income of self $t+1$ is raised by one unit. Altogether, the left-hand-side of (3.11) describes the additional utility that self

[9] See equation (A.5) in the Appendix 3.A.I.

t acquires through the subsidy if she increases healthy consumption by one unit. The optimal rate z is set such that this subsidy-induced utility gain equals the bias in the evaluation of future health benefit, thereby correcting for the bias.

Comparing (3.10) to the optimal subsidy rate in case of immediate subsidization in (3.7), one first notices the discount factor δ and the marginal health impact of consumption $h'(c^*)$ in the numerator of (3.7). These differences reflect the facts that the immediate subsidy is paid one period earlier and based on consumption of the healthy good rather than on health outcome. Moreover, both forms of subsidy differ because of two other, less obvious, effects which are generated by present-biased behavior.

The first effect, which we label as *discounting effect*, arises because self t, who takes the decision on healthy consumption, evaluates period $t+1$ income differently from self $t+1$, who receives the subsidy. Since for self t, this additional income accrues in the future, she disregards the fraction $1-\beta$ of the benefits procured by marginal spending on both goods. Formally, this is expressed by the fact that in the denominator of (3.10) the term $(1-\beta)[w'(c^*)\cdot c^{*\prime}(y)+v'(d^*)\cdot d^{*\prime}(y)]$ is subtracted from the marginal utility of income. This effect raises the optimal future subsidy rate z compared to the optimal current rate $\tilde{z}$.

The second effect, which we label as *instrumental effect*, occurs since the future subsidy allows self t to shift self $t+1$'s spending in a way self t appreciates. From self t's perspective, there should be no additional discounting of health benefit from period $t+2$ to period $t+1$. Since self $t+1$ takes her decision subject to such a bias, the current self anticipates that the future self spends less on healthy consumption than what the current self considers optimal.

To see that the future subsidy provides an instrument for self t to correct this bias, use (3.5) in the denominator of (3.10) and observe that $c^{*\prime}(y)+d^{*\prime}(y)=1$. The optimal future subsidy rate can then be written as

$$z=\frac{1-\beta}{\beta v'(d^*)+(1-\beta)\delta h'(c^*)c^{*\prime}(y)}\,. \tag{3.12}$$

If the marginal propensity to consume the healthy good $c^{*\prime}(y)$ is zero, then the denominator of (3.12) reduces to $\beta v'(d^*)$, self t's evaluation of self $t+1$'s marginal utility of

income. However, when some of the additional income is spent on the healthy good ($c^{*\prime}(y) > 0$), self t values an additional unit of subsidy higher than $\beta v'(d^*)$. From self t's perspective, the health benefit in period $t+2$ is undervalued by self $t+1$, and hence self t welcomes any additional spending on the healthy good. Consequently, the optimal subsidy rate decreases in the marginal propensity to consume the healthy good $c^{*\prime}(y)$. Compared to the immediate rate $\tilde{z}$, this effect tends to reduce the optimal future subsidy rate z.

The importance of the income effect is illustrated by means of two special cases, to which we now turn.

3.4.3 Special Cases

The cases we consider are characterized by quasi-linear preferences regarding consumption goods c and d respectively. To focus on the shape of utility functions w and v, we assume that $h(c) = c$ in both cases.

In the first special case, the numeraire enters utility linearly, so that $v(d) = d$, whereas the healthy good has decreasing marginal utility, $w''(c) < 0$. Then $c^{*\prime}(y) = 0$ and $d^{*\prime}(y) = 1$, and the optimal subsidy rates from (3.7) and (3.10) are $\tilde{z} = \delta(1-\beta)$ and $z = (1-\beta)/\beta$, implying $\tilde{z} = \delta\beta z$. As explained above, the optimal future subsidy rate reflects both the current self's biased valuation of future income (the discounting effect) and her benefit from changing the future self's behavior (the instrumental effect). In this special case, where no part of additional income is spent on the healthy good, only the discounting effect is present. Therefore, the ratio of the current to the future subsidy rate simply reflects the current self's discounting of future income.

The second special case is given by a utility function where the healthy good enters linearly, while numeraire d has decreasing marginal utility, $v''(d) < 0$. In order to make both cases comparable, we again fix the total marginal utility of the linear good to unity, implying, with $h(c) = c$, that $w(c) = (1-\delta)c$. Then $c^{*\prime}(y) = 1$ and $d^{*\prime}(y) = 0$ and the optimal subsidy rates from (3.7) and (3.10) are $\tilde{z} = \delta(1-\beta)$ and $z = (1-\beta)/[\beta+\delta(1-\beta)]$, yielding $\tilde{z} = z[\delta\beta+\delta^2(1-\beta)]$. Hence, in this case, the ratio of current to future subsidy rate is larger.

To explain this, we first observe that the immediate subsidy rates are the same

regardless of whether the preferences are linear in the healthy good or in the numeraire. Considering future subsidy rates, one notices that the term $\delta(1-\beta)$, which describes the disregarded health effect in the future, appears only in the case when the healthy good enters utility linearly. In this case, the additional income procured by the future subsidy will be entirely used for consumption of the healthy good. Hence, the instrumental effect is maximal in this case. Therefore, the subsidy appears more useful to self t than in the general case when both goods are normal, and even more useful compared to the first special case where no such correction is achieved. As a consequence, a smaller future subsidy rate is sufficient to induce unbiased behavior.

3.5 Tax Revenues

We now compare the tax revenue necessary to induce first-best behavior by immediate subsidies on consumption of the healthy good with the taxes required to reach the same goal by future subsidies on the health outcome. Using $\tilde{z}$ and z from (3.7) and (3.10), the present values of immediate ($\tilde{T}$) and future (T) subsidies are

$$\tilde{T} = \sum_{t=0}^{\infty} \delta^t \tilde{z} c^* = \frac{\tilde{z} c^*}{1-\delta}, \tag{3.13}$$

$$T = \sum_{t=1}^{\infty} \delta^t z h(c^*) = \frac{\delta z h(c^*)}{1-\delta}. \tag{3.14}$$

This comparison is motivated by the fact that in general, taxes induce some welfare loss. Instead of modeling such costs explicitly, we simply assume that the government prefers the form of subsidization that results in lower present value of taxes. That is, formally, the government has lexicographic preferences over the allocation (c, d) and tax revenues.

Alternatively, one can easily introduce a simple type of excess burden consisting of a product of two factors: exogenous marginal costs of public funds (α) and tax revenue. Integrating this excess burden, per period utility from (3.1) is modified to

$$u_t = w(c_t) + v(d_t) + h_t - \alpha\tau_t.$$

Using this formulation, one obtains the government's evaluation of the intertemporal utility which the individual achieves when an immediate subsidy according to Proposition 1 is paid:

$$\tilde{U} = h_0 + \frac{1}{1-\delta}\big[w(c^*) + v(d^*) + \delta h(c^*)\big] - \alpha\tilde{T}.$$

In the same way we compute the intertemporal utility with future subsidy according to Proposition 2:

$$U = h_0 + \frac{1}{1-\delta}\big[w(c^*) + v(d^*) + \delta h(c^*)\big] - \alpha T.$$

Clearly, in both cases utility is decreasing in the present value of tax payments. Therefore, it is worthwhile to ask which of the two subsidy schemes induces the first-best consumption at lower cost to the government.[10]

The following proposition answers this question.

Proposition 3. *If $\beta < 1$, inducing unbiased choices by immediately subsidizing healthy consumption requires a lower present value of taxes than inducing unbiased choices by subsidizing the future health outcome, $\tilde{T} < T$.*

Proof: Inserting (3.7) and (3.12) in (3.13) and (3.14) shows that $\tilde{T} < T$ is equivalent to

$$\frac{\delta}{1-\delta} \cdot \frac{(1-\beta)h'(c^*)c^*}{v'(d^*)} < \frac{\delta}{1-\delta} \cdot \frac{(1-\beta)h(c^*)}{\beta v'(d^*) + (1-\beta)\delta h'(c^*)c^{*\prime}(y)}.$$

For $\beta < 1$, this inequality is equivalent to

$$h'(c^*)c^* \cdot \big[\beta v'(d^*) + (1-\beta)\delta h'(c^*)c^{*\prime}(y)\big] < v'(d^*)h(c^*)\,. \tag{3.15}$$

Now observe that $c^{*\prime}(y) = v''/(w'' + \delta h'' + v'') \leq 1$ and that the concavity of $h(c^*)$ implies with $h(0) = 0$ that $h'(c^*)c^* \leq h(c^*)$. Therefore, one has

$$h'(c^*)c^* \cdot \big[\beta v'(d^*) + (1-\beta)\delta h'(c^*)c^{*\prime}(y)\big] \leq h(c^*) \cdot \big[\beta v'(d^*) + (1-\beta)\delta h'(c^*)\big].$$

[10] In a more elaborate set-up, one might account for the excess burden when determining the optimal policy. This will probably result in lower subsidy rates, since there is then a trade-off between the welfare cost of taxation and the health-improvement. However, this trade-off is not the subject of this paper, where we instead focus on costs of corrective policies which implement the first-best choice.

With (3.5), this inequality is equivalent to

$$h'(c^*)c^* \cdot \left[\beta v'(d^*) + (1-\beta)\delta h'(c^*)c^{*\prime}(y)\right] \leq h(c^*) \cdot \left[v'(d^*) - (1-\beta)w'(c^*)\right].$$

For $\beta < 1$, this inequality implies (3.15) and hence $\tilde{T} < T$. *Q.E.D.*

This result shows that the future subsidy is more expensive in terms of tax revenues required than the immediate subsidy. Politically this means that those subsidy instruments which reward health-conscious behavior such as immediate financial rewards are preferable to instruments which reward health outcomes such as health insurance rebates.

Proposition 3 is a consequence of the various effects determining the optimal subsidy rates explained after Proposition 2. The dominating force is the discounting effect, that is the difference in discounting between the individual decision maker and the government. The future subsidy achieves the same behavioral response as the present subsidy only if it is sufficiently high to compensate for the current self's present bias. The government, in contrast, is unbiased in its intertemporal evaluation of tax revenues. This is most clearly seen in the special case where the numeraire enters utility in a linear fashion. In this case, to achieve the first-best, the future subsidy rate must be $1/\beta\delta$ times higher than the immediate subsidy rate. When calculating the present values, the government discounts the future subsidy only with δ. Therefore, the present value of the future subsidy still exceeds the present value of the immediate subsidy by the factor $1/\beta$.

In general, however, this result is mitigated by the instrumental effect. As discussed after Proposition 2, the future subsidy allows self t to counteract the present bias of self $t+1$. Since this makes the future subsidy more valuable for self t, a smaller subsidy rate is sufficient to achieve first-best consumption. Consequently, the present value of the future subsidy is reduced to some extent. As Proposition 3 shows, however, the discounting effect dominates the instrumental effect, so that the immediate subsidy still involves lower taxes.

3.6 Present-biased Policymaker

In this section we take a different perspective on the political decision making process. Until now we considered a paternalistic government with time-consistent preferences which induces the individual to behave as if there was no present bias. In the following we assume that the present-biased individual herself forms the government. This means that the government on the one hand, like the individual, discounts more heavily between the current and the following periods than between later periods. On the other hand, like the paternalistic government, it can impose taxes and pay subsidies. In this scenario two questions arise: Will the present-biased policymaker commit to a subsidy scheme which implements the first-best choice? If so, which form of the subsidy will she prefer?

3.6.1 Commitment

To address the first issue, we compare the intertemporal utility that the present-biased government obtains with the first-best allocation to the utility it would achieve without any intervention.[11] This laissez-faire allocation is given by the solution to the hyperbolic self's decision problem analyzed in Subsection 3.4.1 with $\tilde{z}_t = \tilde{\tau}_t = 0$. From the first-order condition (3.6), using $\partial V(h_{t+1}, x_{t+1})/\partial h_{t+1} = 1$, a stationary solution (c_β, d_β) satisfies $w'(c_\beta) + \beta\delta h'(c_\beta) = v'(d_\beta)$ and the budget constraint $c_\beta + d_\beta = y$. In this solution, healthy consumption is lower and numeraire consumption is larger than in the first-best, $c_\beta < c^*, d_\beta > d^*$.

The present-biased policymaker will commit to a subsidy scheme if the intertemporal utility U^* provided by first-best consumption (c^*, d^*) exceeds the intertemporal utility U_β provided by (c_β, d_β). Inserting c^* and d^* or, respectively, c_β and d_β in (3.1) and

[11] In a more general approach one could allow the present-biased government to choose an allocation which is different from these two. While it would be easy to characterize such an allocation, we restrict attention to the choice between first-best and laissez-faire. By keeping the subsidy rates and the resulting allocation unchanged, we focus on the choice of timing of the subsidy, thus maintaining comparability with the previous analysis.

using (3.2) yields

$$\begin{aligned} U^* &= h_0 + w(c^*) + v(d^*) + \beta\delta h(c^*) + \beta \sum_{t=1}^{\infty} \delta^t [w(c^*) + \delta h(c^*) + v(d^*)] \\ U_\beta &= h_0 + w(c_\beta) + v(d_\beta) + \beta\delta h(c_\beta) + \beta \sum_{t=1}^{\infty} \delta^t [w(c_\beta) + \delta h(c_\beta) + v(d_\beta)] \end{aligned}$$

Inspecting these two expressions one notices the trade-off faced by the present-biased policymaker in period 0. By implementing the subsidy as of period 0, she induces her own consumption to change to the first-best values which, due to her present-bias, she deems inferior to the laissez-faire values. Hence, she loses

$$\Delta_\beta = w(c_\beta) + v(y - c_\beta) + \beta\delta h(c_\beta) - [w(c^*) + v(y - c^*) + \beta\delta h(c^*)] > 0$$

in period 0. In return she gains

$$\Delta^* = w(c^*) + v(y - c^*) + \delta h(c^*) - [w(c_\beta) + v(y - c_\beta) + \delta h(c_\beta)] > 0$$

in every period $t = 1, 2, ...$ This gain arises since self $t = 0$ likes the consumption of future selves $t = 1, 2, ...$ to be changed to the first-best values. Calculating present values, one finds

$$U^* - U_\beta = -\Delta_\beta + \frac{\beta\delta}{1 - \delta}\Delta^*. \tag{3.16}$$

The present-biased policymaker chooses to implement the subsidy if this expression is non-negative.

Figure 3.1 illustrates the loss Δ_β and the gain Δ^*. For self 0, the marginal benefit of her own healthy consumption is reduced by present-bias, as shown by the line labeled $w'(c) + \beta\delta h'(c)$ in Figure 3.1. For consumption levels between c_β and c^*, this marginal benefit falls short of the marginal benefit of the numeraire, expressed by the line labeled $v'(y - c)$ in Figure 3.1. The resulting loss Δ_β is depicted by the vertically shaded area between these two curves. In contrast, self 0 evaluates the marginal benefit of healthy consumption by future selves $t = 1, 2, ...$ without a present-bias. This marginal benefit is given by the line labeled $w'(c) + \delta h'(c)$ in Figure 3.1. It exceeds the marginal utility of

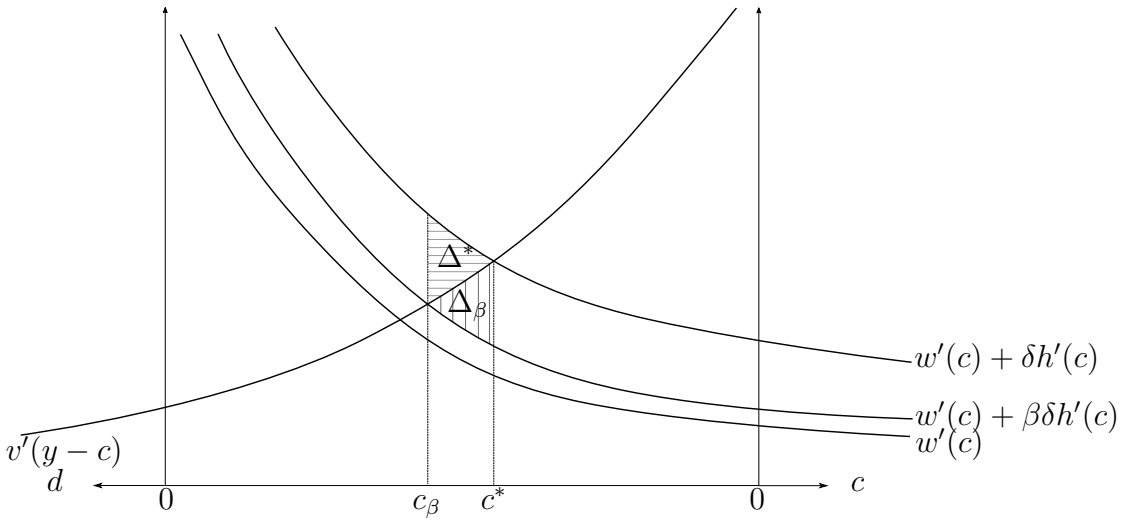


Figure 3.1: Gain (Δ^*) and loss (Δ_β) procured to the present-biased policymaker when committing to a subsidy scheme

the numeraire for consumption levels between c_β and c^*. The gain Δ^* is then represented by the horizontally shaded area between the $w'(c) + \delta h'(c)$ and $v'(y-c)$ curves.

Clearly, the sign of $U^* - U_\beta$ in (3.16) depends on the exact sizes of these areas, and hence on the shape of the marginal utility schedules involved. However, if these do not display too strong curvatures, the decision on the subsidy scheme is essentially determined by the discount rates, as the following proposition shows.

Proposition 4. *Assume that $w''' = v''' = h''' = 0$. Then the present-biased policymaker will commit to one of the subsidy schemes described in Propositions 1 and 2 if and only if*

$$\frac{\beta\delta}{1-\delta} \geq \frac{h'(c^*)}{h'(c_\beta)}. \tag{3.17}$$

Proof: With $w''' = v''' = h''' = 0$, the marginal cost curves $w + \beta\delta h'$, $w + \delta h'$, and v' are linear. Hence, $\Delta^* = (c^* - c_\beta)(1-\beta)\delta h'(c_\beta)/2$ and $\Delta_\beta = (c^* - c_\beta)(1-\beta)\delta h'(c^*)/2$ (see Figure 3.1). Inserting in (3.16) shows that $U^* - U_\beta \geq 0$ is equivalent to (3.17). *Q.E.D.*

Note that this result can easily be extended to the case where marginal utilities are not exactly linear, as long as their curvatures are not too strong. Thus, the present-biased policymaker will still choose to implement (not to implement) the first-best

consumption levels if $\beta/(1-\beta) > (<)\, h'(c^*)/h'(c_\beta)$ as long as the third derivatives of the utility functions are not too large.

Moreover, we note that this choice is time-consistent. Since the policymaker in period 1 is in the same situation as the policymaker in period 0, she will commit to the subsidy if and only if the policymaker in period 0 commits to it. Therefore, she will not abolish the subsidy once it is introduced. Conversely, it would not be time-consistent if the policymaker in period 0 decided to implement the subsidy scheme only from period 1 onward. While this is the best choice for this policymaker, the subsequent government would behave in the same way and hence postpone the starting date for the subsidy scheme by one more period, and so on.

Proposition 4 shows that a government which is formed by the individuals sometimes fails to implement the first-best policy. This occurs if the future gains from committing to increased healthy consumption count too little compared to the immediate loss from forcing a change of behavior upon oneself. On the other hand, if discounting is not too strong, that is, if inequality (3.17) is satisfied, one does not have to appeal to an outside paternalistic government in order to induce first-best.

3.6.2 Tax Revenues

We now turn to the choice between immediate and future subsidies. We assume that the present-biased government, like the paternalistic government, prefers the form of subsidies that requires the lower present value of taxes. However, the present values of immediate ($\widetilde{T}_\beta$) and future (T_β) subsidies are calculated taking into account the policymaker's present-bias:

$$\widetilde{T}_\beta = \delta^0 \widetilde{z} c^* + \beta \sum_{t=1}^{\infty} \delta^t \widetilde{z} c^* = \frac{[1-\delta(1-\beta)]\widetilde{z}c^*}{1-\delta}\,, \tag{3.18}$$

$$T_\beta = \beta \sum_{t=1}^{\infty} \delta^t z h(c^*) = \frac{\beta\delta z h(c^*)}{1-\delta}\,. \tag{3.19}$$

The following proposition shows that the evaluation of tax revenues by the policymaker now depends on the extent of her bias towards the present.

Proposition 5. *There is a critical* $\hat{\beta}$ *with* $0 \leq \hat{\beta} \leq 1$ *such that, for all* $0 < \beta < 1$

$$T_\beta \left\{ \begin{matrix} \leqq \\ > \end{matrix} \right\} \widetilde{T}_\beta \text{ if and only if } \beta \left\{ \begin{matrix} \leqq \\ > \end{matrix} \right\} \hat{\beta}.$$

Proof: See the Appendix 3.A.II.

In contrast to Proposition 5 the comparison of tax revenues by the present-biased policymaker does not always favor the immediate subsidy scheme. This can be understood by considering how the different computation of present values by the paternalistic and the present-biased government interacts with the effects determining optimal subsidy rates. The present-biased government discounts future tax payments more heavily, therefore in its computation the discounting effect raises the cost of future subsidies less than in the computation of the paternalistic government. As a consequence it is possible that the instrumental effect dominates.

This can be illustrated by considering the two special cases presented in Subsection 3.4.3. In the first special case, where $c^{*\prime}(y) = 0$, the instrumental effect is absent. As shown in the proof of Proposition 5, in this case $\hat{\beta} = 0$ implying that the future subsidy is always more expensive that the immediate subsidy. Thus in this case both types of government evaluate tax revenues in the same way.

In the second special case, where $c^{*\prime}(y) = 1$, the instrumental effect is strongest. In this case $T_\beta < \tilde{T}_\beta$ for all $0 < \beta < 1$,[12] or equivalently $\hat{\beta} = 1$. Thus the strong instrumental effect outweighs the discounting effect, so that the future subsidy is always less expensive compared to the immediate subsidy.

Finally, we note that this example shows that there are parameter constellations such that in the same time inequality (3.17) holds and $T_\beta < \tilde{T}_\beta$. Thus, even if the present-biased policymaker agrees with the paternalistic government that the first-best allocation is preferable to laissez-faire, she may choose a different form of subsidy.

[12] To see this, insert $v'(d^*) = h'(c^*) = \varepsilon = 1$ in equation (A.7) in the Appendix 3.A.II.

3.7 Conclusion

This paper analyzes the intrapersonal game that arises when a consumer with present-biased preferences faces an intertemporal consumption decision. In this setting, we examine two forms of subsidizing health conscious behavior: immediate subsidies related to healthy consumption and future subsidies paid for a good health outcome. We show that while both subsidies can achieve the first-best outcome, it then very much depends on the policymaker's preferences which one of the subsidy schemes will be implemented.

This choice is driven by the balance of two effects determining the effectiveness of the future subsidy. On the one hand, present-biased consumers perceive future subsidy payments as less valuable, and hence, the effectiveness of rewards for health-conscious behavior declines the further they are in the future. On the other hand, procuring later selves with additional income raises future health-conscious consumption via an income effect, which makes future subsidies more effective. Our first result shows that, for a paternalistic government, the first effect always dominates so that the future subsidy results in higher costs measured in present value terms. Politically, this implies that the paternalistic government should concentrate on rewarding health-conscious behavior, e.g. by lowering the prices of healthy goods, rather than promising future rewards for successful health investments.

We contrast this result with the decision taken by a present-biased government which is formed by time-inconsistent individuals. We show that such a government will commit to a subsidy scheme which implements the unbiased choice when future gains from commitment to increased healthy consumption exceed the immediate loss from forcing a behavior change. However, in contrast to the first result, the comparison of tax revenues by the biased government does not always favor the immediate subsidy scheme. Since the present-biased government uses a stronger discount factor, the cost of future subsidies count for relatively less. If the income effect on future behavior is strong enough, the present-biased government will therefore favor the future subsidy scheme.

Our results suggest a number of extensions, two of which we briefly discuss. The first extension concerns the effectiveness of subsidies, which may not only depend on

the timing, but also on the type of reward. As the examples given in the introduction illustrate, premia for health-related activities are often awarded in kind. This matters since we could imagine that for present-biased consumers money is a more attractive immediate reward than, say, a free fitness card. In the same time, money can be spent on many things including consumption which damages health, whereas a fitness card procures an additional health benefit. Hence, it might be interesting to find out the optimal combination of monetary and in-kind subsidies for health.

As a second extension, one can ask the general question of whether or not the intervening government can be assumed to have time-consistent preferences. If, as in our model, all individuals have to some extent present-biased preferences, who will form the unbiased government? Alternatively, if there is a minority of unbiased individuals, can we expect them to be elected by the biased majority? Conversely, when the government is composed of biased individuals, how likely is that it will implement a policy that actually goes against the current preferences of its members? And if so, will it be able to put an end to its own postponing game? Finally, do the paternalistic or present-biased governments have to give a reason for their action?

In our view, a convincing theory of policy intervention for correcting biased preferences should address such political economy issues. These considerations are, however, beyond the scope of this paper and will be the subject of further research.

Appendix to Chapter 3

Appendix 3.A.I: Derivation of the Euler Equation

Computing the first-order conditions for a solution to (3.8) and using $\partial h_{t+1}/\partial c_t = h'(c_t)$ and $\partial x_{t+1}/\partial c_t = z_{t+1}h'(c_t)$, we obtain

$$\lambda_t = w'(c(x_t)) + \beta\delta\left(\frac{\partial V(h_{t+1}, x_{t+1})}{\partial h_{t+1}} + \frac{\partial V(h_{t+1}, x_{t+1})}{\partial x_{t+1}} \cdot z_{t+1}\right) \cdot h'(c(x_t)) = v'(d(x_t))\,, \tag{A.1}$$

where λ_t is the Lagrange variable associated to the budget constraint. Inserting optimal choices in the current-value function for the present-biased consumer gives

$$W(h_t, x_t) = w(c(x_t)) + v(d(x_t)) + h_t + \beta\delta V(h_{t+1}, x_{t+1}). \tag{A.2}$$

Using the laws of motion for x_{t+1} and h_{t+1}, one derives from (A.2)

$$\frac{\partial W(h_t, x_t)}{\partial x_t} = \left[w'(c(x_t)) + \beta\delta\left(\frac{\partial V(h_{t+1}, x_{t+1})}{\partial h_{t+1}} + \frac{\partial V(h_{t+1}, x_{t+1})}{x_{t+1}} \cdot z_{t+1}\right) \cdot h'(c(x_t))\right] \cdot c'(x_t) + v'(d(x_t)) \cdot d'(x_t),$$

$$\frac{\partial W(h_t, x_t)}{\partial h_t} = 1.$$

From the envelope theorem, we have

$$\frac{\partial W(h_t, x_t)}{\partial x_t} = \lambda_t = v'(d(x_t)). \tag{A.3}$$

Substituting next period's optimal choices $c(x_{t+1})$ and $d(x_{t+1})$ into the continuation-value function yields

$$V(h_{t+1}, x_{t+1}) = w(c(x_{t+1})) + v(d(x_{t+1})) + h_{t+1} + \delta V(h_{t+2}, x_{t+2}). \tag{A.4}$$

From (A.4) and the equivalent of (A.2) for period $t+1$, the current-value function and the continuation-value function are linked by the equation[13]

$$\beta \cdot V(h_{t+1}, x_{t+1}) = W(h_{t+1}, x_{t+1}) - (1-\beta)[w(c(x_{t+1})) + v(d(x_{t+1})) + h_{t+1}].$$

By differentiation and substitution of the version of (A.3) for period $t+1$ we get

$$\beta \cdot \frac{\partial V(h_{t+1}, x_{t+1})}{\partial x_{t+1}} = v'(d(x_{t+1})) - (1-\beta)\left[w'(c(x_{t+1})) \cdot c'(x_{t+1}) + v'(d(x_{t+1})) \cdot d'(x_{t+1})\right] \tag{A.5}$$

with $w'(c(x_{t+1})) \cdot c'(x_{t+1}) + v'(d(x_{t+1})) \cdot d'(x_{t+1})$ as the current marginal utility of income. From the equation (A.1) we have

$$\beta \cdot \frac{\partial V(h_{t+1}, x_{t+1})}{\partial x_{t+1}} = \frac{v'(d(x_t)) - w'(c(x_t)) - \beta\delta \frac{\partial V(h_{t+1}, x_{t+1})}{\partial h_{t+1}} \cdot h'(c_t)}{\delta z_{t+1} \cdot h'(c_t)}. \tag{A.6}$$

Combining (A.5) and (A.6), and using $\partial V(h_{t+1}, x_{t+1})/\partial h_{t+1} = 1$, we obtain the Euler equation (3.9).

[13] See also Harris and Laibson (2001), p. 940.

Appendix 3.A.II: Proof of Proposition 5

Insert $\tilde{z}$ from (3.7) and z from (3.12) in (3.18) and (3.19) and define the elasticity of health with respect to healthy consumption at the first-best value by $\varepsilon = h'(c^*)c^*/h(c^*)$. It then follows that for all $0 < \beta, \delta < 1$

$$T_\beta \left\{ \begin{matrix} \leqq \\ \geqq \end{matrix} \right\} \widetilde{T}_\beta \Longleftrightarrow \frac{\beta v'(d^*)}{\beta v'(d^*) + (1-\beta)\delta h'(c^*)c^{*\prime}(y)} \left\{ \begin{matrix} \leqq \\ \geqq \end{matrix} \right\} \varepsilon[1-\delta(1-\beta)] \tag{A.7}$$

We denote the left-hand side of (A.7) by $\Theta(\beta)$ and the right-hand side by $\widetilde{\Theta}(\beta)$.

To evaluate $\Theta(\beta) \left\{ \begin{matrix} \leqq \\ \geqq \end{matrix} \right\} \widetilde{\Theta}(\beta)$ we collect several properties of these functions. The boundary behavior at $\beta \to 0$ is given by

$$\lim_{\beta\to 0} \Theta(\beta) = \begin{cases} 0 & \text{if} \quad c^{*\prime}(y) > 0 \\ 1 & \text{if} \quad c^{*\prime}(y) = 0, \end{cases} \tag{A.8}$$

$$\lim_{\beta\to 0} \widetilde{\Theta}(\beta) = \varepsilon(1-\delta) \tag{A.9}$$

where in the last line of (A.8) we use L'Hôpital's rule. At $\beta \to 1$ we find

$$\lim_{\beta\to 1} \Theta(\beta) = 1. \tag{A.10}$$

$$\lim_{\beta\to 1} \widetilde{\Theta}(\beta) = \varepsilon \tag{A.11}$$

Computing the derivatives yields

$$\Theta'(\beta) = \frac{\delta h'(c^*)c^{*\prime}(y)v'(d^*)}{[\beta v'(d^*) + (1-\beta)\delta h'(c^*)c^{*\prime}(y)]^2} \geq 0 \tag{A.12}$$

$$\widetilde{\Theta}'(\beta) = \delta\varepsilon > 0 \tag{A.13}$$

Since $v'(d^*) > \delta h'(c^*)c^{*\prime}(y)$, we have $\Theta''(\beta) \leq 0$, with strict inequality if $c^{*\prime}(y) > 0$.

Consider first $c^{*\prime}(y) = 0$. According to (A.8), (A.10) and (A.12) it follows that $\Theta(\beta) = 1$ for all $0 < \beta < 1$. From (A.11) and (A.13) we have $\widetilde{\Theta}(\beta) < \varepsilon$ for all $0 < \beta < 1$. Since $\varepsilon \leq 1$ this implies $\Theta(\beta) > \widetilde{\Theta}(\beta)$ for all $0 < \beta < 1$. Hence the claim is true for $\hat{\beta} = 0$.

Consider now $c^{*\prime}(y) > 0$ and assume first that $\varepsilon < 1$. From (A.8) and (A.9) it holds $\Theta(\beta) < \widetilde{\Theta}(\beta)$ for β close to zero. From (A.10) and (A.11) one has $\Theta(\beta) > \widetilde{\Theta}(\beta)$ for β close to one. Hence there is an odd number of intersections of the functions $\Theta(\beta)$ and $\widetilde{\Theta}(\beta)$ in the interval $(0, 1)$. Since $\Theta''(\beta) < 0$ and $\widetilde{\Theta}''(\beta) = 0$ there can be at most two such intersections. Altogether, we conclude that there is a unique intersection $\hat{\beta} \in (0, 1)$ such that $\Theta(\hat{\beta}) = \widetilde{\Theta}(\hat{\beta})$. For $\beta < \hat{\beta}$ ($\beta > \hat{\beta}$), we have $\Theta(\beta) < \widetilde{\Theta}(\beta)$ ($\Theta(\beta) > \widetilde{\Theta}(\beta)$), and hence $T_\beta < \widetilde{T}_\beta$ ($T_\beta > \widetilde{T}_\beta$) as claimed.

Finally we consider $c^{*\prime}(y) > 0$ and $\varepsilon = 1$. From (A.8) and (A.9) one sees again $\Theta(\beta) < \widetilde{\Theta}(\beta)$ for β close to zero. From (A.10) and (A.11) one obtains $\lim_{\beta \to 1} \Theta(\beta) = \lim_{\beta \to 1} \widetilde{\Theta}(\beta)$. As before, since $\Theta(\beta)$ is strictly concave and $\widetilde{\Theta}(\beta)$ is linear, there can be at most one intersection of both functions within the interval $(0, 1)$. If such an intersection exists it is $\hat{\beta}$ as claimed in the proposition. Otherwise $\hat{\beta} = 1$ and $T(\beta) < \widetilde{T}(\beta)$ for all $0 < \beta < 1$. *Q.E.D.*

Chapter 4

Determinants of Students' Success at University*

4.1 Introduction

The number of students in higher education worldwide is constantly increasing. Today's students are more heterogeneous than ever before and possess a wide and diverse range of characteristics and abilities. They often differ in educational background, social status, skills, and academic potential, among others. As the diversity of the student population increases, factors predicting students' academic performance become a matter of concern for institutions in the educational sector (Burton and Dowling, 2005; Simpson, 2006). For example, knowledge about factors affecting academic success is relevant for universities when selecting the most promising students. At an aggregate level, based on such knowledge, policy can decide to what extent investment in tertiary education should be directed towards those fields where large numbers of students can expect to succeed, or be concentrated in fields which rather cater to a minority of excellent students.

Our study addresses this concern by focusing on the question of whether, and if so to what extent student characteristics can be used for predicting academic success. We find a highly significant and positive effect of the high school leaving grade on academic performance. Additionally, we narrow our view towards differences between fields of study, grouped by faculties. We find that the importance of the high school leaving grade differs strongly between fields. In some faculties graduation is less difficult to achieve, but not necessarily associated with a good final grade. However, in other faculties, graduation seems to be less likely, but among those students who graduate, the

*This chapter originates from joint work with Johannes Meya, Katharina Suntheim and Robert Schwager (see Danilowicz-Gösele et al. (2014)).

final university grade is on average better and less differentiated. This points towards diverging teaching and examination cultures among faculties. Some of them specialize in preparing a positive selection of students to science or demanding employment, whereas others provide an education which is accessible for large numbers of high school graduates with average abilities.

The probability of academic success and the reasons for dropping out of university are subject of the continuously expanding research literature in many areas, notably economics of education, psychology and sociology. These studies provide a consistent picture of previous high school performance as the most prominent predictor of university success (Baron-Boldt, 1989; Betts and Morel, 1999; Cyrenne and Chan, 2012). Furthermore, various other personal characteristics are found to affect students' academic performance, for instance gender (McNabb et al., 2002), age (Hong, 1984) or socio-economic status (Arulampalam et al., 2005). Besides, the type of high school visited is shown to influence both the probability of entering a college (Altonji et al., 2005) as well as the probability of obtaining a good degree (Smith and Naylor, 2005). At university, also the chosen field of study might matter (Achen and Courant, 2009).

Although there is a vast amount of literature on factors predicting academic success, our paper differs from previous work in this area in a number of ways. Firstly, to the best of our knowledge, this is the first paper that analyzes a comprehensive administrative data set of student population, that aims to be an encompassing analysis of students' characteristics as predictors for academic success at university in Germany. In contrast to much of the earlier work, we can track students' academic careers from the admission day onward. For instance, we observe changes in fields of study. Secondly, we analyze not only one but three dimensions of academic success: graduation from the university, graduation within a chosen field of study and final grade of the university degree. Thirdly, differentiating between faculties allows us to observe different examination cultures.

The remainder of the paper is structured as follows: In Section 4.2 we present a brief overview of the related literature. In Section 4.3 we describe our dataset, explain the variables used, and lay out the empirical setup. We turn our attention to our empirical results in Section 4.4 and conclude with a discussion of the implications of these results

in Section 4.5.

4.2 Literature

As the universities' selection process is often based on high school performance, almost all literature dealing with students' academic performance examines in the first place whether the high school Grade Point Average (GPA) is a valid predictor for university success. According to the meta-analysis of Robbins et al. (2004), the correlation between secondary school grades and university GPA is on average about 0.41. Trapmann et al. (2007) find a mean corrected validity between 0.26 and 0.53 for high school grades predicting university success by using a meta-analysis approach including studies from Austria, Czech Republic, Germany, Great Britain and Norway. In this sample, the German high school GPA has the highest validity.

However, the predictive effectiveness of secondary school grades on academic performance seems to be different for diverse groups. For instance, Dobson and Skuja (2005) show that high university entrance scores are indeed a good predictor, but not for every field of study. They find a strong correlation between the university entrance scores and students' academic performance in agriculture, engineering and science, and almost no correlation in education and health studies. This corresponds to the results of Trapmann et al. (2007) who find a high predictive power for engineering and natural sciences and a comparatively low validity for psychology.

There is also a large number of contributions showing that students with the same entry grades are often found to perform differently in tertiary education, which suggests the importance of other factors when predicting university success. Based on an analysis of about 300 students in a regional equity and access program of Monash University, Australia, Levy and Murray (2005) report that an appropriate coaching program can reduce the impact of discrepancy in university entrance scores. Consequently, the entrance scores themselves may not be able to capture all relevant student characteristics.

In a study by Grebennikov and Skaines (2009) at the University of Western Sydney, data relating to about 9,000 students was analyzed in order to determine a set of variables predicting students' academic performance and retention. They find that the

odds of dropping out without applying to other educational institutions are significantly higher for part-time and mature students, who tend to have less time for studying and face stricter financial constraints. Furthermore, the probability of early withdrawal from university is particularly high for students from an English-speaking background and with a low grade point average.

An analysis of academic, psychological, cognitive, and demographic predictors for academic performance can be found in McKenzie and Schweitzer (2001). For this purpose, they examine a group of about 200 first-year students and find significant coefficients for the university entry score (accounting for 39% of the variance in GPA), student institution integration (accounting for 3% of the variance in GPA) and self-efficacy (accounting for 8% of the variance in the GPA). When both the measure of integration and the measure of self-efficacy are included in the model, the prediction of GPA at university is improved by 12%.

Looking at a data set of the population of newly enrolled students at the University of Brussels, Arias Ortiz and Dehon (2008) examine the probability of succeeding the first year at university by accounting for individual characteristics, prior schooling and socio-economic background. According to their results, socio-economic background, especially the mother's level of education and the father's occupational activity, matters for students' academic success. In addition, they observe differences in academic performance between students coming from different high school programs.

Other factors mentioned in the literature that may help identify students at risk of failing include: standardized pre-university tests (Cohn et al., 2004), study skills (Robbins et al., 2004), the ability to adapt to the university environment (McInnis et al., 2000; Peat et al., 2001) or first-year experience at the university (Krause et al., 2005). Further studies emphasize the importance of psychosocial variables such as goal and institutional commitment (Tinto, 1975), emotional intelligence (Parker et al., 2004), relationship with the faculty (Girves and Wemmerus, 1988) and social support (Gerdes and Mallinckrodt, 1994).

Altogether it appears to be generally accepted that high school performance is the best predictor for university success. We confirm this result using a new and comprehensive dataset from a German university. Contrary to the mixed results about the

link between high school GPA and success in specific fields, we find that such a link is present in all faculties, albeit in different forms. Specifically, by distinguishing between several measures of success, we are able to describe in detail how this relationship varies across fields. Finally, again contrasting with some of the results cited, our data does not support the view that social origin or income have strong additional impact on university success once high school grades are taken into account.

4.3 Data and Approach

In our analysis we use an extensive administrative dataset from Göttingen University, Germany, which encompasses detailed, anonymized information on more than 12,000 students. One part of the data is collected when students enroll at university and contains information about the student's high school leaving certificate, her parental address, gender and type of health insurance. The other part includes information about the student's university career, such as the field of study, the reason for her leaving university, whether she obtained a degree and if so, which one.

In addition, we use data on the purchasing power of the German zip-code areas which is provided by *GfK*, a market research firm.[15] The index is based on data provided by the German tax offices as well as other relevant statistics, for instance regarding pensions and unemployment benefits.

Detailed information on data filtering and processing can be found in Appendix 4.A.I.

4.3.1 Variable Description and Institutional Background

We use the following three measures of university success: the probability of finishing studies with a degree, the probability of finishing a chosen field of study with a degree and the grade of the final university degree. For the first two measures, it is necessary to distinguish between students who drop out and those who change institution. For this reason, we exclude students who mention that they leave Göttingen University in order to continue studying at another university from the sample.

[15] *GfK* is one of the biggest companies worldwide in the field of market research and collects information on people's lifestyle and consumption behavior.

As one is generally considered to be a successful student if one holds some degree after finishing university, we first examine a binary variable which describes whether the student graduates at all from university. The variable is equal to one for all students who finish their studies with any kind of degree at Göttingen University, and zero otherwise.

However, since in Germany students have to decide on their field of study as soon as they register for university, it is not uncommon that more than one subject is chosen or that the major is changed within the first few years. Therefore, we narrow down the definition of university success by using an additional outcome variable, labeled 'graduation within faculty', measuring success in each program the student enrolled in. This implies that when a student changes her field of study or enrolls in more than one degree program, several observations are generated. Thereby, success or failure are registered individually for every observation dependent on whether the student obtained a degree in this specific field of study or not. For example, for a student who changed her subject of study once during her university career and completed only the second study subject, the dataset will contain two observations. For the first observation, the variable describing success equals zero, and for the second, it is one. However, as study programs within the same faculty are typically quite similar with respect to their content or required abilities, a change of subject is only seen as a failure if it also implies a change of the faculty.

The third outcome variable is the grade of the university degree. As some students are enrolled in more than one study program or complete two consecutive degree programs, we create individual observations for every final university degree obtained. Furthermore, we transform grades into the U.S. grading scale in order to make results internationally comparable and easier to interpret. In Germany, the grading schedule traditionally ranges from 1.0 to 5.0, with 1.0 being the best grade to achieve and 4.0 the worst grade that is still a pass. This implies that the better the performance, the lower the grade. The outcome variable *university GPA*, which we use in our analysis, is a transformation of the actual grade achieved. It ranges from 1.0 to 4.0 with 4.0 being the best grade to obtain and 1.0 the worst that is still a pass.[16]

[16]We transformed the grades into the U.S. grading scale by subtracting the final university grade from five. For legal studies the special grade *vollbefriedigend* is treated as a 2.5.

The central exogenous variable used in the analysis is the high school GPA, a transformation of the grade of the high school leaving certificate. Similar to the grade of the university degree, it is converted to the U.S. grading scale with 4.0 being the best and 1.0 the worst passing grade.

The students' socio-economic background is captured by two variables: the type of health insurance and the purchasing power of the parents' zip-code area.

Due to a particular institutional feature of the German health insurance system, the type of health insurance can be used as a proxy for the students' educational and socio-economic background. In order to choose a private instead of the generally compulsory public health insurance, one has to earn more than a certain amount of income (2013 : 52,200 Euro gross income per year), be self-employed or work as civil servant. As most students are insured through their parents, the type of health insurance a student holds contains information about whether her parents satisfy at least one of the above criteria. Specifically, a large group of civil servants are teachers, and many self-employed and high earners hold a university degree. Overall, in 2008, 56.7 percent of the people being privately insured held a degree enabling registration at a university or a university of applied science, 38.0 percent had completed university or university of applied science with a degree or a Ph.D (Finkenstädt and Keßler, 2012). Within the total German population, these shares were much lower, amounting to 24.4 and 13.0 percent respectively (Statistisches Bundesamt, 2009).

The second socio-economic variable we use is an index of the purchasing power within the zip-code area of the student's home address evaluated in the year 2007. The index, provided by *GfK*, is measured relative to the German average, and normalized to 100. For example, an index value of 110 means that the purchasing power of this area is 10% higher than the German average. Since German zip-code areas are fairly small, with the biggest cities like Hamburg or Berlin encompassing up to 191 different zip-codes, and assuming a certain degree of residential sorting according to income, we are confident that this local measure approximates the students' economic background reasonably well.

As additional covariates we include indicator variables for male students, the sixteen German states and the university's thirteen faculties.

To get a more diversified picture of the determinants of university success, we also divide the data into sub-samples by faculty. At Göttingen University the various fields of study are assigned to thirteen faculties: theology, law, medicine, humanities, mathematics, physics, chemistry, geology/geography, biology, forestry, agriculture, economic sciences, and social sciences. A detailed analysis of individual faculties seems worthwhile since they may differ with regard to scientific approach, organizational structure and general conditions of studying.

4.3.2 Summary Statistics

The final dataset contains 12,315 students out of which 48% obtained a degree at Göttingen University. The remaining 52% left Göttingen University without completing a degree. Taking into account that students might be enrolled in more than one degree program or change fields of study during their university career increases the number of observations to 16,931. For 49% of these observations the respective field of study is completed with a degree (Table 4.1).

When taking a look at those students who graduated, we see that a final grade is registered for 8,204 observations. This implies that around one third of the students who finished their studies obtained more than one university degree. The reason for this could be the introduction of the consecutive study programs which by definition leads to more than one degree for many students.

The mean university GPA is 2.97 and hence, higher than the mean high school GPA of all students in the dataset which is 2.50. Furthermore, the standard deviation of the final university grade is smaller than the standard deviation of the high school GPA. This indicates that compared to the grade of the high school leaving certificate, the distribution of the final university grade is compressed and shifted to the upper end of the grading scale.

With regard to the other covariates, we see that 47% of the students are male and 22% hold a private health insurance. The mean purchasing power index is 98.50, meaning that the mean purchasing power in our sample is 1.5% lower than the German average.

Taking a look at the distribution of students across faculties, we see that the highest

Table 4.1: Summary statistics

Variable	N	Mean	Std. Dev.	Min	Max
High school GPA	12315	2.50	0.63	1.10	4.00
Graduation (university)	12315	0.48	0.50	0.00	1.00
Graduation (within faculty)	16931	0.49	0.50	0.00	1.00
Final grade	8204	2.97	0.59	1.00	4.00
Male	12315	0.47	0.50	0.00	1.00
Private health insurance	12315	0.22	0.42	0.00	1.00
Purchasing power index	12315	98.50	11.79	64.72	186.99
Theology	16931	0.02	0.13	0.00	1.00
Law	16931	0.07	0.26	0.00	1.00
Medicine	16931	0.09	0.28	0.00	1.00
Humanities	16931	0.20	0.40	0.00	1.00
Mathematics	16931	0.04	0.19	0.00	1.00
Physics	16931	0.03	0.18	0.00	1.00
Chemistry	16931	0.04	0.19	0.00	1.00
Geology/Geography	16931	0.03	0.18	0.00	1.00
Biology	16931	0.08	0.28	0.00	1.00
Forest sciences	16931	0.04	0.19	0.00	1.00
Agriculture	16931	0.09	0.29	0.00	1.00
Economic sciences	16931	0.16	0.37	0.00	1.00
Social sciences	16931	0.11	0.31	0.00	1.00

Grades transformed to 1-4 Scale, with 4 being the best grade and 1 being the worst grade that is still a pass.

share of students is studying at the faculty of humanities (20%). Theology, on the other hand, is the smallest faculty with a share of 2%.

4.3.3 Empirical Setup

We start by examining the broadest measure of academic success, namely, whether or not a student graduates from university at all. Afterwards, we narrow our view towards graduation within fields, considering a change of field as a failure in the abandoned subject. Finally, we focus on the final grade of the university degree. This grade is a measure of the relative success within the group of successful students completing their studies.

For each of the three outcome variables we start with the GPA achieved at high school as independent variable only and continue by adding the full set of controls. These also include indicator variables for all 16 German states excluding Lower Saxony, the state where Göttingen is located, so as to reflect potential differences between the states concerning schooling systems and grading standards. Afterwards, we allow for differing effects by faculties. The binary outcome, graduation, is analyzed using probit models. For the continuous outcome variable, university grade, we use simple OLS models. In all the regressions we cluster standard errors by administrative district.

In order to interpret the regression results of the probit models right away, we display marginal effects for a benchmark student.[17] For categorical variables the effects are calculated as discrete changes from the base category. Our benchmark student is characterized by the average high school leaving grade and income, and the mode of categorical variables. Accordingly, the student is female, holds a public health insurance and finished high school in Lower Saxony.

[17]The coefficients of the probit regressions can be found in Tables 4.A.1-4.A.3.b in Appendix 4.A.II.

4.4 Results

There is a strong ex ante expectation that the better the high school leaving grade is, the better the performance at university should be. High income as well as a private health insurance status are expected to have positive effects on academic success. Low family income, proxied by the purchasing power index, might inhibit academic success through channels different from performance in high school. Students from low income families might lack sufficient monetary support and thus have to earn their living expenses outside university, such as working in bars, shops or factories, and thus would have less time to study. They might be less able to buy books that are not (numerous) in the libraries or other auxiliary devices such as software packages. However, payments according to the Federal Training Assistance Act (*BAföG*) should at least partly counteract this effect by providing financial support for students from poorer families.[18] We do not have a clear ex ante expectation about the influence of gender and the different faculties.

4.4.1 University Level

Table 4.2 shows the expected highly significant and positive effect of the high school leaving grade on academic success. A marginal improvement of this grade increases the probability of the benchmark student to graduate at all from university by about 21 percentage points per grade, and within fields by about 16 percentage points. An improvement of the high school leaving certificate by one full grade is associated with an improvement of the expected final grade by slightly below 0.4 grades.

The controls are of lesser importance: All else being equal, coming from a family that provides a student with private health insurance increases the estimated probability of the benchmark student of graduating at all or within a faculty by 5 or 4 percentage points respectively. This effect is highly significant but relatively small: Being privately insured raises the graduation probability by as much as having a 0.25 better grade at

[18]These payments are based on the income of the parents and the student. They can amount to up to 670 Euro per month (2010) of which only 50% are to be repaid, capped at a maximum amount due of 10,000 Euro. In winter term 2009/2010 almost 20% of all students in Göttingen received payments according to this act.

Table 4.2: University level

	Graduation -All Faculties- Probit		Graduation -Within Faculty- Probit		Final Grade OLS	
	(1)	(2)	(3)	(4)	(5)	(6)
High school GPA	0.210***	0.210***	0.165***	0.161***	0.371***	0.386***
	(28.121)	(28.444)	(21.810)	(26.022)	(0.010)	(0.010)
Male		-0.006		-0.009		-0.019
		(-0.548)		(-1.077)		(0.014)
Private health insurance		0.053***		0.036***		0.014
		(4.825)		(3.826)		(0.015)
Purchasing power index		0.001		0.000		0.001
		(0.748)		(0.423)		(0.001)
Constant					1.986***	1.902***
					(0.027)	(0.070)
States included	No	Yes	No	Yes	No	Yes
Pseudo-R^2	0.048	0.051	0.031	0.033		
Log Likelihood	-8120	-8093	-11368	-11338		
R^2					0.155	0.169
Observations	12315	12315	16931	16931	8204	8204

Columns 1-4: marginal effects for benchmark student, z-statistic in parentheses; columns 5-6: coefficients, standard errors in parentheses; clustered by counties; $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

high school. Conditional on graduating, there is no significant effect of the health insurance on the final grade.

The income variable does not show significant effects in any of the regressions presented in Table 4.2. This might indicate that financial aid, provided according to the Federal Training Assistance Act, is performing well. It could also mean that income alone is not very important for academic success if aspects such as the educational family background, as captured by the health insurance status, are accounted for. Another explanation could be that those who are negatively affected by their low family income have never even started university education in the first place.

Finally, the higher importance of the high school leaving GPA with respect to overall graduation compared to graduation within a field might indicate that being a good (high school) student does not help to find the most preferred field of study right away. Obviously, re-orientation at an early stage of the studies towards a field that fits the student's own preferences or abilities better should not be seen as severe as an overall failure to graduate. This is especially true with respect to international comparisons. For instance in the U.S. a major might be chosen only after trying several fields whereas in Germany students select their field prior to entering university.

4.4.2 Faculties

Some students change their field of study while being enrolled. This might reflect some change in their preferences or time needed to search for the perfect match. At the same time it might also reflect differences in the (perceived) degree of difficulty to graduate or to get a good grade. Every now and then a discussion arises in Germany about whether or not some faculties give good grades too easily. The faculties in question will usually defend themselves by pointing out the high ability of their student body (see for instance Krass and Scherf, 2012). In order to address this issue, we allow for differing effects by faculties. Firstly, we add indicator variables for the 13 faculties excluding the base category/faculty, humanities. Afterwards we present separate regressions for each of the faculties.

Column (1) of Table 4.3 shows marginal effects for a probit regression, estimating the probability of graduation, for the benchmark student. Column (2) presents

Table 4.3: Faculties

	Graduation Probit (1)	Final Grade OLS (2)
High school GPA	0.190***	0.373***
	(25.212)	(0.011)
Male	-0.016*	0.049***
	(-2.000)	(0.012)
Private health insurance	0.047***	0.023*
	(5.040)	(0.011)
Purchasing power index	0.000	0.001*
	(0.638)	(0.001)
Theology	-0.073**	-0.648***
	(-2.580)	(0.086)
Law	-0.004	-1.164***
	(-0.233)	(0.024)
Medicine	0.075**	-0.267***
	(2.989)	(0.024)
Mathematics	-0.060***	-0.123***
	(-3.504)	(0.030)
Physics	-0.059**	0.020
	(-2.820)	(0.029)
Chemistry	-0.020	0.042
	(-0.946)	(0.033)
Geology/Geography	0.103***	0.073*
	(4.959)	(0.032)
Biology	0.119***	0.063***
	(7.784)	(0.019)
Forest sciences	0.283***	-0.327***
	(13.612)	(0.027)
Agriculture	0.259***	-0.204***
	(15.286)	(0.021)
Economic sciences	0.185***	-0.414***
	(12.445)	(0.018)
Social sciences	0.066***	0.001
	(4.541)	(0.019)
Constant		2.056***
		(0.065)
States included	Yes	Yes
Pseudo R^2	0.062	
Log Likelihood	-11005	
R^2		0.423
Observations	16931	8204

Column 1: marginal effects for benchmark student, z-statistics in parentheses; column 2: coefficients, standard errors in parentheses; clustered by county; $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

corresponding OLS results for the final university grade given graduation.

Many indicator variables of faculties show effects that are significant at the 0.1 percent level. For the benchmark student the predicted probability of graduating, given she started studying at the faculty of humanities, is about 39%; given successful graduation, her expected final grade is 3.1. A male student is almost 2 percentage points less likely to graduate within the given faculty compared to the benchmark. Ceteris paribus, if he does, he receives slightly better grades. The private health insurance status is associated with both better grades and a higher probability of graduating.

All else being equal, the predicted probability of graduating at the faculty of economic sciences is about 19 percentage points higher than at the faculty of humanities; at the faculty of mathematics it is 6 percentage points lower than at the base faculty. Given graduation, the faculty of economic sciences awards, ceteris paribus, a final grade that is more than 0.4 grades worse than the respective grade at the faculty of humanities. This difference is greater than the expected change in the degree associated with an improvement of the high school leaving certificate by one full grade. The worst grades are awarded by the faculty of law.[19]

Doing the same regressions separately by faculties, the picture gets more differentiated. Tables 4.4.a and 4.4.b reveal strong differences with respect to how important the high school GPA is for the probability of graduating at the different faculties of Göttingen University. The effect is not significantly different from zero at the faculty of geology and geography, and it is strongest at the medical school and the faculty of chemistry. For the benchmark student at these two faculties, a marginal increase in the GPA earned in high school is associated with an increase in the graduation probability by almost 29 percentage points per grade. At the faculty of social sciences, the effect is only about one third of that size.

Private health insurance status, which proxies a high socio-economic background, is significant and has a positive sign for about half of the faculties, while being insignificant for the other faculties. Purchasing power is also of little importance for the probability of graduating at the faculty level. It is significant only at the faculty of social sciences.

For illustration and further comparison of faculties, Table 4.5 provides predicted

[19]The faculty of law is traditionally known to only rarely award very good grades. Accordingly, not too much attention should be given to this fact.

Table 4.4

Table 4.4.a: Graduation by faculties

	Graduation by Faculties						
	Theology	Law	Medicine	Humanities	Mathematics	Physics	Chemistry
High school GPA	0.180***	0.256***	0.285***	0.187***	0.279***	0.209***	0.285***
	(4.558)	(11.171)	(9.357)	(12.378)	(6.412)	(7.185)	(9.016)
Male	0.112	0.007	0.019	-0.114***	0.060	0.110*	0.043
	(1.789)	(0.231)	(0.685)	(-6.688)	(1.677)	(2.257)	(1.004)
Private health insurance	0.184*	0.019	0.080**	0.068***	0.131*	-0.013	0.011
	(2.507)	(0.611)	(2.923)	(3.541)	(2.464)	(-0.310)	(0.253)
Purchasing power index	0.002	-0.000	-0.001	0.002	-0.001	-0.000	0.001
	(0.633)	(-0.031)	(-0.729)	(1.522)	(-0.575)	(-0.246)	(0.349)
States included	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.109	0.076	0.125	0.059	0.164	0.111	0.137
Log Likelihood	-167	-774	-896	-2128	-367	-345	-378
Observations	284	1246	1481	3342	660	567	644

Marginal effects for benchmark student, z-statistics in parentheses; clustered by county; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 4.4.b: Graduation by faculties

	Graduation by Faculties					
	Geology/Geography	Biology	Forest Sciences	Agriculture	Economic Sciences	Social Sciences
High school GPA	0.069	0.176***	0.152***	0.132***	0.159***	0.086***
	(1.875)	(8.304)	(3.971)	(5.451)	(8.061)	(4.521)
Male	-0.127*	-0.016	0.031	0.049	-0.022	-0.027
	(-2.151)	(-0.587)	(0.651)	(1.471)	(-1.071)	(-1.225)
Private health insurance	0.061	0.037	0.040	-0.038	0.064**	0.011
	(1.070)	(1.113)	(1.056)	(-1.071)	(3.110)	(0.388)
Purchasing power index	0.004	-0.002	-0.000	-0.002	0.001	0.004***
	(1.436)	(-1.826)	(-0.029)	(-1.390)	(0.553)	(3.340)
States included	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.039	0.047	0.043	0.024	0.032	0.019
Log Likelihood	-360	-923	-425	-1004	-1819	-1198
Observations	542	1410	666	1546	2740	1778

Marginal effects for benchmark student, z-statistics in parentheses; clustered by county; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

probabilities of graduation based on the estimation results underlying Tables 4.4.a and 4.4.b. The predictions for the benchmark student are presented in the middle column (mean high school GPA). The remaining predictions deviate from the usual benchmark by the high school GPA used. We define low and high high school GPA as the mean GPA minus two standard deviations and mean GPA plus two standard deviations respectively.

Although we do not want to put too much emphasis on these predictions, they serve to illustrate the rather large differences between faculties. The predicted probability of graduation for the benchmark student is between roughly 20 and 60 percent. Based on these predictions, a student with a low high school GPA can hardly expect to graduate at some of the faculties, such as mathematics and physics. At other faculties chances to graduate are still relatively high; the predicted probabilities for such a student are 45 and 39 percent at the faculties of agriculture and economic sciences respectively. For an otherwise identical student with a high high school GPA the predictions vary between about 50 and 80 percent.

Tables 4.6.a and 4.6.b show corresponding regression results for final grades at graduation. There is a highly significant positive effect of the high school GPA at every faculty. However, the importance of this GPA differs strongly. It is highest at the faculty of mathematics, where the expected grade at graduation is more than half a grade better for every full grade of the high school leaving certificate. At the faculty of chemistry, where the coefficient of high school GPA is the smallest, the effect is only about half that size. Given graduation, male students can expect slightly better grades than their female fellow students in about half of the faculties. The effects of health insurance status and purchasing power are indistinguishable from zero at most faculties.[20]

Figure 4.1 visualizes the relationship between the GPA earned at university and at high school across selected faculties. The red lines represent fitted values for female students who are publicly insured, come from a zip code area with average purchasing

[20]There is a surprisingly large, highly significant, positive effect of the private health insurance status on the final grade at university at the faculty of Theology. Taking this coefficient at face value, a reason for this strong effect could be that children of pastors in Germany are privately insured. However, due to the small sample size of the underlying regression, we refrain from emphasizing this finding.

Table 4.5: Predicted probabilities of graduation by faculties

	High School GPA		
	Low	Mean	High
Theology	0.10	0.27	0.53
Law	0.14	0.40	0.72
Medicine	0.19	0.52	0.83
Humanities	0.21	0.42	0.66
Mathematics	0.04	0.24	0.67
Physics	0.05	0.21	0.54
Chemistry	0.06	0.30	0.69
Geology/Geography	0.41	0.50	0.59
Biology	0.30	0.51	0.72
Forest sciences	0.38	0.57	0.75
Agriculture	0.45	0.62	0.77
Economic sciences	0.39	0.59	0.77
Social sciences	0.34	0.45	0.56

Predicted probability of graduating at a faculty for female students who are publicly insured, come from a zip code area with average purchasing power, and finished high school in Lower Saxony. Low and high high school GPA are defined as the mean GPA minus two standard deviations and mean GPA plus two standard deviations, respectively.

Table 4.6

Table 4.6.a: Grades by faculties

	Final Grade by Faculties						
	Theology	Law	Medicine	Humanities	Mathematics	Physics	Chemistry
High school GPA	0.434**	0.428***	0.279***	0.393***	0.503***	0.291***	0.270***
	(0.157)	(0.030)	(0.044)	(0.019)	(0.043)	(0.051)	(0.052)
Male	-0.078	0.090*	-0.066	0.080**	0.150*	0.166*	0.099
	(0.208)	(0.038)	(0.050)	(0.024)	(0.068)	(0.067)	(0.065)
Private health insurance	0.536***	0.016	0.053	0.036	0.081	0.018	-0.052
	(0.146)	(0.052)	(0.049)	(0.023)	(0.062)	(0.059)	(0.059)
Purchasing power index	-0.019*	0.001	0.005**	0.002	0.004	-0.002	-0.004
	(0.008)	(0.003)	(0.002)	(0.001)	(0.003)	(0.002)	(0.003)
Constant	3.125***	0.743**	1.739***	1.971***	1.177**	2.551***	2.957***
	(0.797)	(0.278)	(0.210)	(0.128)	(0.371)	(0.287)	(0.312)
States included	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.324	0.234	0.097	0.277	0.421	0.184	0.171
Observations	86	502	776	1365	253	249	270

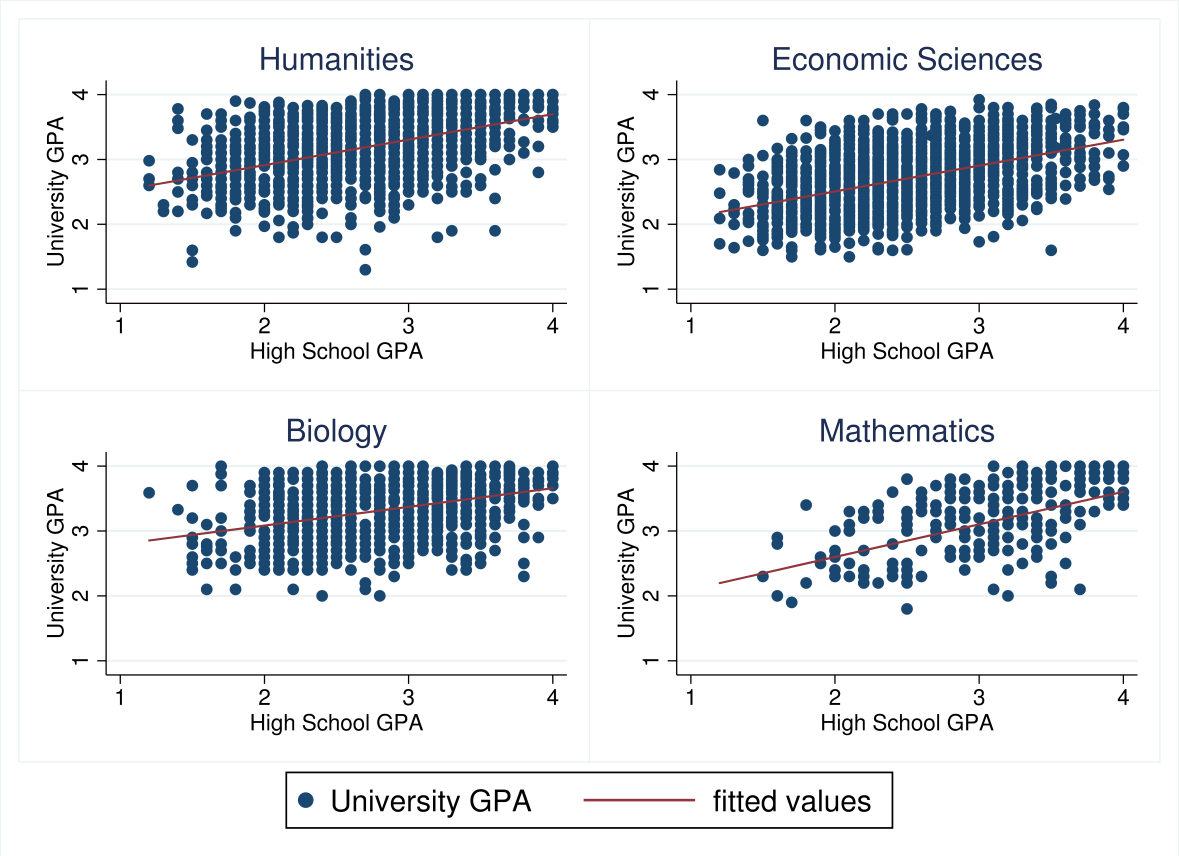
Coefficients, standard errors in parentheses; clustered by county; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 4.6.b: Grades by faculties

	Final Grade by Faculties					
	Geology/Geography	Biology	Forest Sciences	Agriculture	Economic Sciences	Social Sciences
High school GPA	0.293***	0.288***	0.352***	0.386***	0.398***	0.398***
	(0.057)	(0.029)	(0.041)	(0.024)	(0.019)	(0.026)
Male	-0.064	0.113***	0.116**	0.018	0.017	0.054
	(0.051)	(0.033)	(0.043)	(0.029)	(0.022)	(0.027)
Private health insurance	0.011	0.013	0.041	-0.087*	0.014	0.050
	(0.053)	(0.030)	(0.053)	(0.037)	(0.035)	(0.035)
Purchasing power index	-0.003	0.002	0.003	-0.000	0.001	0.000
	(0.003)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)
Constant	2.807***	2.339***	1.573***	2.030***	1.575***	2.110***
	(0.335)	(0.159)	(0.250)	(0.171)	(0.129)	(0.171)
States included	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.212	0.158	0.184	0.232	0.247	0.250
Observations	250	784	408	953	1534	774

Coefficients, standard errors in parentheses; clustered by county; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

power and finished high school in Lower Saxony. We can notice from the upper two panels of this figure that grades in humanities are generally better than in economic sciences. The lower two panels show that the relationship between high school GPA and university grade is much steeper in mathematics than in biology.



Dots represent one or several observations. Fitted values are the predicted university GPA for female students who are publicly insured, come from a zip code area with average purchasing power, and finished high school in Lower Saxony.

Figure 4.1: Grades at selected faculties

Comparing the faculties with the highest number of students, humanities and economic sciences, it seems to be easier to graduate in economic sciences whereas the expected grade conditional on graduation is worse. This pattern can also be found for a couple of other faculties and might suggest differences in grading and examination culture between the faculties. It seems that at some faculties it is more difficult to obtain

a degree while the grades given differentiate less strongly between students. However, at others achieving a degree is more likely while the grades obtained vary more within the grading scale.

There are a number of possible mechanisms which might contribute to these faculty-specific results. Firstly, students may self-select into faculties on unobservable characteristics related to the outcome variables. For example, some students may be more motivated to obtain good grades at university than they were in high school. If such students disproportionately choose humanities rather than mathematics or economics, we will find better grades in the former faculty conditional on high school GPA. While we cannot exclude such self-selection with the data at hand, in our view it is not very plausible that students of various faculties should differ precisely in this respect.

Alternatively, and arguably more convincingly, the results may be driven by features of the teaching and grading system in the respective faculties. A first explanation along this line is based on the similarity between curricula in high school and in university. The high school grade is a composite of a comprehensive variety of subjects whereas university studies are more specialized. Since students likely choose subjects which fit their specific abilities, one may expect that in highly specialized fields, university grades are better and less closely associated with high school GPA than in broader subjects. Given that the impact of high school GPA on university grades is largest in mathematics, which is a more specialized field than social science or economics, this explanation, however, does not find much support in the data.

Instead, the differences in grades are likely to reflect different grading cultures. Some faculties may simply be willing to award good grades to most students without differentiating strongly among good and mediocre performance. More subtly, an upward drift of average grades may be built in the structure of some degree programs. When a program grants ample choice among electives, students can avoid difficult or unpleasant courses while still obtaining the degree. Moreover, if students can freely choose courses, teachers might have an incentive to attract students by grading leniently. As a result, grades from such a program will be compressed at the upper end of the scale compared to programs with a more rigid structure of compulsory courses.

Although we have some sympathy for the last explanation, our data do not permit

to conclusively distinguish between these mechanisms. Instead, we confine ourselves to pointing out the main result of this paper: The relationship between high school grades and university success varies in a statistically discernible manner among faculties, which hints at some differences in grading, teaching, and examination cultures.

4.5 Discussion and Policy Implications

In this paper, the determinants of studying successfully are analyzed using data from more than 12,000 students from Göttingen University. Two main results are shown. Firstly, the high school leaving grade is by far the best predictor of both the probability of graduating and the final grade obtained at university. Other factors, notably gender or social origin, play only a minor role. Secondly, differences emerge among the various faculties regarding grading and graduation policies. In some faculties, like humanities or social sciences, the rate of graduation is low but those who graduate can expect to obtain quite good grades even when they start from a weaker academic base as measured by the high school GPA. In other faculties, such as economic sciences or forest sciences, the chance of obtaining a degree is relatively high whereas grades are moderate, and strongly linked to high school GPA. Finally, in some faculties such as mathematics and physics, graduation appears to be very difficult and good grades are hard to obtain, especially for weaker students.

These findings carry a number of implications both for the university and for the students individually as well as for education policy in general. Most obviously, our results support the current process of admission to German universities, which is based primarily on high school GPA. Clearly, this practice contributes to improving the academic success of those admitted. We do not find any evidence that adding other information can improve the selection. Specifically, variables capturing income or social background have a comparatively low explanatory power. This suggests that students from disadvantaged social backgrounds do not, on average, have abilities relevant for success at university which are undocumented by high school grades. Consequently, granting privileged access to minorities or providing universities with financial incentives to admit more students from poor districts can be a useful policy to raise equity in higher education, but will not enhance the overall quality of the students. It appears

that most of the impact of social origin on university achievement is already absorbed in the high school leaving grade. Consequently, policy should start addressing social imbalances in educational outcomes at earlier stages of the academic career.

For prospective students, the faculty specific results, summarized in Table 4.5, may give useful hints about what subject to choose. A student with mean high school GPA has a higher chance of graduating if she chooses agriculture or economic sciences rather than humanities or social sciences. If obtaining some degree irrespective of the field is very important for her, such a student should enroll in the former rather than in the latter faculties. Considering mathematics, physics, or chemistry, the recommendation is even clearer: The average student will graduate in these faculties with a probability of 30% or less. For weaker students with high school GPA substantially below the mean this probability falls below 10%. This suggests that these three fields are almost unfeasible for students in the bottom half of the ability distribution and that such students are well advised to opt for other fields.

Extending the principle of selection on academic merit to the aggregate level obviously raises a consistency issue: Not every university or field can be restricted to the best students, since the weaker ones also will have to be placed somewhere, or else must be told not to study. This points out a basic choice which education policy must make: Should universities provide an excellent education for the most able individuals at a level defined by the current state of knowledge, or should tertiary education be targeted to large numbers of students and settle for an academic level accessible for these? Related to this, there are competing views on the main purpose of university studies. On one hand, in Humboldt's tradition, one may see academic studies mainly as a tool of personal intellectual enhancement, where knowledge, understanding and academic debate are rewards in themselves. On the other hand, studies may be seen as an investment in productivity, whose main reward comes in the form of a higher wage. In the former view, graduation and examination grades are of lesser importance. In the latter case, the signaling value of a degree is likely to be essential for employers. As a consequence, the labor market will honor only completed degrees, and a wage premium will be paid for good grades as long as these are rare enough so as to convey credible information.

The results presented in this paper suggest that faculties take different sides in this debate. In humanities, graduation rates are relatively low and individual grades are less differentiated than in other fields. This corresponds to the idea that one does not study for the sake of the examination or for a higher wage, but for intrinsic motivation. Quite possibly these fields specifically attract students with such expectations. In this view, a low completion rate in such subjects should not be seen as a sign of failure. These fields offer students an education tailored to their abilities and preferences and students use this offer to the extent which is individually optimal. On the other end of the scale, examinations in mathematics, physics and chemistry are highly selective. Thereby, these fields cannot cater to large numbers of students, but they prepare those who make it for demanding sections of the labor market. Similarly, economic sciences serve the labor market by awarding differentiated grades while still being accessible for large numbers of weaker students.

These considerations shed some light on the recommendation, repeatedly voiced by the OECD (see for instance OECD, 2013, p. 151), that Germany should produce more university graduates and the corresponding complaint by employers' organizations that German industry faces a shortage of graduates from mathematics, natural sciences, and engineering (see Anger et al., 2013). It is certainly conceivable that reforms in secondary schooling can raise the number of students entering university. It appears far-fetched, however, that a large fraction of those additional students will display academic abilities superior to those of the average current student. Our results show that average or below average students will typically be unable to successfully complete a degree in mathematics, physics or chemistry. Therefore it seems highly unlikely that an increase in university enrollment will produce substantial numbers of additional graduates in the subjects required by industry, at least as long as the concerned faculties are unwilling to lower their academic standards. If this does not occur, any increase in university enrollment will lead to larger numbers of graduates in those fields which cater to the preferences and abilities of the majority of students but not in those fields which firms demand.

Appendix to Chapter 4

Appendix 4.A.I: Data Processing

We exclude students for whom not all information is available as well as students for whom we observe pure data errors, such as when the grade of the high school leaving certificate is not within the possible interval. Ph.D. students are also dropped from the dataset. The reason for this is that they form a highly selective group and their success may be influenced by other factors than regular students' performance. Furthermore, we only take into account students who either finished university with a degree or dropped out of their study program. Since students are asked to give the reason for dropping out when they leave university, we can distinguish between real drop outs and students who intend to continue their studies at another university. We exclude these students from the sample in order not to register a drop out for the latter group.

As German and foreign high school leaving grades may not be comparable and university success of students with a foreign educational background may be influenced by additional factors such as language skills, we only take into account students who hold a German high school leaving certificate. In addition, we exclude students with a high school leaving grade of 4.0, the worst grade still allowing a student to pass. This is done as in our dataset a high school leaving grade of 4.0 was often found for students, in particular for foreign students, who enrolled in fields of study without admission restriction. This strongly suggests that the grade is sometimes used as a place holder when the real grade seemed not to be important for the admission procedure. However, we are confident that we have only deleted a very small number of students who actually have a high school leaving grade of 4.0 by imposing this restriction.

In addition, students have to provide information about their home address, usually their parents' address, and their semester address, usually the place students live by themselves. Since most students move to Göttingen when starting university, home and semester address should differ. Nonetheless, for some students in our dataset the two zip-codes are identical. As we make use of the parents' address in our analysis it is important that the correct zip-code is used. To deal with this problem, we look at all students for whom the zip-code of their home and semester address are the same. If both zip-codes belong to a place outside of Göttingen, it is very likely that this student

is still living with her parents. If the zip-codes are identical and from Göttingen, it might be that the student did not provide any information about her parents' home address. Therefore, we take a look at the administrative district the student went to school in. If she graduated from a high school in Göttingen, we have no reason to doubt that her parents also live there. On the other hand, if she went to school outside of Göttingen, it is not entirely clear that the information about the home address really corresponds to the parental address. Consequently, we exclude these students from the sample.

Appendix 4.A.II: Coefficients

Table 4.A.1: University level - Coefficients for Table 4.2

	Graduation -All Faculties- Probit		Graduation -Within Faculty- Probit	
	(1)	(2)	(3)	(4)
High school GPA	0.528***	0.527***	0.414***	0.405***
	(0.019)	(0.018)	(0.019)	(0.016)
Male		-0.014		-0.022
		(0.025)		(0.021)
Private health insurance		0.134***		0.091***
		(0.028)		(0.024)
Purchasing power index		0.002		0.001
		(0.002)		(0.002)
Constant	-1.359***	-1.513***	-1.076***	-1.142***
	(0.048)	(0.271)	(0.079)	(0.240)
States included	No	Yes	No	Yes
Pseudo-R^2	0.048	0.051	0.031	0.033
Log Likelihood	-8120	-8093	-11368	-11338
Observations	12315	12315	16931	16931

Coefficients, standard errors in parentheses; clustered by counties; $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 4.A.2: Faculties - Coefficients for Table 4.3

	Graduation Probit (1)
High school GPA	0.493***
	(0.016)
Male	-0.040*
	(0.020)
Private health insurance	0.119***
	(0.023)
Purchasing power index	0.001
	(0.002)
Theology	-0.196*
	(0.080)
Law	-0.009
	(0.039)
Medicine	0.191**
	(0.065)
Humanities	base
Mathematics	-0.160***
	(0.047)
Physics	-0.157**
	(0.057)
Chemistry	-0.052
	(0.055)
Geology/Geography	0.261***
	(0.052)
Biology	0.302***
	(0.039)
Forest sciences	0.730***
	(0.058)
Agriculture	0.663***
	(0.046)
Economic sciences	0.467***
	(0.038)
Social sciences	0.167***
	(0.037)
Constant	-1.587***
	(0.216)
States included	Yes
Pseudo R^2	0.062
Log Likelihood	-11005
Observations	16931

Coefficients, standard errors in parentheses; clustered by county; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 4.A.3

Table 4.A.3.a: Graduation by faculties - Coefficients for Table 4.4.a

	Graduation by Faculties						
	Theology	Law	Medicine	Humanities	Mathematics	Physics	Chemistry
High school GPA	0.539***	0.663***	0.714***	0.479***	0.889***	0.717***	0.820***
	(0.115)	(0.057)	(0.076)	(0.038)	(0.083)	(0.088)	(0.089)
Male	0.310	0.019	0.048	-0.304***	0.181	0.336*	0.121
	(0.167)	(0.084)	(0.070)	(0.047)	(0.112)	(0.150)	(0.120)
Private health insurance	0.496**	0.048	0.204**	0.173***	0.373**	-0.044	0.032
	(0.188)	(0.078)	(0.071)	(0.048)	(0.139)	(0.145)	(0.125)
Purchasing power index	0.006	-0.000	-0.003	0.005	-0.003	-0.001	0.002
	(0.009)	(0.004)	(0.003)	(0.003)	(0.006)	(0.006)	(0.005)
Constant	-2.518**	-1.900***	-1.496***	-1.848***	-2.571***	-2.438***	-2.742***
	(0.953)	(0.419)	(0.406)	(0.364)	(0.623)	(0.666)	(0.551)
States included	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.109	0.076	0.125	0.059	0.164	0.111	0.137
Log Likelihood	-167	-774	-896	-2128	-367	-345	-378
Observations	284	1246	1481	3342	660	567	644

Coefficients, standard errors in parentheses; clustered by county; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 4.A.3.b: Graduation by Faculties - Coefficients of Table 4.4.b

	Graduation by Faculties					
	Geology/Geography	Biology	Forest Sciences	Agriculture	Economic Sciences	Social Sciences
High school GPA	0.172	0.441***	0.388***	0.346***	0.410***	0.218***
	(0.092)	(0.053)	(0.104)	(0.068)	(0.049)	(0.048)
Male	-0.325*	-0.040	0.081	0.132	-0.057	-0.069
	(0.151)	(0.069)	(0.123)	(0.090)	(0.053)	(0.056)
Private health insurance	0.154	0.093	0.102	-0.098	0.168**	0.028
	(0.145)	(0.083)	(0.097)	(0.092)	(0.055)	(0.072)
Purchasing power index	0.010	-0.006	-0.000	-0.005	0.002	0.009***
	(0.007)	(0.003)	(0.005)	(0.004)	(0.003)	(0.003)
Constant	-1.375	-0.472	-0.772	-0.031	-0.971**	-1.589***
	(0.715)	(0.369)	(0.546)	(0.423)	(0.374)	(0.308)
States included	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.039	0.047	0.043	0.024	0.032	0.019
Log Likelihood	-360	-923	-425	-1004	-1819	-1198
Observations	542	1410	666	1546	2740	1778

Coefficients, standard errors in parentheses; clustered by county; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Chapter 5

"A is the Aim?"*

5.1 Introduction

Academic grades are said to reflect students' achievement and thereby the effectiveness of educational institutions and their accountability to potential employers. However, in the past decades, confidence in the reliability of the grades has been badly shaken by studies exposing the trend towards grade inflation. Several studies have shown that rise in grades has become an issue in both secondary and tertiary education across many countries, thereby stressing the need to provide explanations for the phenomenon of grade inflation. Rojstaczer and Healy (2010) conducted a study of grading patterns in more than 160 American colleges and universities and found a nationwide rise in average grades of nearly a tenth of a point change per decade, with A being the most commonly awarded grade at American colleges and universities.

This paper addresses this concern by focusing on the differences in grading policies between professors assigned to the same mandatory first-year course in Economics. Firstly, the analysis reveals that there are huge differences in grading even within the same course. Secondly, the effect of having a certain professor in the mandatory first-year course on student's later performance is highly significant and cannot be solely explained by differences in professors' grading. However, the sign of this effect is ambiguous, and depends on the mathematical rigor of the course and the examination style. Furthermore, the results demonstrate a highly significant effect of having the same professor for many classes, although switching from a tough to an easy grader seems to be the best strategy for improving grades. Our analysis shows that the obtained effects are quite meaningful. All else being equal, having a certain professor in Microeconomics I is associated with an improvement of the expected grade in a follow-

*See Danilowicz-Gösele (2016).

on course by up to 1.385 grades. The overall result indicates that both grading policies and learning outcomes vary between professors within the same course.

When talking about grade inflation, it is important to distinguish between awarding higher grades *per se* and improvement in grades as a result of better performing students who are learning more and/or being taught better. The difficulty with the latter explanation lies in the fact that a growing number of educational researchers claim that it is raining A's in the education system without continuous evidence of increasing academic performance.

Strong evidence that students are indeed doing worse today relative to a decade ago is provided by researchers from the National Center for Education and Statistics in 2015, who claim that SAT scores in critical reading, writing and math have dropped each year within the analyzed period from 2004 till 2012 (Kena et al., 2015). Moreover, according to the report published by the National Bureau of Economic Research of the University of California, students decreased their class and studying time from 40 hours per week in 1961 to 27 hours per week in 2003 (Babcock and Marks, 2011). Since concerns about grade inflation are not new, researchers have already offered many explanations for the upward trend in grades. Some of these explanations focus on changes in educational institutions including changes in enrollment patterns (Prather and Kodras, 1979), curricula (Prather and Kodras, 1979) or grading policies (Birnbaum, 1977).

Although the teaching body is expected to individually regulate grading policies, because of their effect on the reputation of the institutions they work for and students' career chances, it is commonly known that unregulated institutions are very often challenged to maintain certain continuous standards. This is also the case for the colleges and universities, which have troubles maintaining academic standards in the absence of any regulation. Against this background, educational researchers found out that teachers' characteristics such as teaching quality (De Paola, 2009), teaching experience (Rivkin et al., 2005; Clotfelter et al., 2007), gender (Neumark and Gardecki, 1998; Bettinger and Long, 2005; Carrell et al., 2010) and age (Hong, 1984), have statistically significant effects on students' grades. However, there is less agreement on the influence of the instructor rank (Sonner, 2000), part-time or full-time status or salary (Nelson and

Lynch, 1984; Pressman, 2007; Hoffmann and Oreopoulos, 2009). In addition, there is strong evidence that changes in the use of student evaluations (Krautmann and Sander, 1999; Stratton et al., 1994; Johnson, 2003; Eiszler, 2002) or the public availability of median grades (Bar et al., 2009) may also influence the extent to which instructors exaggerate students' grades.

Other kinds of explanations draw attention to changes in students' behavior including students' freedom in choosing departments, courses or certain professors. At many universities, there is a visible trend towards learning that is more relevant to students' interests and goals. Therefore, today's students have much more freedom in designing their study paths, being able to choose from a wide range of major/elective courses dependent on their interest, abilities, difficulty level, instructor, work load or examination structure. In some cases, they can even decide whether the grade from a taken class will appear on their transcript of records or not. Thus, in order to improve the overall grade, students may act strategically by taking advantage of the mentioned differences, which will result in attending carefully-selected courses or in opting for non-visible grades.

The study of Sabot and Wakeman-Linn (1991) shows that students are significantly more likely to enroll in a subsequent course of a department where they have already received a relatively higher grade. Another finding of their study is that grades obtained in low-grading departments are better predictors of students' later performance than grades received from grade-deflating departments. A similar finding is reported in a study of Ost (2010), who found out that low grades in science classes can be used as a predictor for students' participation in subsequent science courses.

Given all this, it is not surprising that, in many countries, such as United States, Canada, England, Scotland and Wales, online professor rating sites, such as "Rate My Professors.com", become so popular. In this case, students have the possibility to rate their professors according to *easiness*, *helpfulness*, *clarity*, *hotness* and the rater's *interest* in the class, in order to help fellows to choose the appropriate classes and/or professors. Some studies, such as Miles and Sparks (2013), examined the effect of online professor ratings and found out, that such websites indeed have an influence on students' choices for selecting professors, however it is not very clear to what extent.

Although there is a vast amount of research on grade inflation, there is little at-

tention to grading differences within higher education, especially within the same field of study or within the same course. This paper contributes to this body of literature by assessing the effect of grading and teaching differences from a mandatory first-year course on students' performance in follow-on courses at a German University. Even though all professors assigned to the mandatory first-year course have a very similar teaching and examination style, and students in most cases follow the curriculum, there is no random placement of students into the classroom and thus we have to be aware of students' self-selection toward certain professors. For this reason, this paper proposes an instrumental variable (IV) strategy by instrumenting student's choice of a professor through a random assignment of professors, on the semester basis, to the mandatory first-year course. In this case, we follow the faculty's recommendation to write the exam in the second semester of studies. Therefore, taking the exam with a professor who was assigned to the course in the student's second semester will influence the student's later performance. On the contrary, the fact that a professor is assigned to the course in the student's second semester does not affect the student's later achievements if a student decides, against the faculty's recommendation, to write her exam in an earlier or later semester.

The paper is organized as follows: Section 5.2 provides a brief overview of the institutional background. The data set, variables used and the empirical framework are presented in Section 5.3. Section 5.4 presents the results and Section 5.5 concludes with summarizing the findings of the analysis.

5.2 Institutional Background

Since the aim of the paper is to study the differences in professors' grading and its effect on students' performance in follow-on classes, this paper chooses one of the mandatory first-year courses offered at the faculty of Economics, namely Microeconomics I. It is an introductory undergraduate course that teaches the fundamentals of microeconomics. The reasons for our choice are three-fold. Firstly, this course is mandatory for all students enrolled at the faculty of economic sciences at Göttingen University, thus providing us with extensive and diverse observations. Secondly, within the faculty of economic sciences, Microeconomics I is the course with the greatest number of

professors assigned, thus indicating variation according to their characteristics such as grading policy. Last but not least, since Microeconomics I is one of the first courses economic students are required to take in their undergraduate programs, it also serves as a prerequisite for many follow-on classes. Therefore, its accompanying effect of different grading policies, if existent, should be strong enough to be observed in students' later achievements.

Based on the explicit information provided on students' university records, we are able to restrict our analysis to students enrolled at the faculty of economic sciences who have participated in the Microeconomics I exam either once or multiple times but always with the same professor. Although professors' assignment to the Microeconomics I course is to a greater extent random and thus not known in advance to the students, we are still aware of the self-selection bias toward certain professors. The reason for this is that, students are, to some extent, free to choose when they want to take this course. The only restricting factor is the examination regulation, according to which the credits for this course must be earned by the end of the student's fifth semester. For this reason, they can postpone taking the course to a later semester or even after taking the class in their second semester they can still decide to drop-out of the exam. In fact, it might be tempting for some students to postpone the Microeconomics I exam until a professor, known for "easy grading", will offer it.

In order to control for students' potential self-selection bias, we follow the recommendation of the Economics faculty. Hence, in our analysis we will distinguish between students who followed faculty's recommendation and students who postponed the exam to a later date. In addition to this we could also think about other factors, such as illness or other circumstances beyond students' control, that lead to postponing the exam.

The follow-on courses that we decided to look at are Microeconomics II and Public Finance, the latter one combines the introductory and the advanced course which are both taught by the same professor. Both courses aim to further deepen the study in microeconomics, thus having Microeconomics I as a prerequisite. All examined courses have a similar structure and consist of lecture and tutorials on a weekly basis. The lectures are taught by one of the university professors, tutorials, however, are taught

either by scientific assistants (Public Finance) or trained students who already passed the respective course (Microeconomics I and II).

5.3 Data and Methodology

5.3.1 Dataset and Descriptives

Dataset

This paper employs a unique administrative data set of 2,920 students enrolled at the faculty of economic sciences of the Göttingen University who participated in the Microeconomics I exam between 2006 and 2011.[22] The detailed and anonymized data includes students' characteristics such as high school leaving degree, gender, type of health insurance or parental address, as well as students' university records such as chosen field of study, grades, examiners, attempts and examination dates.

Dependent Variables

Our outcome variables are students' grades from three undergraduate courses offered at the faculty of economic sciences, namely Microeconomics I, Microeconomics II and Public Finance.

The data set is restricted to students who either took the exam in Microeconomics I one time or multiple times but always with the same professor. In case of multiple examination attempts, we will use the grade a student received on her first examination attempt in this course, implying that every student in our data has only one grade in Microeconomics I. For the two other courses, Microeconomics II and Public Finance, we will use the best grade the student received in all her attempts. In order to make the results internationally comparable, German grading scale, with 1.0 as the best possible grade and 4.0 as the minimum passing grade, has been translated to the U.S. grading scheme.

[22] Göttingen is a small city in Lower Saxony (Germany) where students account for 22% of the city's population (Statistisches Bundesamt, 2009). According to the statistics of the University of Göttingen, there are currently more than 4,300 students enrolled at the Faculty of Economic Sciences, which amounts to around 15 percent of all students on campus.

Independent Variables

Within the analyzed period from 2006 to 2011, five different professors were assigned to teach the Microeconomics I course. To include these in our analysis, we create an indicator variable for each professor, reflecting the student's enrollment decision concerning Microeconomics I. For example, if a student took her first Microeconomics I exam with *Professor 1*, her indicator variable for this professor will equal one. Consequently, the remaining four indicator variables (*Professor 2*; *Professor 3*; *Professor 4*; *Professor 5*) will be zero.

Assuming that students can benefit from having the same professor in different subjects, we include one further indicator variable *Same Professor* to capture the effect of having the same professor in the two analyzed courses.

The high school grade point average (GPA) will be used as a control for students' ability. GPA serves as a predictor of academic success believing that this measure captures more than just the students' abilities or achievements but also certain behavioral factors. Likewise course grades, high school GPA is converted to the U.S. grading scale.

In the analysis, we also account for students' socio-economic background by using the students' health insurance type and the purchasing power index of her parents' zip-code area. The type of health insurance is suited as a proxy for students' educational and socio-economic background because of the organization of the German health care insurance system, which is characterized by the dual system of public and private health insurance. While almost everyone is eligible for public health insurance, the private health insurance can be chosen only due to certain income criteria or employment status.[23]. Since in most cases students are insured through their parents, their health insurance status also provides information about their educational background.

Additionally, the purchasing power index within the zip-code area of the students' home address serves as another compelling measure of students' socio-economic background, considering that the German zip-code areas are fairly small and assuming that there is some residential sorting due to income. It relates to the per-capita income of a zip-code area with the average per-capita income of Germany, thus expressing the pur-

[23]In 2008 the number of people being privately insured (civil servants, self-employed or high earners) with a university degree was almost three times as high as that within the total German population (Finkenstädt and Keßler, 2012; Statistisches Bundesamt, 2009)

chasing power of a region. The index is normalized to 100, meaning that an area with an index value of 110 has a purchasing power of greater than ten percent as compared to the German average.

Furthermore, since there is a lack of agreement in educational research about gender differences we include gender as a control variable in our analysis. Gender is measured as a dichotomous variable coded as 1 for female and 0 for male.

Summary Statistics

The summary statistics in Table 5.1 show that the number of observations in the two analyzed follow-on courses, Microeconomics II and Public Finance, does not equal our sample size in Microeconomics I. This is due to the fact that study and examination regulations vary among degree programs at the faculty of economic sciences, meaning that some courses are not necessarily required as part of undergraduate studies. For Microeconomics I, the mean-value for a sample of 2,920 students is 2.12. For Microeconomics II, the mean-value for a sample of 1,255 students is 2.03 and thus only slightly lower than the mean-value of Microeconomics I. The highest mean-value of 2.33 appeared for a sample of 964 students in Public Finance.

Nearly half of the students in the sample took their Microeconomics I exam with *Professor 1*, and one third of the students with *Professor 5*. Possible explanations for these differences in the attendance rates include the fact, that these both professors offered this course more often.

In our sample the mean-value of High School GPA is approximately 2.5 which is higher, meaning better, than the mean of the grades obtained in Microeconomics I, Microeconomics II and Public Finance, respectively. This can be explained by the fact that a grade of 0.0, meaning failed, is not possible for the high school leaving certificate. The share of female students in all courses is about 40 percent. The purchasing power index is slightly lower for students in the Public Finance class and almost the same for students in Microeconomics I and in Microeconomics II. Microeconomics I has the lowest share of students with a private health insurance.

Table 5.1: Summary statistics

Variable	**Observations**	**Mean**	**Std. Dev.**	**Min**	**Max**
Grade in Microeconomics I	2920.00	2.12	1.11	0.00	4.00
High School GPA	2920.00	2.46	0.55	1.20	4.00
Female	2920.00	0.41	0.49	0.00	1.00
Private Health Insurance	2920.00	0.17	0.38	0.00	1.00
Purchasing Power Index	2920.00	99.09	11.76	70.16	258.82
Prof. 1 Dummy	2920.00	0.43	0.49	0.00	1.00
Prof. 2 Dummy	2920.00	0.10	0.30	0.00	1.00
Prof. 3 Dummy	2920.00	0.07	0.25	0.00	1.00
Prof. 4 Dummy	2920.00	0.11	0.31	0.00	1.00
Prof. 5 Dummy	2920.00	0.30	0.46	0.00	1.00
Easy Graders Dummy	2920.00	0.41	0.49	0.00	1.00
Tough Graders Dummy	2920.00	0.59	0.49	0.00	1.00
Grade in Microeconomics II	1255.00	2.03	1.13	0.00	4.00
High School GPA	1255.00	2.46	0.58	1.20	4.00
Female	1255.00	0.39	0.49	0.00	1.00
Private Health Insurance	1255.00	0.19	0.40	0.00	1.00
Purchasing Power Index	1255.00	99.10	11.71	70.16	186.99
Prof. 1 Dummy	1255.00	0.38	0.48	0.00	1.00
Prof. 2 Dummy	1255.00	0.06	0.25	0.00	1.00
Prof. 3 Dummy	1255.00	0.03	0.16	0.00	1.00
Prof. 4 Dummy	1255.00	0.18	0.39	0.00	1.00
Prof. 5 Dummy	1255.00	0.35	0.48	0.00	1.00
Easy Graders Dummy	1255.00	0.53	0.50	0.00	1.00
Tough Graders Dummy	1255.00	0.47	0.50	0.00	1.00
Grade in Public Finance	964.00	2.23	1.07	0.00	4.00
High School GPA	964.00	2.50	0.57	1.20	4.00
Female	964.00	0.40	0.49	0.00	1.00
Private Health Insurance	964.00	0.19	0.39	0.00	1.00
Purchasing Power Index	964.00	98.93	11.54	70.16	169.56
Prof. 1 Dummy	964.00	0.37	0.48	0.00	1.00
Prof. 2 Dummy	964.00	0.05	0.22	0.00	1.00
Prof. 3 Dummy	964.00	0.03	0.17	0.00	1.00
Prof. 4 Dummy	964.00	0.19	0.39	0.00	1.00
Prof. 5 Dummy	964.00	0.35	0.48	0.00	1.00
Easy Graders Dummy	964.00	0.54	0.50	0.00	1.00
Tough Graders Dummy	964.00	0.46	0.50	0.00	1.00

Grades transformed to 1-4 Scale, with 4 being the best grade and 1 being the worst grade that is still considered a pass. A grade of 0 means that the student failed the respective course.

5.3.2 Empirical Approach

In fact, we do not have course placement in the sense of a random assignment, meaning that the students could up to some extent freely decide on the timing of their exams and hence choose professors they want to write the exams with. Thus, although we cannot observe any assignment pattern, students' long-term expectations about who will teach Microeconomics I in following semesters should not be underestimated. In order to control for the potential self-selection bias of the students, we will follow the faculty's recommendation to take the Microeconomics I exam in the second semester of undergraduate studies.

In our analysis, we first examine the grading policy of the five professors assigned to Microeconomics I. Since we first want to analyze how does *tough* or *easy* grading from a fundamental undergraduate course correspond to students' later university performance, we divide the sample of professors in two groups (*Tough Graders* vs. *Easy Graders*) according to their grading standards - professors are classified by whether their mean grade is above or below the sample average. To examine the effect of having a *Tough Grader/Easy Grader* in Microeconomics I on the grade obtained in the respective follow-on class, ordinary least square (OLS) regressions are applied. Thus, in the baseline empirical model, student's performance can be described using a linear relationship with student's grade from a single class as the dependent variable and a vector of independent variables. The baseline model is then given by

$$y_{ic} = \beta_0 + \beta_1 GPA_i + \beta_2 S_i + \beta_3 P_i + \varepsilon_i \tag{5.1}$$

where y_{ic} is the grade for student i in course c; GPA_i is the high school GPA of student i; S_i is a vector of individual characteristics of i such as gender and socio-economic background, P_i is the dummy variable either for the professor the student i wrote her Microeconomics I exam with or for the group (*Tough Graders/Easy Graders*) her professor belongs to; ε_i is an error term. In all regressions, robust standard errors are clustered by semester. However, since the number of clusters is less than 30, it is possible that the estimated standard errors are biased downward. Therefore, we follow (Cameron et al., 2008) and report the wild bootstrap p-values below the coefficient estimates in brackets.

Instrumental Variables Regression (IV)

Estimates of β_3, the coefficient of professor's dummy variable from Microeconomics I exam, may be biased under OLS regressions due to the potential self-selection bias of students toward certain professors. For instance, if all weak students chose to write the Microeconomics I exam with an *Easy Grader*, the professor's effect in Microeconomics II would be much more negative. For this reason we treat the professor's dummy P_i as an endogenous regressor, assuming that P_i and ε_i are somehow correlated. Since we are treating P_i as endogenous, we need one or more additional variables that are correlated with P_i but not correlated with ε_i. When analyzing grades obtained in the Microeconomics II and the Public Finance exam, we group up professors according to their grading standards, so that P_i from the baseline model will be then replaced by $P(tough)_i$ representing the *Tough Graders*.

To account for the potential self-selection bias, we propose standard IV approach. Implementing the IV approach requires a two stage least squares estimation (2SLS) to be performed. This approach starts with the first stage of analysis, which is necessary given that the potential self-selection of students toward certain professors may affect both independent and dependent variables. In the first stage $P(tough)_i$ becomes the dependent variable and the independent variables include all control variables from the second stage as well as the instrumental variable. Addressing the faculty's recommendation to write the Microeconomics I exam in the second semester, we create an instrumental variable, that is equal to one if a student took the Microeconomics I exam with a professor who was supposed to offer this course in her second semester of studies, and equal to zero otherwise. Therefore, taking the exam with a professor who was assigned to the course in the student's second semester will influence the student's later performance. On the contrary, the fact that a professor is assigned to the course in the student's second semester does not affect the student's later achievements if a student decides, against the faculty's recommendation, to write her exam in an earlier or later semester. The instrument should not have any influence on the outcome variable. Moreover, these excluded exogenous variables must not influence grade y_{ic} directly, otherwise they should be included in the Equation 5.1. Since we have five dif-

ferent professors assigned to the Microeconomics I exam, we will have five instrumental variables, one for each professor. Therefore, the endogenous regressor *Tough Graders* will be instrumented by three (*Professor 1; Professor 2 and Professor 3*) additional exogenous variables.

The first stage equation looks as follows:

$$P(tough)_i = \pi_o + \pi_1 GPA_i + \pi_2 S_i + \pi_3 Professor1_i + \pi_2 Professor2_i + \pi_3 Professor3_i + \varepsilon_i \tag{5.2}$$

The second stage implements the Equation 5.1, in which the dependent variable is regressed on the predicted values from the first stage regression plus the control variables. It is assumed that the instrumental variables are uncorrelated with any omitted variables, thus removing the bias in the relationship between student's grade in a course and student's choice of the professor to write the exam with. In the following we will apply the above instrumental approach to the subsamples of students who obtained, besides the grade from the Microeconomics I, at least one grade in the Microeconomics II and/or in the Public Finance exam.

An important condition to obtain consistent estimation is that the instruments are not weak. This can be tested with the Kleibergen-Paap Wald rk F statistic, which is a robust analog to the Cragg Donald statistic and thus superior in the presence of heteroskedasticity, autocorrelation or clustering (Baum et al., 2007). It is an F statistic for the joint significance of the instruments in the first stage regression, which tests whether the instruments jointly explain a sufficient amount of the variation in the endogenous regressor. If the instruments are weak the standard errors can become considerably larger and the t statistics considerably smaller than those from OLS, indicating the loss of precision.

The other general specification is the Hansen test that implements a test of overidentifying restrictions and is robust to heteroskedasticity. The null hypothesis of this test is that overidentifying restrictions are valid (Cameron and Trivedi, 2010). Therefore the rejection of the null hypothesis is an indication that at least one of the instruments is not valid.

5.4 Results

Having a certain professor in the Microeconomics I exam may be associated with a different grading and teaching style, which in turn will influence students' performance in follow-on courses. On the one hand, it is likely that students who took their Microeconomics I exam with a *Tough Grader*, and thus learned more, will perform better in the Microeconomics II exam. This would suggest that besides different grading policies, we also find significant differences in learning outcomes. On the other hand, it may also be that they learned the same as they would with an *Easy Grader* and just got a worse grade, which simply refers to differences in grading policies.

5.4.1 Microeconomics I

By looking at students' performance in Microeconomics I, we estimate the effect of having a certain professor on the grade obtained from the first attempt in the Microeconomics I exam. This effect may arise from a number of professors' characteristics, for instance grading policy, teaching or examination style, or combination of those. For this course, our benchmark student is male, holds a public health insurance and is average with regard to all continuous variables.

The OLS estimation results can be found in Table 5.2. According to the results, the choice of a professor in Microeconomics I has a significant impact on the grade achieved in this subject. Hence, taking Microeconomics I exam with *Professor 1*, *Professor 2* or *Professor 3* results in lower grades than writing the exam with *Professor 5* (baseline). Thus, professors seem to vary considerably in their grading standards, even within a single course. This finding raises further questions whether the observed grading differences have a significant influence on students' later achievements.

Estimating the effect of professors' grading standards on students' later achievements assumes that these standards are relatively consistent over time, that they are not affected by the composition of students attending the course. To analyze this, we divided the sample of professors into two groups according to their grading standards each semester and analyzed if their position changes over time. The first group, *Tough Graders*, consists of *Professor 1*, *Professor 2* and *Professor 3*, and the second group, *Easy Graders*, of *Professor 4* and *Professor 5*. We found that there was no movement

between those groups, which led us, since we are not primarily interested in professors' individual characteristics, to use this classification for our subsequent analysis. It is not surprising that, when including these two groups as control variables, one finds a highly significant and negative effect of the *Tough Graders* on students' obtained grade in Microeconomics I.

In addition, we find the expected highly significant and positive effect of the high school leaving grade as shown in Table 5.2. The higher the high school leaving grade, the better the grade the student obtains in Microeconomics I exam. Although the size of the coefficient may appear to be somewhat large, it provides an indication that this course is based on skills and methods already used in high school. For instance, Mathematics serves as a prerequisite for all analyzed courses and is generally counted more heavily in the calculation of the GPA. This result is consistent with the findings of a large body of existing research (see e.g. Cyrenne and Chan, 2012; Girves and Wemmerus, 1988). Other controls are of lesser importance. The gender as well as the socio-economic variables, private health insurance and the purchasing power index of parents' zip-code area, do not show a significant effect in any of the regressions based on conventional hypothesis tests that are presented in Table 5.2. However, the purchasing power index is significant at the 5 percent level based on wild bootstrap p-values. These overall findings are in line with Danilowicz-Gösele et al. (2014) who found that socio-economic factors are, if at all, poor indicators of students' university performance.

Table 5.2: OLS regression estimates for Microeconomics I

Dependent variable: Grade in Microeconomics I				
	(1)	(2)	(3)	(4)
High School GPA			0.716***	0.713***
			(0.056)	(0.055)
			[0.000]	[0.000]
Female			-0.0731	-0.0707
			(0.060)	(0.061)
			[0.360]	[0.360]
Private Health Insurance			0.0631	0.0604
			(0.038)	(0.037)
			[0.160]	[0.160]
Purchasing Power Index			0.00243	0.00258
			(0.001)	(0.001)
			[0.040]	[0.040]
Course Assignment of Prof. 1	- 0.639***		-0.732***	
	(0.088)		(0.089)	
	[0.040]		[0.040]	
Course Assignment of Prof. 2	-0.622**		-0.728***	
	(0.219)		(0.180)	
	[0.200]		[0.080]	
Course Assignment of Prof. 3	-0.564***		-0.611***	
	(0.118)		(0.128)	
	[0.040]		[0.040]	
Course Assignment of Prof. 4	-0.226		-0.245	
	(0.179)		(0.178)	
	[0.200]		[0.200]	
Tough Graders		-0.568***		-0.653***
		(0.072)		(0.068)
		[0.040]		[0.040]
Constant	2.512***	2.452***	0.581***	0.508**
	(0.068)	(0.047)	(0.178)	(0.203)
	[0.000]	[0.000]	[0.000]	[0.000]
R^2	0.0673	0.0637	0.193	0.189
Observations	2920	2920	2920	2920
Cluster	18	18	18	18
F stat.	28	62	278	431

Notes: Stars indicate significance levels at 10%(*), 5%(**) and 1%(***). Standard errors clustered at a semester level are given in parentheses below each coefficient estimate. Wild bootstrap p-values are given in brackets below each coefficient estimate.

5.4.2 Microeconomics II

Knowing that there are relatively big differences in grading between *Tough Graders* and *Easy Graders* in Microeconomics I, we want to analyze whether the grades obtained in this fundamental course correspond to students' performance in subsequent courses. Did the students who obtained better grades in Microeconomics I in fact learn more? Or do some of the grades just mirror the grade inflation trend?

As already mentioned before, today's students very often have the possibility to decide on courses or even professors within one single course. Hence, the endogenous variables *Tough Graders* will be instrumented by the respective exogenous variables, namely the assignment of professors to the Microeconomics I course. Our instrumental variables fulfill the two usual conditions: (1) they are correlated with the endogenous variable and (2) do not affect students' performance in subsequent courses independently.

Table 5.3 presents the OLS estimation results and Tables 5.4 and 5.5 the two-stage least squares estimates using the assignment-based instrument for professors. First stage F-statistic and Kleibergen-Paap rk Wald F-statistic jointly confirm that the instrument is not weak. With the Hansen test, denoted as Hansen's J statistic, the validity of overidentifying restrictions cannot be rejected in specification 1 at the five percent level and in specifications 2 and 3, when controlling for having the same professor in both courses, at the ten percent level.

There is a strong ex ante expectation that the better the grade in Microeconomics I, the better the performance in Microeconomics II. Surprisingly, Table 5.4 shows significant and positive effect of *Tough Graders* on students' grade in Microeconomics II in all specifications. The size of the effect depends on whether or not the control variable for having the same professor in both analyzed courses is included. In the first specification, having a *Tough Grader* in Microeconomics I is associated with an improvement of the expected grade in Microeconomics II exam by 0.453 grades. This effect becomes less important when controlling for the effect of the *Same Professor*. At first view, this result appears to be straightforward: students profit from the familiar teaching and examination style when taking several courses with the same instructor. However, this effect becomes less obvious when distinguishing between both professors'

groups. According to the estimation results from specification 3, a student who took both Microeconomics exams with the same *Tough Grader* is worse off than a student who wrote both exams with a *Tough Grader* but not the same one. However, most disadvantaged are students who took both Microeconomics exams with an *Easy Grader*. Best off are students who had a *Tough Grader* in Microeconomics I and an *Easy Grader* in Microeconomics II. These results can partly be explained by the the fact that the *Tough Graders* from Microeconomics II are exactly the same ones we had in Microeconomics I. Still, there seem to be a significant difference between the both professors assigned to *Tough Graders* due to some unobservable characteristics.

From the above results, we conclude that, other things being equal, students who wrote their Microeconomics I exam with one of the *Tough Graders* are performing better in Microeconomics II exam. Hence, having a *Tough Grader* in the first-year course is positively related to the student's performance in follow-on courses. Therefore, grades obtained in classes with low-grading professors seem to be better predictors of students' later achievements than grades received from grade-deflating professors. These result are in line with the findings of Sabot and Wakeman-Linn (1991) and Ost (2010) and also confirm our speculation on the grade inflation within Microeconomics I course.

Furthermore, we find the expected highly significant and positive effect of the high school leaving degree on students' performance in Microeconomics II exam. An improvement of the high school GPA by one full grade is associated with an improvement of the expected grade in Microeconomics II by slightly more than 0.6 grades, which is only slightly lower than for Microeconomics I. This comparison indicates that the Microeconomics II course is based on concepts and skills that go somewhat beyond the high school level. In addition, we now find a significant negative effect for female, which is consistent with the existing literature on gender gap in Mathematics (Ellison and Swanson, 2010; Xie and Shauman, 2003). In our case, this result can be explained by the composition of students within a single course. The Microeconomics I course is mandatory for all students enrolled at the faculty of economic sciences. The Microeconomics II course, on the contrary, only to the students majoring in Economics. Furthermore, our data reveals that due to some unobserved characteristics female business students perform better in Mathematics than their female colleagues from Economics. These both

Table 5.3: Student performance in Microeconomics II (OLS)

	Grade in Microeconomics II		
	(1)	(2)	(3)
High School GPA	0.656***	0.635***	0.653***
	(0.067)	(0.065)	(0.066)
Female	-0.218**	-0.215***	-0.210***
	(0.076)	(0.072)	(0.070)
Private Health Insurance	0.0628	0.0448	0.0437
	(0.086)	(0.084)	(0.079)
Purchasing Power Index	0.00394*	0.00403*	0.00432**
	(0.002)	(0.002)	(0.002)
Micro I: Tough Grader	-0.00995	-0.0348	0.617***
	(0.085)	(0.078)	(0.156)
Same Professor		0.489***	
		(0.118)	
Same Professor - Easy Grader			0.898***
			(0.184)
Same Professor - Tough Grader			-0.0208
			(0.123)
Constant	0.0909	-0.198	-0.548*
	(0.253)	(0.292)	(0.311)
Observations	1230	1230	1230
Cluster	17	17	17

Notes: Stars indicate significance levels at 10%(*), 5%(**) and 1%(***). Standard errors clustered at a semester level are given in parentheses below each coefficient estimate.

insights may explain the significant negative effect for female in the Microeconomics II exam. However, this conclusion should be qualified only to some extent, because the wild bootstrap p-value for female is only significant in the last specification.

Our socio-economic variables, students' health insurance type and the purchasing power index of her parents' zip-code area, are of lesser importance. This results are in line with the findings of Danilowicz-Gösele et al. (2014) implying that socio-economic factors do not determine students' academic achievements, even if they are significantly associated with it.

Table 5.4: Student performance in Microeconomics II (IV) (second stage)

Second stage	Grade in Microeconomics II		
	(1)	(2)	(3)
High School GPA	0.614***	0.604***	0.644***
	(0.061)	(0.063)	(0.059)
	[0.000]	[0.000]	[0.000]
Female	-0.172*	-0.180**	-0.159*
	(0.088)	(0.083)	(0.094)
	[0.160]	[0.200]	[0.200]
Private Health Insurance	0.0570	0.0412	0.037
	(0.090)	(0.086)	(0.085)
	[0.560]	[0.640]	[0.720]
Purchasing Power Index	0.00221	0.00274*	0.003
	(0.002)	(0.002)	(0.002)
	[0.280]	[0.120]	[0.120]
Micro I: Tough Grader	0.453***	0.311**	1.715***
	(0.114)	(0.145)	(0.589)
	[0.000]	[0.080]	[0.000]
Micro II: Tough Grader			-1.047***
			(0.255)
			[0.040]
Same Professor		0.469***	
		(0.123)	
		[0.000]	
Same Professor - Tough Grader			-0.657*
			(0.357)
			[0.040]
Constant	0.132	-0.155	-0.831**
	(0.243)	(0.278)	(0.370)
	[0.440]	[0.400]	[0.480]
Observations	1230	1230	1230
Cluster	17	17	17
Kleibergen-Paap Wald F stat	230.162	172.419	174.991
Hansens J statistic	4.619	2.135	1.642
Hansen p-value	0.099	0.344	0.440

Notes: Stars indicate significance levels at 10%(*), 5%(**) and 1%(***). Standard errors clustered at a semester level are given in parentheses below each coefficient estimate. Wild bootstrap p-values are given in brackets below each coefficient estimate.

Table 5.5: Student performance in Microeconomics II (IV) (first stage)

First Stage	Tough Graders		
	(1)	(2)	(3)
High School GPA	0.0829***	0.0819***	0.006
	(0.026)	(0.025)	(0.018)
	[0.040]	[0.040]	[0.720]
Female	-0.0881**	-0.0876**	-0.038*
	(0.036)	(0.036)	(0.018)
	[0.120]	[0.120]	[0.120]
Private Health Insurance	0.00924	0.00893	0.000
	(0.025)	(0.025)	(0.016)
	[0.480]	[0.520]	[0.800]
Purchasing Power Index	0.00356***	0.00357***	0.001
	(0.001)	(0.001)	(0.001)
	[0.000]	[0.000]	[0.160]
Course Assignment of Prof. 1	0.0881	0.0784	0.037
	(0.086)	(0.085)	(0.053)
	[0.520]	[0.600]	[0.560]
Course Assignment of Prof. 2	0.612***	0.605***	0.225***
	(0.046)	(0.048)	(0.030)
	[0.000]	[0.000]	[0.000]
Course Assignment of Prof. 3	0.643***	0.627***	0.060
	(0.047)	(0.054)	(0.055)
	[0.000]	[0.000]	[0.240]
Micro II: Tough Grader			0.248***
			(0.078)
			[0.000]
Same Professor		0.033	
		(0.042)	
		[0.560]	
Same Professor - Tough Grader			0.601***
			(0.094)
			[0.000]
Constant	-0.150	-0.166*	0.226**
	(0.095)	(0.093)	(0.088)
	[0.200]	[0.160]	[0.880]
Observations	1230	1230	1230
Cluster	17	17	17
F first-stage	175	199	2788

Notes: Stars indicate significance levels at 10%(*), 5%(**) and 1%(***). Standard errors clustered at a semester level are given in parentheses below each coefficient estimate. Wild bootstrap p-values are given in brackets below each coefficient estimate.

5.4.3 Public Finance

The second subsequent course we analyze is the Public Finance course. The OLS estimation results can be found in Table 5.6. Tables 5.7 and 5.8 report two-stage least squares estimates for our outcome variable across two different specifications. Instrumentation is strong, as indicated by the first stage F-statistic and Kleibergen-Paap rk Wald F-statistic. Hansen's J statistic is far from rejection of its null, implying the validity of overidentifying restrictions.

Here again, we have a strong ex ante expectation that the grade obtained in Microeconomics I can be used as a predictor for student's performance in the Public Finance class. In the first specification, where we do not control for the effect of having the same professor in both analyzed courses, we find a highly significant but, surprisingly, negative effect of the *Tough Graders*. This conclusion has to be qualified to some extent, since the wild bootstrap p-value for the *Tough Graders* is insignificant. This negative effect becomes smaller once the control variable for having the same professor in both courses is included.

Both coefficients, *Tough Graders* and *Same Professor* are highly significant in the last specification, both based on conventional and wild bootstrap hypothesis tests. In this case, the interpretation is a little bit different than in the case of Microeconomics II, since there is only one professor assigned to teach this course. In order to understand the obtained results, we need to take into account that the professor assigned to teach the Public Finance course is the one who gives the worst grades in Microeconomics I and thus a *Tough Grader*. Hence, the students who wrote their both exams, Microeconomics I and Public Finance, with the same professor are slightly better of than those students who took their Microeconomics I exam with one of the *Easy Graders*.

In contrast to the results found for Microeconomics II, the worst off are now students who have their Microeconomics I grade from one of the other two *Tough Graders*. A reason for the partly inconsistent results may be that, although both courses are strongly related to Microeconomics I, the Public Finance course is less mathematical and has a different examination style than Microeconomics II. In Microeconomics I and Microeconomics II students complete exam problems that are either similar to the multiple choice question type (true/false) or graded on the basis of the final result (fill-ins).

However, exam problems in Public Finance include essay questions and calculations which gives professors more freedom in grading their students. For this reasons, we did not expect to find such a strong effect of having the same professor in both courses.

Table 5.6: Student performance in Public Finance (OLS)

	Grade in Public Finance	
	(1)	(2)
High School GPA	0.594***	0.576***
	(0.067)	(0.070)
Female	0.0210	0.0190
	(0.075)	(0.073)
Private Health Insurance	0.251***	0.273***
	(0.057)	(0.057)
Purchasing Power Index	0.00330	0.00287
	(0.003)	(0.003)
Micro I: Tough Grader	-0.0931	-0.690***
	(0.103)	(0.108)
Same Professor		0.732***
		(0.138)
Constant	0.404	0.487
	(0.317)	(0.349)
Observations	961	961
Cluster	14	14

Notes: Stars indicate significance levels at 10%(*), 5%(**) and 1%(***). Standard errors clustered at a semester level are given in parentheses below each coefficient estimate.

Table 5.7: Student performance in Public Finance (IV) (second stage)

Second stage	Grade in Public Finance	
	(1)	(2)
High School GPA	0.671***	0.576***
	(0.057)	(0.068)
	[0.000]	[0.000]
Female	-0.0757	0.0178
	(0.071)	(0.070)
	[0.240]	[0.920]
Private Health Insurance	0.271***	0.274***
	(0.070)	(0.055)
	[0.000]	[0.000]
Purchasing Power Index	0.00645**	0.00290
	(0.003)	(0.002)
	[0.080]	[0.120]
Micro I: Tough Grader	-0.926***	-0.723***
	(0.329)	(0.110)
	[0.120]	[0.040]
Same Professor		0.761***
		(0.142)
		[0.000]
Constant	0.313	0.490
	(0.295)	(0.337)
	[0.400]	[0.280]
Observations	961	961
Cluster	14	14
Kleibergen-Paap Wald F stat	208.172	3468.381
Hansens J statistic	2.365	2.652
Hansen p-value	0.306	0.266

Notes: Stars indicate significance levels at 10%(*), 5%(**) and 1%(***). Standard errors clustered at a semester level are given in parentheses below each coefficient estimate. Wild bootstrap p-values are given in brackets below each coefficient estimate.

Table 5.8: Student performance in Public Finance (IV) (first stage)

First Stage	Tough Graders	
	(1)	(2)
High School GPA	0.0970***	0.0124
	(0.024)	(0.008)
	[0.000]	[0.160]
Female	-0.108**	-0.0129*
	(0.046)	(0.007)
	[0.040]	[0.080]
Private Health Insurance	0.0135	0.0291
	(0.037)	(0.020)
	[0.840]	[0.120]
Purchasing Power Index	0.00446***	0.00132*
	(0.001)	(0.001)
	[0.000]	[0.120]
Course Assignment of Prof. 1	0.109	0.0139
	(0.104)	(0.019)
	[0.280]	[0.440]
Course Assignment of Prof. 2	0.659***	0.925***
	(0.047)	(0.042)
	[0.000]	[0.000]
Course Assignment of Prof. 3	0.684***	0.705***
	(0.043)	(0.009)
	[0.000]	[0.000]
Same Professor		0.936***
		(0.040)
		[0.000]
Constant	-0.296**	-0.120
	(0.101)	(0.068)
	[0.040]	[0.120]
Observations	961	961
Cluster	14	14
F first- stage	158	20274

Notes: Stars indicate significance levels at 10%(*), 5%(**) and 1%(***). Standard errors clustered at a semester level are given in parentheses below each coefficient estimate. Wild bootstrap p-values are given in brackets below each coefficient estimate.

This finding suggests that there are differences in professors' characteristics, not only between the *Easy Graders* and the *Tough Graders*, but also within those groups. In order to analyze this, we create two sub-samples: the first one includes students who wrote their Microeconomics I exam either with the professor who gives the worst grades (*Professor 1*) or with the professor who gives the best grades (*Professor 5*). The second sub-sample consists of students who took their Microeconomics I exam again with the toughest grader (*Professor 1*) or with the second toughest grader (*Professor 2*). The results of estimating the effect of a single professor on the grade in Public Finance exam can be found in Tables 5.9 and 5.10. The baseline category is now represented by *Professor 1*. For both professors, *Professor 2* and *Professor 5*, we find highly significant and negative effect on student's performance in Public Finance. According to this result, having a tough professor in Microeconomics I does not always positively affect student's later performance in follow-on courses. Here, a student who took her Microeconomics I exam with a professor who "gives out" grades is better off than a student with the second toughest grader. On the one hand, it looks like, some of the *Tough Graders* teach better and demand higher performance from their students, others just give lower grades. On the other hand, *Easy Graders* do not always "only" inflate grades. Furthermore, looking at some other follow-on courses suggests that students benefit from having an *Easy Grader* in less theoretical courses where mathematical skills are not that essential and examinations thus require less calculations.

In addition, we find the expected highly significant and positive effect of the high school leaving degree on students' grade in Public Finance exam. Since the Public Finance course is a less mathematical one, we do not find a significant effect for female, which is again consistent with our previous argumentation about prevalent gender gap in mathematics. The effect of student's health insurance type is now highly significant and positive, which can be explained by the differences in students composition between Microeconomics and Public Finance. Microeconomics courses are mandatory for many students enrolled at the faculty of economic sciences. Public Finance, on the contrary, is completed mostly by the students from economics, who in contrast to their business colleagues, are less known for being fast climbers which in turn can relate to their wealthy family status. A good socio-economic background is often related to the level

of educational attainment of the parents. In the case of the Public Finance course students from educated families can benefit from discussions and dialogue at home.

Table 5.9: Student performance in Public Finance (IV) - Comparison (second stage)

Second stage	Grade in Public Finance	
	Professor 5	Professor 2
High School GPA	0.523***	0.616***
	(0.065)	(0.093)
	[0.000]	[0.000]
Female	0.104	-0.0698
	(0.067)	(0.075)
	[0.080]	[0.320]
Private Health Insurance	0.291***	0.385***
	(0.072)	(0.097)
	[0.000]	[0.000]
Purchasing Power Index	0.00260	0.00119
	(0.003)	(0.003)
	[0.520]	[0.840]
Micro I: Professor 5	-0.396***	
	(0.147)	
	[0.040]	
Micro I: Professor 2		-0.522***
		(0.137)
		[0.040]
Constant	0.836*	0.601
	(0.459)	(0.474)
	[0.120]	[0.160]
Observations	697	410
Cluster	13	13
Kleibergen-Paap Wald F stat	69.702	206.523
Hansens J statistic	0.000	0.000

Notes: Stars indicate significance levels at 10%(*), 5%(**) and 1%(***). Standard errors clustered at a semester level are given in parentheses below each coefficient estimate. Wild bootstrap p-values are given in brackets below each coefficient estimate.

Table 5.10: Student performance in Public Finance (IV) - Comparison (first stage)

First Stage	Professor 5	Professor 2
High School GPA	-0.101***	0.00980
	(0.031)	(0.014)
	[0.080]	[1.000]
Female	0.129**	0.0146
	(0.059)	(0.015)
	[0.120]	[0.640]
Private Health Insurance	-0.0199	0.0526
	(0.051)	(0.039)
	[0.680]	[0.280]
Purchasing Power Index	-0.00442***	0.00203
	(0.001)	(0.001)
	[0.040]	[0.200]
Course Assignment of Prof. 5	0.329***	
	(0.394)	
	[0.000]	
Course Assignment of Prof. 2		0.917***
		(0.064)
		[0.000]
Constant	1.062***	-0.208
	(0.136)	(0.149)
	[0.000]	[0.400]
Observations	697	410
Cluster	13	13
F first- stage	26	747

Notes: Stars indicate significance levels at 10%(*), 5%(**) and 1%(***). Standard errors clustered at a semester level are given in parentheses below each coefficient estimate. Wild bootstrap p-values are given in brackets below each coefficient estimate.

5.5 Discussion

In recent years a great number of studies on higher education have emphasized factors that might influence students' performance. While most of those studies focus on students' characteristics such as previous academic performance, age or socio-economic status, less attention has been paid to the impact of professors' characteristics or professors' grading.

Therefore in this paper, we analyze the impact of professors' characteristics on students' achievement using data from almost 3,000 students from Göttingen University. By looking at three courses at the Faculty of Economic Sciences, we first capture the professor's effect from the mandatory first-year course (Microeconomics I) and then analyze its effect on student's performance in follow-on courses (Microeconomics II and Public Finance). Two main results emerge from our analysis. Firstly, we find that it matters, for student's later university performance in micro-related courses, which professor was teaching and giving the Microeconomics I exam. Our results suggest that students can benefit from having a tough or an easy grader in a fundamental course, with the effects depending on the student's prior academic performance and design of the follow-on course, in particular the mathematical content of the course and the resulting examination form. Secondly, we show that the effect of having the same professor in Microeconomics I and in one of the two analyzed follow-on courses is significant and highly relevant. Thus, students benefit much more from the familiar teaching practices and the familiar examination style. In some cases, this positive effect can even compensate or partially compensate for the negative professor's effect from Microeconomics I course.

From the consideration of the above results two questions arise: How can we explain the arising differences in grading? and Do the *Tough Graders* indeed prepare students for more rigorous mathematical and analytical standards by teaching something different than their easy grading colleagues? In order to answer the first question, we should consider if there are, besides the grading, other substantial differences between both professor's groups included in our analysis. When looking at the groups individually, professors assigned to the *Easy Graders* are of higher age, have been employed longer at a university and do less research than their younger colleagues. Together, our results

and these facts lead us to hypothesize that professor's age and teaching experience may be, above certain threshold, positively associated with students' grades. Older professors very often exude grandfather's mildness and thus have more understanding for the students. Another explanation for this result may stem from the fact that the longer a professor was teaching the same subject, the easier he is to predict.

Furthermore, we can think about other factors, besides individual characteristics of the teaching body, that potentially affect the grades and may differ between professors. One of them is the class size. Since all professors structure the course in the same way, offering lecture and tutorials on a weekly basis should not be an issue. Also the credentials of the faculty teaching team should not determine our results, since we do not observe striking differences in education level between the assigned professors and their assistants. In addition, all professors included in our data set have very similar status, implying that none of them are in a position of having to earn good evaluations in order to be able to keep their job. Nonetheless, some of the professors may generally tend to keep the students happy by giving them good, possibly inflated, grades. In particular, those professors, who are close to retire, are less stringent since they want to leave a good impression.

From the students' point of view, the benefit of having an *Easy Grader* or a *Tough Grader* in Microeconomics I depends to a great extent on the student's course choices and her timing. If grades were the only aim, students may act strategically by taking advantage of the mentioned differences, which will result in postponing the exams until the desired professor is offering the class or in attending only carefully-selected courses. Our analysis shows that the obtained effects are quite meaningful. All else being equal, having a certain professor in Microeconomics I is associated with an improvement of the expected grade in a follow-on course by up to 1.385 grades. Therefore, making smart strategic choices about study pathways may lead to a considerable improvement of the final grade.

From the above results we conclude that there is much variability in grading not only between universities or faculties, but also from one professor to the next and between courses within a field. The increasing diversity especially within higher education system is valuable but comes at the expense of transparency in grading policies. Al-

though grades are used to motivate students and to report the quality of student's performance for employers, researchers and politicians consistently raise arguments in favor of abandoning grades and rather encouraging students to pursue more than a perfect transcript.

Bibliography

Achen, A. C. and Courant, P. N. (2009). What are Grades Made of? *The Journal of Economic Perspectives*, 23(3):77–92.

Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, 113(1):151–184.

Angeletos, G.-M., Laibson, D., Repetto, A., Tobacman, J., and Weinberg, S. (2001). The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation. *The Journal of Economic Perspectives*, 15(3):47–68.

Anger, C., Demary, V., Koppel, O., and Plünnecke, A. (2013). Zu wenig Nachwuchs. *iw-dienst*, 19. May 9, 2013. `https://www.iwd.de/artikel/zu-wenig-nachwuchs-111260/` (accessed September 17, 2014).

Arias Ortiz, E. and Dehon, C. (2008). What Are the Factors of Success at University? A Case Study in Belgium. *CESifo Economic Studies*, 54(2):121–148.

Aronsson, T. and Thunström, L. (2008). A note on optimal paternalism and health capital subsidies. *Economics Letters*, 101(3):241–242.

Arulampalam, W., Naylor, R. A., and Smith, J. P. (2005). Effects of In-Class Variation and Student Rank on the Probability of Withdrawal: Cross-Section and Time-Series Analysis for UK University Students. *Economics of Education Review*, 24(3):251–262.

Babcock, P. and Marks, M. (2011). The Falling Time Cost of College: Evidence from Half a Century of Time Use Data. *The Review of Economics and Statistics, MIT Press*, 93(2):468–478.

Bar, T., Kadiyali, V., and Zussman, A. (2009). Grade Information and Grade Inflation: The Cornell Experiment. *Journal of Economic Perspectives*, 23(3):93–108.

Baron-Boldt, J. (1989). *Die Validität von Schulabschlussnoten für die Prognose von Ausbildungs- und Studienerfolg: Eine Metaanalyse nach dem Prinzip der Validitätsgeneralisierung*. Europäische Hochschulschriften. Frankfurt am Main: Lang.

Bartlett, S., Jacob, K., and Lauren, O. (2014). Evaluation of the Healthy Incentives Pilot (HIP): Final Report. Prepared by Abt Associates for the U.S. Department of Agriculture, Food and Nutrition Service, September 2014.

Baum, C., Schaffer, M., and Stillman, S. (2007). Enhanced routines for instrumental variables/GMM estimation and testing. *Stata Journal*, 4(7):465–506.

Becker, G. S. and Murphy, K. M. (1988). A Theory of Rational Addiction. *Journal of Political Economy*, 96(4):675–700.

Bettinger, E. P. and Long, B. T. (2005). Do Faculty Serve as Role Models? The Impact of Instructor Gender on Female Students. *American Economic Review*, 95(2):152–157.

Betts, J. R. and Morel, D. (1999). The Determinants of Undergraduate Grade Point Average: The Relative Importance of Family Background, High School Resources and Peer Group Effects. *Journal of Human Resources*, 34(2):268–293.

Birnbaum, R. (1977). Factors Related to University Grade Inflation. *The Journal of Higher Education*, 48(5):519–539.

Brezis, M. and Marans, R. (2010). Do People Support Sin Taxes? A Population Survey of Attitudes. *World Medical and Health Policy*, 2(4):33–43.

Burton, L. J. and Dowling, D. (2005). In Search of the Key Factors that Influence Student Success at University. In *Proceedings of the 28th Higher Education Research and Development Association Conference: Higher Education in a changing world.* 68-78.

Business Roundtable (2007). Doing well through wellness: 2006-07 survey of wellness programs at Business Roundtable member companies. Washington, DC: Business Roundtable. `http://www.businessroundtable.org/uploads/studies-reports/downloads/BR_Doing_Well_through_Wellness_09192007.pdf` (accessed February 26, 2013).

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics*, 90(3):417–427.

Cameron, A. C. and Trivedi, P. K. (2010). *Microeconometrics Using Stata*. Stata Press, College Station, Texas.

Carrell, S. E., Page, M. E., and West, J. E. (2010). Sex and Science: How Professor Gender Perpetuates the Gender Gap. *The Quarterly Journal of Economics*, 125(3):1101–1144.

Chetty, R. (2015). Behavioral Economics and Public Policy: A Pragmatic Perspective. *American Economic Review*, 105(5):1–33.

Clotfelter, C. T., Ladd, H. F., and Vigdor, J. L. (2007). Teacher Credentials and Student Achievement: Longitudinal Analysis with Student Fixed Effects. *Economics of Education Review*, 26(6):673 – 682.

Cohn, E., Cohn, S., Balch, D. C., and Bradley, J. (2004). Determinants of Undergraduate GPAs: SAT scores, High-School GPA and High-School Rank. *Economics of Education Review*, 23(6):577–586.

Conlisk, J. (1996). Why Bounded Rationality? *Journal of Economic Literature*, 34(2):669–700.

Cremer, H., De Donder, P., Maldonado, D., and Pestieau, P. (2012). Taxing Sin Goods and Subsidizing Health Care. *The Scandinavian Journal of Economics*, 114(1):101–123.

Cyrenne, P. and Chan, A. (2012). High School Grades and University Performance: A Case Study. *Economics of Education Review*, 31(5):524–542.

Danilowicz-Gösele, K. (2016). "A" is the Aim? *Cege Discussion Papers*. No. 291.

Danilowicz-Gösele, K., Meya, J., Schwager, R., and Suntheim, K. (2014). Determinants of Student Success at University. *Cege Discussion Papers*. No. 214.

Danilowicz-Gösele, K. and Schwager, R. (2016). Subsidizing Health-Conscious Behavior Now or Later. *CESifo Working Paper*. No. 5734.

De Paola, M. (2009). Does Teacher Quality Affect Student Performance? Evidence from an Italian University. *Bulletin of Economic Research*, 61(4):353–377.

DellaVigna, S. (2009). Psychology and Economics: Evidence from the Field. *Journal of Economic Literature*, 47(2):315–72.

Diamond, P. and Kőszegi, B. (2003). Quasi-Hyperbolic Discounting and Retirement. *Journal of Public Economics*, 87(910):1839–1872.

Dobson, I. R. and Skuja, E. (2005). Secondary Schooling, Tertiary Entry Ranks and University Performance. *People and Place*, 13(1):53–62.

Eiszler, C. F. (2002). College Students' Evaluations of Teaching and Grade Inflation. *Research in Higher Education*, 43(4):483–501.

Ellison, G. and Swanson, A. (2010). The Gender Gap in Secondary School Mathematics at High Achievement Levels: Evidence from the American Mathematics Competitions. *Journal of Economic Perspectives*, 24(2):109–28.

European Commission (2012). Evaluation of the European School Fruit Scheme. `http://ec.europa.eu/agriculture/evaluation/market-and-income-reports/2012/school-fruit-scheme/fulltext_en.pdf` (accessed February 11, 2015).

Finkenstädt, V. and Keßler, T. (2012). Die sozioökonomische Struktur der PKV-Versicherten: Ergebnisse der Einkommens- und Verbrauchsstichprobe 2008. *WIP Discussion Paper*. No. 3/2012.

French, S. A., Jeffery, R. W., Story, M., Breitlow, K. K., Baxter, J. S., Hannan, P., and Snyder, M. P. (2001). Pricing and Promotion Effects on Low-Fat Vending Snack Purchases: the CHIPS Study. *American Journal of Public Health*, 91(1):112–117.

Gabaix, X., Laibson, D., Moloche, G., and Weinberg, S. (2006). Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model. *American Economic Review*, 96(4):1043–1068.

Gerdes, H. and Mallinckrodt, B. (1994). Emotional, Social, and Academic Adjustment of College Students: A Longitudinal Study of Retention. *Journal of Counseling and Development*, 72(3):281–288.

Gigerenzer, G. (2015). On the Supposed Evidence for Libertarian Paternalism. *Review of Philosophy and Psychology*, 6(3):361–383.

Girves, J. E. and Wemmerus, V. (1988). Developing Models of Graduate Student Degree Progress. *The Journal of Higher Education*, 59(2):163–189.

Grebennikov, L. and Skaines, I. (2009). University of Western Sydney Students at Risk: Profile and Opportunities for Change. *Journal of Institutional Research*, 14(1):58–70.

Gruber, J. and Kőszegi, B. (2001). Is Addiction Rational? Theory And Evidence. *The Quarterly Journal of Economics*, 116(4):1261–1303.

Gruber, J. and Kőszegi, B. (2004). Tax Incidence When Individuals Are Time-inconsistent: The Case of Cigarette Excise Taxes. *Journal of Public Economics*, 88(9-10):1959–1987.

Harris, C. and Laibson, D. (2001). Dynamic Choices of Hyperbolic Consumers. *Econometrica*, 69(4):935–957.

Hoffmann, F. and Oreopoulos, P. (2009). A Professor like Me: The Influence of Instructor Gender on College Achievement. *The Journal of Human Resources*, 44(2):479–494.

Hong, S.-M. (1984). The Age Factor in the Prediction of Tertary Academic Success. *Higher Education Research and Development*, 3(1):61–70.

Johnson, V. E. (2003). *Grade Inflation: A Crisis in College Education*. New York: Springer.

Kahneman, D. and Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2):263–291.

Karp, L. (2005). Global warming and hyperbolic discounting. *Journal of Public Economics*, 89(23):261–282.

Kena, G., Musu-Gillette, L., Robinson, J., Wang, X., Rathbun, A., Zhang, J., Wilkinson-Flicker, S., Barmer, A., and Dunlop Velez, E. (2015). *The Condition of Education 2015*. (NCES 2015-144) Washington, DC: U.S. Department of Education, National Center for Education Statistics. `http://nces.ed.gov/pubs2015/2015144.pdf` (accessed January 1, 2016).

Krass, S. and Scherf, M. (2012). Warum die Einser-Inflation nicht überrascht. *Süddeutsche Zeitung*. November 19, 2012. `http://www.sueddeutsche.de/bildung/gute-noten-an-hochschulen-warum-die-einser-inflation-nicht-ueberrascht-1.1526490` (accessed July 14, 2014).

Krause, K.-L., Hartley, R., James, R., and McInnis, C. (2005). *The First Year Experience in Australian Universities: Findings From a Decade of National Studies*. Centre for the Study of Higher Education, University of Melbourne, Melbourne.

Krautmann, A. C. and Sander, W. (1999). Grades and Student Evaluations of Teachers. *Economics of Education Review*, 18(1):59 – 63.

Laibson, D. (1994). *Essays in Hyperbolic Discounting*. Ph.D. dissertation, MIT, USA.

Laibson, D. (1997). Golden Eggs and Hyperbolic Discounting. *Quarterly Journal of Economics*, 112(2):443–477.

Laibson, D. (1998). Life-cycle Consumption and Hyperbolic Discount Functions. *European Economic Review*, 42(35):861–871.

Levitt, S. D. and List, J. A. (2007). What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World? *The Journal of Economic Perspectives*, 21(2):153–174.

Levy, S. and Murray, J. (2005). Tertiary Entrance Scores Need Not Determine Academic Success: An Analysis of Student Performance in an Equity and Access Program. *Journal of Higher Education Policy and Management*, 27(1):129–141.

List, J. A. (2004). Neoclassical Theory versus Prospect Theory: Evidence from the Marketplace. *Econometrica*, 72(2):615–625.

McInnis, C., James, R., and Hartley, R. (2000). *Trends in the First Year Experience in Australian Universities*. Canberra: Department of Education, Training and Youth Affairs.

McKenzie, K. and Schweitzer, R. (2001). Who Succeeds at University? Factors Predicting Academic Performance in First Year Australian University Students. *Higher Education Research and Development*, 20(1):21–33.

McNabb, R., Pal, S., and Sloane, P. (2002). Gender Differences in Educational Attainment: The Case of University Students in England and Wales. *Economica*, 69(275):481–503.

Miles, D. A. and Sparks, W. (2013). Examining Social Media and Higher Education: An Empirical Study On Rate My Professors.Com and Its Impact On College Students. *International Journal of Economy, Management and Social Sciences*, 2(7):513–524.

Nelson, J. P. and Lynch, K. A. (1984). Grade Inflation, Real Income, Simultaneity, and Teaching Evaluations. *The Journal of Economic Education*, 15(1):21–37.

Neumark, D. and Gardecki, R. (1998). Women Helping Women? Role Model and Mentoring Effects on Female Ph.D. Students in Economics. *The Journal of Human Resources*, 33(1):220–246.

O'Donoghue, T. and Rabin, M. (1999). Incentives for Procrastinators. *The Quarterly Journal of Economics*, 114(3):769–816.

O'Donoghue, T. and Rabin, M. (2003). Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes. *American Economic Review*, 93(2):186–191.

O'Donoghue, T. and Rabin, M. (2006). Optimal Sin Taxes. *Journal of Public Economics*, 90(10-11):1825–1849.

OECD (2013). *Economic Policy Reforms 2013: Going for Growth*. Paris: OECD Publishing.

Ost, B. (2010). The Role of Peers and Grades in Determining Major Persistence in the Sciences. *Economics of Education Review*, 29(6):923 – 934.

Parker, J. D., Summerfeldt, L. J., Hogan, M. J., and Majeski, S. A. (2004). Emotional Intelligence and Academic Success: Examining the Transition from High School to University. *Personality and Individual Differences*, 36(1):163–172.

Peat, M., Dalziel, J., and Grant, A. M. (2001). Enhancing the First Year Student Experience by Facilitating the Development of Peer Networks through a One-Day Workshop. *Higher Education Research and Development*, 20(2):199–215.

Powell, L. M., Chriqui, J. F., Khan, T., Wada, R., and Chaloupka, F. J. (2013). Assessing the Potential Effectiveness of Food and Beverage Taxes and Subsidies for Improving Public Health: A Systematic Review of Prices, Demand and Body Weight Outcomes. *Obesity Reviews*, 14(2):110–128.

Prather, James E., G. S. and Kodras, J. E. (1979). A Longitudinal Study of Grades in 144 Undergraduate Courses. *Research in Higher Education*, 10(1):11–24.

Pressman, S. (2007). The Economics of Grade Inflation. *Challenge*, 50(5):93–102.

Rivkin, S. G., Hanushek, E. A., and Kain, J. F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica*, 73(2):417–458.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., and Carlstrom, A. (2004). Do Psychosocial and Study Skill Factors Predict College Outcomes? A Meta-Analysis. *Psychological Bulletin*, 130(2):261–288.

Rojstaczer, S. and Healy, C. (2010). Grading in American Colleges and Universities. *Teachers College Record*, 4.

Sabot, R. and Wakeman-Linn, J. (1991). Grade Inflation and Course Choice. *Journal of Economic Perspectives*, 5(1):159–170.

Simon, H. A. (1947). *Administrative Behavior. A Study of Decision-making Processes in Administrative Organizations.* Macmillan. New York.

Simon, H. A. (1985). Human Nature in Politics: The Dialogue of Psychology with Political Science. *The American Political Science Review*, 79(2):293–304.

Simpson, O. (2006). Predicting Student Success in Open and Distance Learning. *Open Learning*, 21(2):125–138.

Smith, J. and Naylor, R. (2005). Schooling Effects on Subsequent University Performance: Evidence for the UK University Population. *Economics of Education Review*, 24(5):549–562.

Sonner, B. S. (2000). A is for "Adjunct": Examining Grade Inflation in Higher Education. *Journal of Education for Business*, 76(1):5–8.

Statistisches Bundesamt (2009). *Bildungsstand der Bevölkerung: Ausgabe 2009*. Wiesbaden.

Stratton, R. W., Myers, S. C., and King, R. H. (1994). Faculty Behavior, Grades, and Student Evaluations. *The Journal of Economic Education*, 25(1):5–15.

Thaler, R. H. (1981). Some empirical evidence on dynamic inconsistency. *Economics Letters*, 8(3):201 – 207.

Thaler, R. H. and Sunstein, C. R. (2003). Libertarian Paternalism. *American Economic Review*, 93(2):175–179.

Tinto, V. (1975). Dropout from Higher Education: A Theoretical Synthesis of Recent Research. *Review of Educational Research*, 45(1):89–125.

Trapmann, S., Hell, B., Weigand, S., and Schuler, H. (2007). Die Validität von Schulnoten zur Vorhersage des Studienerfolgs - eine Metaanalyse. *Zeitschrift für Pädagogische Psychologie*, 21(1):11–27.

Volpp, K., John, L., Troxel, A., Norton, L., Fassbinder, J., and Loewenstein, G. (2008). Financial Incentive-Based Approaches for Weight Loss. *Journal of the American Medical Association*, 300(22):2631–2637.

Volpp, K., Troxel, A., Pauly, M., Glick, H., Puig, A., Asch, D., Galvin, R., Zhu, J., Wan, F., DeGuzman, J., Corbett, E., Weiner, J., and Audrain-McGovern, J. (2009). A Randomized Trial of Financial Incentives For Smoking Cessation. *New England Journal of Medicine*, 350(7):699–709.

Xie, Y. and Shauman, K. A. (2003). *Women in Science: Career Processes and Outcomes.* Cambridge, MA: Harvard University Press.