Empirical Studies from the Fields of Health, Education and Labour Economics

DISSERTATION
of the University of St. Gallen,
School of Management,
Economics, Law, Social Sciences
and International Affairs
to obtain the title of
Doctor of Philosophy in Economics and Finance

submitted by

Darjusch Tafreschi
from
Germany

Approved on the application of

Prof. Dr. Michael Lechner
and
Prof. Eddy van Doorslaer, PhD

Dissertation no. 4197

epubli GmbH, Berlin 2013

The University of St. Gallen, School of Management, Economics, Law, Social Sciences and International Affairs hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St. Gallen, April 29, 2013

The President:

Prof. Dr. Thomas Bieger

Acknowledgements

I dedicate this dissertation to my family and my woman, they are my backbone. The articles reflect my academic learning process between the years 2007-2012 and I thank my supervisor Michael Lechner for the opportunity. I also want to thank my co-authors Sofie Cabus, Eva Deuchert, Sharon Pfister, Benjamin Schuenemann and Petra Thiemann for their dedication and willingness to collaborate - my friends and colleagues Dirk Burghardt, Philipp Denter, Christina Felfe, Martin Huber, Giovanni Mellace, Markus Olapade, Philip Schuster, Andreas Steinmayr, Thomas Vischer and Conny Wunsch for interesting discussions and their steady support – and Eddy Van Doorslaer, Owen O' Donnell and Tom Van Ourti for hosting me during the later stages of my PhD studies and beyond.

Contents

List of Figures i

List of Tables iii

I Introduction 1

II The income body weight gradients in the developing economy of China 8

1 Introduction 9

2 Background 12
2.1 China's Economic Development 12
2.2 The Evolution of Overweight in China 14

3 Data 18

4 Development Index 21
4.1 An Index of Economic-Development at the Community Level 21
4.2 Assessment of Development Index 26

5 Regression Analyses 28
5.1 Set-Up 28
5.2 Results 32

6 Conclusion 33

7 Appendix 38

III Economic Development and Socioeconomic Inequality in Female Body Weight 43

1 Introduction 44

2 Theoretical Background 45

3 Empirical Analysis 47
3.1 Data 47
3.2 Results 49

4 Concluding Discussion 52

5 Appendix 55

IV Retention Effects in Higher Education 59

1 Introduction 60

2 Literature **61**

3 Institutional Setup **62**

4 Data **65**
4.1 General description . . . 65
4.2 What happens during the first year? . . . 68
4.3 Estimation sample . . . 71

5 Identification Strategy and Estimation **71**

6 Results **78**
6.1 The effect of retention on dropout . . . 78
6.2 The effect of repeating on academic outcomes . . . 81

7 Conclusion **85**

8 Appendix **88**

V Labor Market Training and Occupational Mobility **92**

1 Introduction **93**

2 Occupation switching, human capital and training **95**

3 Further Vocational Training Programs in Germany (2000-2004) **99**

4 Data, Sample Selection and the Definition of Occupation Switch **103**
4.1 General Database . . . 103
4.2 Sample Definition and Program Participation . . . 104
4.3 Definition of the Intention to Switch Occupation and Sample Descriptives . . 106

5 Empirical Strategy **111**
5.1 Identification . . . 111
5.2 Estimation . . . 115

6 Results **116**
6.1 The Outcomes . . . 116
6.2 Effects of Training . . . 117
6.3 Sensitivity Analysis . . . 123

7 Conclusion **125**

8 Appendix **128**

VI References **145**

VII Curriculum Vitae **155**

Summary

This dissertation contains four independent chapters from the fields of health, education and labour economics.

1. The first two chapters are concerned with the socio-economic gradient of overweight and obesity. The topic is motivated by the fact that increasing population shares of many countries in the world are overweight and obese, respectively, and that this trend poses a serious threat to societies due to the implied costs of the epidemic. In general, it is shown that - in line with existing theories - overweight is predominantly a problem of the elite at low stages of economic development, but trickles down to socio-economic classes as societies advance. In particular, this holds for the case of China (first chapter), but is also shown to be true when a larger set of countries is investigated (second chapter). It is argued that these findings have major implications for policy makers who are increasingly worried about the overweight epidemic.

2. The third chapter investigates the effect of retention on subsequent student outcomes using the University of St. Gallen in a case study. Knowledge about retention effects is valuable when (re-)designing education policies. Existing studies from the field of education economics provide evidence for the effect of retention at the level of primary and secondary education - however this is the first attempt to quantify the effects at the university level. For the particular example at hand it is found that retention leads only to modest increases in the drop-out probability of students. In addition, students show better grade performance as a result of being retained.

3. The fourth chapter analyzes the effectiveness of vocational training programs with respect to future labour market outcomes of the participants. While similar studies can be found in the literature, the special feature here is the focus on occupational mobility of individuals. Based on existing theories it is argued that unemployed workers who intend to change their occupation should benefit more from vocational training. However, the empirical results which are based on labour market data from Germany do not confirm this assumption.

Zusammenfassung

Diese Dissertation enthält vier unabhängige Kapitel aus den Bereichen Gesundheits-, Bildungs- und Arbeitsökonomie.

1. Die ersten beiden Kapitel befassen sich mit dem sozioökonomischen Gefälle von Übergewicht und Adipositas. Das Thema wird durch die Tatsache, dass ein zunehmender Teil der Bevölkerung vieler Länder der Welt übergewichtig bzw. fettleibig ist, und dass sich dieser Trend, aufgrund der immensen implizierten Kosten, eine ernsthafte Bedrohung für diese Gesellschaften darstellt. Insbesondere wird gezeigt, dass - im Einklang mit bestehenden Theorien - Übergewicht in unterentwickelten Gesellschaften in erster Linie ein Problem der Elite ist, sich im laufe ökonomischer Entwicklung aber zunehmend auf untere Bevölkerungsschichten ausweitet. Insbesondere gilt dies für den Fall von China (erstes Kapitel), trifft aber ebenso zu wenn eine größere Gruppe von Ländern untersucht wird (zweites Kapitel). Es wird argumentiert, dass diese Ergebnisse wichtige Implikationen für die politischen Entscheidungsträger mit sich tragen, die sich mittlerweile zunehmend dem Thema widmen.

2. Das dritte Kapitel untersucht den Effekt der Nichtversetzung und Wiederholung des ersten Jahres im Studium auf zukünftige Leistungen von Studierenden am Beispiel der Universität St. Gallen. Das Wissen über Auswirkungen von Nichtversetzung ist wertvoll bei der (Neu-) Gestaltung bildungspolitischer Maßnahmen. Existierende Studien aus dem Bereich der Bildungsökonomie liefern Erkenntnisse über die Auswirkungen von Nichtversetzung auf den Ebenen der Primär-und Sekundarstufe - dies ist jedoch der erste Versuch, die Auswirkungen auf universitärer Ebene zu quantifizieren. Die empirischen Befunde zeigen, dass eine Nichtversetzung lediglich zu einem mäßigen Anstieg in der "Drop-out" Wahrscheinlichkeit von Studenten führt. Darüber hinaus scheint das Nichtversetzen eine positive Auswirkung auf zukünftige Notenleistungen der Studierenden zu haben.

3. Das vierte Kapitel analysiert die Wirksamkeit von Aus-und Weiterbildungsprogrammen im Bezug auf zukünftige Arbeitsmarktergebnisse ihrer Teilnehmer. Während ähnliche Studien in der Literatur bereits existieren, ist die Besonderheit hier der Fokus auf die berufliche Mobilität der Einzelnen. Basierend auf bestehenden Theorien wird argumentiert, dass Arbeitslose, die ihren Beruf wechseln möchten, stärker von der derartigen Programmen profitieren sollten. Allerdings können die empirischen Ergebnisse der Studie, die auf Arbeitsmarktdaten aus Deutschland beruhen, derartige Annahmen nicht empirisch bestätigen.

List of Figures

2.A.1 CHNS Regions . . . 38
2.A.2 Logarithmic Income vs. Development Index by Wave . . . 38
2.A.3 Distribution of Development Index by Survey Year . . . 39
2.A.4 CHNS Trends: Per Capita Income over Time for Geographical Clusters . . . 39
2.A.5 CHNS Trends: Changes in Activity Levels over Time . . . 40
3.1 Concentration curves . . . 47
4.1 Grades, by type . . . 70
4.2 Histogram of the assignment variable . . . 75
4.3 RDD estimates: Probability of immediate dropout and starting a Bachelor degree . . . 79
4.4 RDD estimates: Credits accumulated by the end of each Bachelor semester . 82
4.5 RDD estimates: Grade point averages (GPA) by the end of each Bachelor semester (standardized) . . . 84
4.A.1 Time line: Institutional setup . . . 89
5.1 Employment probabilities . . . 118
5.2 Program effects for participants with switching intention: Employment probability in unsubsidized employment . . . 119
5.3 Program effects for participants without switching intention: Employment probability in unsubsidized employment . . . 119
5.4 Difference in program effects between participants with and without switching intention . . . 120
5.5 Successful switching probabilities . . . 121
5.6 : Difference in program effects (successful switching) between participants with switching intention and those without . . . 122
5.A.1 Distribution of occupation distances . . . 128
5.B.1 Common support for program participation of switchers – distribution of estimated propensity scores . . . 130
5.B.2 Common support for covariate adjustment of participants with switching intention to non-participants without switching intention – distribution of estimated propensity scores . . . 130
5.B.3 Common support for covariate adjustment of participants with switching intention to participants without switching intention – distribution of estimated propensity scores . . . 131
5.B.4 Program effects for participants with switching intention: Successful occupation switch . . . 131
5.B.5 Program effects for participants without switching intention: Successful occupation switch . . . 132

5.C.1Employment probabilities . . . 137
5.C.2Difference in program effects between participants with and without switching intention . . . 138
5.C.3Difference in program effects (successful switching) between participants with switching intention and those without . . . 138
5.C.4Program effects for participants with switching intention: Employment probability in unsubsidized employment . . . 141
5.C.5Program effects for participants without switching intention: Employment probability in unsubsidized employment . . . 141
5.C.6Difference in program effects between participants with switching intention and those without . . . 142

List of Tables

2.1 Means of Selected Key Variables by Gender, 1991-2006 15
2.2 Attrition and differences in BMI 20
2.3 Details of sample selection process and mean key variables, by gender 22
2.4 Community variables reflecting level of development 24
2.5 Means of Community Variables by Year 25
2.6 Average Development Rank by Province and Year 27
2.7 Average Development Rank by Geographical Clusters and Time 27
2.8 Development Category Transition Probabilities in % 28
2.9 Estimation Results by Gender and Development Level, Adults 18-60 34
2.A.1 Descriptive Statistics for estimation Sample, Adults 18-60 41
2.A.2 Factor Loadings of Community Variables 42
3.1 Sample Descriptives 48
3.2 Estimation Results for Proposition 1 (per capita GDP measured in thousands) 50
3.3 Estimation Results for Proposition 2 (per capita GDP measured in thousands) 51
3.A.1 Estimation Sample - Detailed Descriptive Statistics 56
4.1 Descriptive statistics: Entering first-year students, by type 67
4.2 Description of student types 69
4.3 Descriptive statistics: Estimation sample, by year 72
4.4 Descriptive statistics: Estimation sample, by minus credits 73
4.5 RDD Estimates: Pre-determined Characteristics 77
4.6 RDD estimates: Dropout and Major Choice 80
4.7 RDD estimates of Bachelor outcomes 83
4.A.1 Graduation statistics Switzerland 88
4.A.2 Capacity constraints at the university due to high amount of entering students. 88
4.A.3 Typical Curriculum of a Student 90
4.A.4 Descriptive statistics: Entering first-year students, by year 91
5.1 Further vocational training and other instruments of ALMP in Germany 2000-2004 100
5.2 Most frequent intended occupation switches in the sample 109
5.3 Intention to switch, actual switch, and program participation 110
5.4 Probit estimation of the intention to switch occupation on relevant covariates 112
5.5 Program Effects of FVT on cumulated labor market outcomes 124
5.A.1 Selected mean characteristics by sample subgroups 129
5.B.1 Results of probit estimations of the three covariate adjustments to the target group 'participants with intention to switch' 133
5.C.1 Most frequent intended occupation switches in the sample 139

5.C.2 Program Effects of FVT on cumulated labor market outcomes (sensitivity: distance measure) 140
5.C.3 Selected mean characteristics by sample subgroups 143
5.C.4 Program Effects of FVT on cumulated labor market outcomes (18 months participation window) 144

Part I

Introduction

Applied microeconometric research covers a broad range of subjects from different fields nowadays. Supported by major advances in computational power, as well as data availability, empirical analyses have become increasingly important in order to answer questions which are relevant to policy makers. If used adequately, informative databases paired with meaningful econometric methods can not only provide valuable insights about the underlying dynamics of real word phenomena, but they can (at times) also be used to draw counterfactual situations to be learned from. This thesis provides four examples of such research. It aims to make contributions to the fields of health economics, education economics and labour economics. Despite the fact that the chapters included are thematically diverse they all built on the availability of individual data. Since each of the chapters is going to be submitted to academic journals as an independent article they are self-contained and can be read independently of each other.

Research in Health Economics

Chapters one and two are concerned with the problem of rising overweight and obesity levels. Both provide empirical insights into the relationship between individual income and their weight, and changes thereof, with respect to economic development. They are largely motivated by the fact that obesity is known to significantly increase the spread of non-communicable diseases such as cardiovascular disease, diabetes, hypertension, musculoskeletal disorders and various cancers - thus creating significant costs for the economy. Understanding the determinants and dynamics of the obesity epidemic is a key concern for policymakers. Hence, it is not surprising that the rise of obesity in developed societies (especially in the US) throughout the last decades is one of the most intensively studied subjects in the recent health economics literature (see e.g. Cutler et al. (2003); Rosin (2008); Philipson and Posner (2008)) - and the debate is still ongoing. On the contrary, the overall rise of weight levels in developing countries has long been neglected by economists. At present, the World Health Organization (WHO) projects that by the year 2015 approximately 2.3 billion adults will be overweight and more than 700 million will be obese (Caballero (2007)) - there is no doubt that a significant share is contributed by individuals from developing countries.[1] While epidemiologists like Popkin (1994, 1999) have clearly spotted the ongoing "nutritional transition" in the developing world already some time ago, economic research

[1] A recent paper by Sahn (2009) strengthens this point. It shows that in a sample of more than 80 developing countries, if anything, there are only few cases which are not affected by significant increases in obesity levels.

(traditionally more concerned about the consequences of under-nutrition) that investigates the phenomenon of growing weight levels in low-income countries is still relatively scarce. The few existing papers (Fernald (2007); Case and Menendez (2009); Du et al. (2004)) commonly identify higher levels of obesity for the majority of developing countries, especially affected are females and more wealthy household.

A striking difference between obesity in developed and developing societies has first been pointed out in a meta-study by Sobal and Stunkard (1989) - and has later been updated and reconfirmed by Monteiro et al. (2004b) and Ball and Crawford (2005): While overweight is typically a problem of individuals of lower socioeconomic levels in developed countries (negative correlation between income level body-weight) the opposite is true for developing countries where overweight is relatively more widespread among individuals belonging to higher income classes (positive correlation between income level and body-weight). Empirical evidence for the latter can be found for the country specific cases of Mexico (Fernald (2007)), Egypt (Asfaw (2007)) and China (Popkin (1999)). Philipson and Posner (2003) nicely modeled this inverted U-shape claiming that in poor or early societies the more obese are relatively wealthier, but in wealthy, more modern societies the more obese are relatively poorer. The underlying idea is that starting from a situation where overweight is mainly a problem of higher income classes, with economic development it trickles down to lower income classes. Extrapolating from that latter situation, the study of Philipson and Posner (2003) implies that once even the lowest income ranks in a (developed) society achieve a certain standard of living, the self-limiting nature of the overweight problem will lead to an absolute decrease of the epidemic. However, until recently no study had ever investigated the income body weight gradient within a developing country over time.

The first chapter of this dissertation (single authored) uses the case of China to investigate the issue in-depth. Contrary to a framework where the gradient of interest would be compared between countries at different stages of economic development, the within country approach taken in this dissertation chapter has the advantage of being subject to a more equal institutional setting as it exploits pure within country differentials of development. The data used comes from the China Health and Nutrition Survey (CHNS).[2] The CHNS is longitudinal and, thus, allows a dynamic analysis rather than just a snapshot of a single moment in time.

In particular, China is an interesting case to investigate that relationship mainly for two reasons: First, China's economy grew at an unprecedented annual rate of about 10% over the last two decades and, second, this development was rather unequally distributed over the country (coastal regions were the main beneficiaries of economic growth (Yang

[2]The CHNS is not nationally representative, but only covers nine provinces: Guangxi, Guizhou, Heilongjiang, Henan, Hubei, Hunan, Jiangsu, Liaoning, and Shandong. For more information about the regions covered by the CHNS and its specific survey design see Liu (2008) or visit www.cpc.unc.edu/projects/china.

(1999))).[3] The two sources of variation in development are utilized in the paper - cross-sectional and over-time - to study changes in the income body weight gradient. Parallel to economic development, China has experienced tremendous increases of overweight levels - especially in urban areas (see Popkin et al. (2006)). Ng et al. (2009) provides evidence that the nutrition transition in China is currently progressing at a fast pace with already more than 20% of the Chinese being overweight and another 3% being obese nowadays - translating into more than 200 million Chinese being affected by the epidemic in absolute terms.

The analyses use multidimensional community level data provided by the CHNS in order to characterize relative regional development levels. Similar to Van de Poel et al. (2009) and Jones-Smith and Popkin (2010) factor analysis is used to construct a one-dimensional development index from the community variables that is assumed to describe the relative level of economic development of a community or region, respectively. Based on the index sub-samples of communities are formed which are assumed to be at a comparable level of economic development. Consequently, gender specific weight growth regressions are conducted in each of these strata where individual´s annual growth rates in their BMI between CHNS waves are regressed on their income in the base year as well as a battery of potential confounders. As predicted by theory, the results show a positive income gradient of BMI growth for the economically least developed strata, i.e. within those areas individuals at higher income ranks show the largest BMI growth rates, ceteris paribus. On the contrary, when looking at the results for the most developed areas, a negative relationship between income and BMI growth is found (for females). These patterns concur with the theoretical underpinning of Philipson and Posner (2003).

The second chapter (joint with Sofie J. Cabus and Eva Deuchert) investigates the topic further and should therefore be seen complementary to the first chapter. However, unlike the first chapter that draws on findings from one country, the second chapter takes a more comprehensive stance using individual level databases of females aged 15-49 from more than 50 developing countries. These data are cross sections and originate from the Demographic and Health Surveys (DHS)[4]. Given the similar (to identical) survey designs these data very comparable across countries and should, thus, assure a high degree of consistency throughout the analyses. Due to the lacking of pure income measures in the DHS, this chapter uses summary measure for wealth (the so-called wealth index) to approximate an individuals relative relative socio-economic status. Using regression analyses, once again, two of the main predictions of Philipson and Posner (2003) are tested - and found to be supported by result. First, it is shown that there is a positive and significant relationship between a

[3] Heilig (2006) documents in a global comparison that in terms of regional GDP per capita some Chinese provinces are comparable to industrialized countries while others are still at levels of low-income countries.

[4] For more information about the DHS visit http://www.measuredhs.com/.

countries level of development (this time approximated its level of GDP) and its mean BMI, overweight and obesity level, respectively. Second, building upon concentration indices to approximate the socio-economic gradients of overweight, regression analyses confirm that a higher level of GDP is associated with a falling value in its concentration index meaning that overweight is trickling down to lower income ranks. Stated differently, the level of economic development also determines the distribution of body weight within countries. The chapter concludes that, if the observed trend continues, the obesity epidemic may further burden health systems in developing countries.

Research in Education Economics

The third chapter (joint with Sharon Pfister and Petra Thiemann) relates to the field of education economics and investigates the causal effects of retention policies on educational outcomes of students in higher education. The practice of grade retention is subject to an ongoing debate in many countries, partly because the net effect of retention on students' outcomes is difficult to assess. Following theoretical arguments, outcomes of students can either be positively or negatively influenced by retention: positive effects might include learning gains from repetition as well as improved confidence as students can cope better with their performance requirements while negative effects might occur through retarded learning, low aspirations, stigmatization and necessary adjustments to a new classroom environment. Therefore, finding the net effect of retention on student outcomes ultimately remains an empirical question.

The effects of retention on the level of primary as well as secondary education have already been addressed in the economic literature (see, for example, Greene and Winters (2007), Jacob and Lefgren (2004), Jacob and Lefgren (2009), Manacorda (2012)). However, retention policies also exist at the level of post-secondary education, but the effects on student outcomes might be different from the effects found on the level of primary or secondary education. For example, stigmatization or re-adjustment to a new environment might play a less important role at university. Therefore, the results from existing studies do not directly translate to the higher education context. Since an assessment of retention effects in post-secondary education is still missing in the literature this chapter aims to fill that gap.

Data of six cohorts (for the years 2001 - 2006) of freshmen at the University of St. Gallen, Switzerland, provide a rich source in order to address the following two research questions: First, what is the effect of retention on the probability of dropout from university (Manski (1989))? Second, what is the effect of first-year repetition on academic performance (grades, credits per semester)? The latter outcomes give not only an indication of students' human capital, but might also serve as signals to potential employers. Furthermore, addressing the former question is necessary to quantify the fraction of students opting out due to retention, which might strongly influence their educational outcomes as well.

To overcome selection problems when simply comparing outcomes of retained and not-retained students the effects of retention on later educational outcomes are identified using a sharp regression discontinuity design (Imbens and Lemieux (2008)). A clear-cut retention threshold at the University of St. Gallen divides the students close to a threshold quasi-randomly into treated and control groups. The threshold is based on academic performance during the freshmen year, i.e. students are retained based on a measure that accounts for insufficient performance across all first-year courses. Retained students then either have to repeat all first-year courses and exams, or to leave the university. All non-retained students in turn can directly proceed to the second year.

The results are as follows: Retention significantly increases individual dropout at the threshold - however, the share of students who drop out as a result of being retained is surprisingly small. Moreover, repeating the first year has a positive and weakly statistically significant effect on educational achievements. The grade point average (GPA) of the retained students by the 4th Bachelor semester is somewhat higher if being retained. Moreover, no significant effects on the number of credits obtained in each semester exist. Overall, these results suggest that grade repetition at university can be - if anything - modestly beneficial from a student's perspective in terms of academic success.

Research in Labour Economics

The fourth chapter (joint with Benjamin Schuenemann) relates to the field of labour economics. It evaluates the effectiveness of further vocational training (conducted between 2000 and 2004) on future labour market prospects of unemployed workers in Germany. Until today, training programs for unemployed workers are a major component of Active Labor Market Policy (ALMP) in many OECD countries. At the same time, training programs are costly, thus, motivating researchers to analyze their effectiveness (see e.g. Lechner and Wunsch (2009) or Kluve (2010) for an extensive overview). As summarized in a comprehensive meta-analysis by Card et al. (2010) most training programs exhibit “modestly positive effects”.

The chapter differs from most of the existing literature in that field as it explicitly considers occupational mobility of workers throughout the analysis – i.e. (1) unemployed workers who intend to change their occupation are more likely to select into training and (2) training itself may induce occupation switches. Following the discussion of Kambourov et al. (2012) we agree that the existing literature has largely overlooked a key characteristic of unemployed workers: occupational mobility. Traditionally, evaluation studies use econometric reweighting methods together with rich observational data to estimate the effects of program participation on individuals' labor market outcomes. When establishing evidence, however, evaluation studies usually neglect that participants and non-participants are often characterized by differences in occupational mobility where the former have higher rates than the latter. The chapter provides a in-depth discussion on why differences in occupational mobility need to be accounted for explicitly when the outcomes of participants and non-participants are compared. Moreover, training participation itself may induce occupational mobility - in particular for those who did not intend to change in the first place - we provide evidence that this is indeed the case.

The study is based on data from the Integrated Employment Biographies (IEB) - a compilation of administrative labour market data from different sources for a 2%- sample of the German labour force.[5] To correct for differential mobility patterns between participants and non-participants this study utilizes a unique feature of the data, i.e. information about individuals' target occupation measured at the very beginning of unemployment and, hence, before any decision about training participation is made. Using this feature we can determine the intention to switch occupation without conditioning on being employed and, thus, split occupation switching into an ex-ante (exogenous to training) and an ex-post (endogenous to training) part.

The purpose of this chapter is to investigate mainly two things. First, we ask whether training programs are differently effective for those who want to change their occupation compared to those who do not want to. This is relevant since any significant differences between the two groups would call for reforms towards a more intensive profiling of job seekers with a special focus capturing switching intentions of unemployed workers. Secondly, we analyze the effect of training participation on future patterns of occupational mobility. A further advantage of our data is the possibility to investigate a long outcome horizon. Early studies as summarized in Heckman et al. (1999) could rarely detect any effects other than the well-known lock-in effect (Van Ours (2004)) due to short outcome horizons. Recent evaluations provide a more positive view on program impacts. They show that in the medium-term training programs can increase the employability of participants implying that the post-treatment observation period should be sufficiently long to obtain convincing results.

Using microeconometric matching methods (see Huber et al. (2013)) we estimate the impact of further vocational training on subsequent employment, successful occupation changes and earnings for up to four years after training assignment - separately for those with and without switching intentions. Moreover the respective outcome differences across the sub-groups are tested for mean inequality. The estimated effects show positive impacts of training on future unsubsidized employment for both subgroups - i.e. higher employment probabilities from about three years after program start onwards. With respect to successful

[5] The data is provided by the German Institute for Employment Research (IAB). For further information visit http://fdz.iab.de/en.aspx

occupation switching it is found that those who intend to move to a new occupation are more likely to be employed in a different occupation then their last one due to training participation. This effect of training is even more pronounced for participants without initial switching intention as they are about three months more employed in a new occupation than without training. However, little effect heterogeneity is found (also due to relatively imprecise estimates) - thus we cannot conclude that unemployed most in need of human capital adjustments, i.e. individuals who intend to change their occupation, do benefit more from training than unemployed workers who look for a job in their old occupation.

Part II

The income body weight gradients in the developing economy of China

Abstract

Existing theories predict the income gradient of individual body weight to change sign from positive to negative in process of economic development. However, only few empirical studies support the hypothesis. This paper adds to the literature on that topic by investigating the case of China using data from the 1991-2006 waves of the China Health and Nutrition survey. Using a one-dimensional measure to characterize the level of economic development of a region, regression analyses indicate that higher income is related to larger future growth of individuals' BMI in less developed areas whereas it lowers BMI growth in more developed areas. The switch is somewhat more pronounced for females.

1 Introduction

Overweight induced by excess body fat poses a serious threat to individual health as it significantly increases the risk of non-communicable diseases such as cardiovascular disease, diabetes, hypertension, musculoskeletal disorders (especially osteoarthritis) and various cancers (endometrial, breast and colon). On this account, it is not surprising that the rise of overweight in developed societies throughout the last decades and its potential effects on public health are intensively studied subjects in the recent health economics literature (see e.g. Cutler et al. (2003); Rosin (2008); Philipson and Posner (2008); Bleich et al. (2008)) - and the debate is still ongoing. Determinants identified are technological progress (and directly related more sedentary lifestyles) (Lakdawalla et al. (2005)), changes in food prices (Schroeter et al. (2008)), increased fast-food availability and cigarette taxation (Chou et al. (2004); Currie et al. (2010)), sugar sweetened beverages (Pereira (2006)) and bounded individual knowledge about the potential health consequences (Kan and Tsai (2004)) of overweight.

On the contrary, rising weight levels in developing countries have long been neglected by economists. The World Health Organization (WHO) projects that by 2015 approximately 2.3 billion adults worldwide will be overweight and more than 700 million will be obese (Caballero (2007)) - a significant share is contributed by individuals from developing countries. Sahn (2009) strengthens this point by showing that there are few countries which are not affected by rising weight levels over the last decades.[6] While epidemiologists like Popkin (1994, 1998, 1999) have clearly spotted the ongoing "nutritional transition" in the developing world already some time ago, economic research (traditionally more concerned about the consequences of under-nutrition) investigating the phenomenon of growing weight levels in low-income countries is rather recent. Piecewise evidence on the main determinants has been delivered by some early studies (Galal (2002); Fernald (2007); Case and Menendez (2009); Du et al. (2004)). In a nutshell, these papers commonly find rising levels of overweight for the majority of developing countries, especially affected are females and more wealthy household. Somewhat surprising, papers by Doak and Popkin (2008) and Sahn and Younger (2009) show that there is a growing number of households in which underweight and overweight individuals coexist. The sneaky nature of the epidemic has prompted policy makers to shelve the issue for a long time, but recently there has been a notable increase in the awareness for the urgent need of counter measures (see Gortmaker et al. (2011)).

A striking difference between the general patterns of overweight in developed and developing societies has first been pointed out in the seminal meta-study by Sobal and Stunkard (1989) - and has later been updated and reconfirmed by Monteiro et al. (2004b): While

[6]In his sample of 30 developing countries, 17 show first order statistical dominance when their intra-country weight distribution is compared to past weight distributions. In addition, second order dominance is observed in 11 cases.

overweight is mainly concentrated among poorer individuals in developed countries (i.e. there is a negative correlation between income level body-weight) (see Ball and Crawford (2005)), the opposite is true for developing countries where overweight is relatively more widespread among individuals who belong to higher income classes (i.e. there is a positive correlation between income level and body-weight). Empirical evidence for the latter can be found for the country specific cases of Mexico (Fernald (2007)), Egypt (Asfaw (2007)) and China (Popkin (1999)). Overall, a shift of overweight concentration from the rich to the poor should only be observed if there is a non-monotonic relationship between income and weight gain.

Philipson and Posner (2003) and Lakdawalla and Philipson (2009) modeled the implied inverted U-shape in a dynamic framework of weight management - their model is as the workhorse model in most of the obesity literature until present. Explicitly accentuating the reversal hypothesis, one of its major implications is that *in poor or early societies the more obese are relatively wealthier, but in wealthy, more modern societies the more obese are relatively poorer.* In fact, the few existing empirical studies (e.g, Sarlio-Lahteenkorva et al. (2004)) support the overall presence of a hump shaped income-body-weight relationship. The idea is that starting from a situation where overweight is mainly a problem of higher income classes, with economic development it trickles down to lower income classes. The underlying dynamics are simply described. Individual body-weight is determined by the relative ratio of energy intake to energy expenditure and an increase (decrease) of this ratio, ceteris paribus, will lead to an increase (decrease) in weight. Economic development decreases caloric cost through a reduction in food prices and increases the cost of caloric expenditure through the more sedentary nature of jobs. The role of additional income, is, however, ambiguous. In less developing societies, where most individuals work in physically demanding occupations, food is scarce and body weight is typically is low, one would expect individuals to use additional income to increase the amount of calories consumed (among other things) - and thereby gain weight, ceteris paribus. Here, overweight is only “affordable” to relatively richer individuals who have wider access to food and are more likely to work in physically less demanding jobs, respectively. There are, however, situations where one might expect additional income to be associated with a reduction in weight. For example, Schroeter et al. (2008) argues that a reduction should be observed for heavier individuals who use additional income to substitute a high-caloric diet by a low-caloric one. In this case, calories (e.g. from staple foods) would be a quasi-inferior good and more income would lead to lower (or negative) weight growth (Jensen and Miller (2008)). A similar argument holds with respect to energy expenditure where one might imagine a latent demand for thinness that leads to increasing levels of voluntary and, at times, costly activities (e.g. sports) once a certain income level is achieved. Such behaviours are more likely observed in developed societies where the combination of abundant food availability and mostly sedentary jobs gives rise to concerns of weight control in large parts of the population, in particular among the wealthier parts. Considerations like these have led economists to believe that the growth of overweight and obesity at the population level may be self-limiting

While piecewise evidence supports the hypothesis of an income-body-weight gradient that changes from positive to negative with economic development, no study has ever shown such a relationship empirically in one go. This paper contributes to the discussion by testing the predicted patterns of the reversal hypothesis, i.e. providing empirical insights of how the relationship between income and body-weight changes with economic development. The analysis is carried out using the illustrative example of China and, in particular multiple waves of the China Health and Nutrition Survey (CHNS). In simplified terms, differences in the association between income and changes in body weight are investigated and compared across regions with varying levels of economic development. Thereby, the level of development is approximated by a one-dimensional index that, nevertheless, accounts for several dimensions such as regional infrastructure, available services, labour market structure, etc.. Contrary to a framework where different countries are compared, the within country approach taken here has the advantage of being subject to a harmonized institutional setting. Moreover, the use of longitudinal information allows a more dynamic analysis as compared to just looking at snapshots of a single moment in time - in that sense, it is more in the spirit of existing theoretical models of weight management.

In particular, the Chinese case is well-suited to investigate the linkages between economic development, income and body weight for several reasons. Following significant economic reforms in 1979, China's GDP grew at an unprecedented average rate of 10% per annum. At the same time, this growth was rather skewed towards coastal provinces while provinces in inner China were left behind dramatically (Yang (1999)) which has created significant heterogeneity in development levels across provinces. Furthermore, China has experienced a continuous increase in weight levels - especially due to the rise of overweight in urban areas (see Popkin et al. (2006) and Zhu and Jones (2010)) - which translates into more than 20% of the Chinese being overweight and another 3% being obese at present (Wu (2006), Levine (2008)). The duality of both, tremendous (but unequal) economic growth and significant increases in body weight patterns, provides an interesting set-up to investigate if theory is right in predicting the role of income in weight management to differ across different stages of economic development.

In short, the main findings of this paper are in line with the reversal hypothesis of the Philipson and Posner (2003) model. Following the main predictions, body weight levels are higher among higher income sub-groups and in the more developed areas of China. At the same time, lower income groups and less developed areas are catching up - this trickling down appears to reverse the socio-economic gradient of overweight. Looking at the micro-data, per capita income levels of Chinese adults are found to be positively associated with subsequent growth in body weight when focusing on individuals from the least developed areas, but the association is non-significant for males and significant and negative for females when only the most developed areas are considered. It needs to be acknowledged from the very beginning, however, that due to the lack of long-lasting exogenous variations in household income the estimates lack a causal interpretation, but, instead, reflect conditional associations. Having said this, the analyses deliver new insights that are relevant from a policy point of view. First, they provide a comprehensive picture about the evolution and the current state of overweight in the Chinese context as they are based on high quality data, apply to a large share of the country's adult population (i.e. those represented by the CHNS) and reflect a rather long time horizon. Second, the patterns that emerge can help to guide policy makers to better identify sub-groups in the population that are most susceptible to the disease, predict its expected diffusion and, thus, help to design counter measures (e.g. information campaigns) more efficiently. Third, the results are relevant with respect to the evolution of socio-economic health inequalities in developing countries, here particularly in China, as a trickling down of obesity to lower income groups is likely to trigger bad health states among the poorest parts of the population and, thus, increase socio-economic health inequalities (Monteiro et al. (2004a)).

In what follows, section two will provide a broader context by discussing the patterns of China's recent economic growth as well as trends in body weight, nutrition and physical activity levels. First descriptive investigations will show how weight levels vary across income terciles and different levels of development, respectively. After that, section three describes the data source used in this paper, namely the CHNS and how the estimation sample is constructed. Section four explains how community data taken from the CHNS are used to conceptualize economic development. In particular, it is illustrated how a one dimensional measure that aims to characterize individuals' living environments at the time when surveyed is created and validated. Section five presents a simple econometric set-up for the analysis at the level of individuals. The sample is divided according to different strata of economic development and weight growth equations are estimated based on observable characteristics. After a discussion of the results, section seven concludes this paper.

2 Background

2.1 China's Economic Development

China was economically frail for well over the first half of the twentieth century and hit the rock bottom between 1957 and 1962 when - known as the *Great Leap Forward* - governmental policies to boost industrial growth resulted in the worst (documented) famine in the history of mankind.[7] During this period food production was clearly insufficient.

[7] Jisheng (2010) estimates this devastating famine to have claimed 36 million lives!

Correspondingly, caloric intake was generally very low throughout China and while under-nutrition was wide-spread, little evidence about increases of overweight is known. For most of that time there was limited access to progressive technologies and occupations (e.g. in agriculture or heavy industry) were physically demanding - leaving no space for excess body weight. While the economy remained at low levels throughout that period, it was just in 1978 when broad structural reforms opened the path for future economic development of the country.[8] Subsequently, China's GDP grew on average 9.6% per year between 1978 and 2005 - and growth still persists at comparable levels to date (10,3% in 2011).

Yet, the development process has been distributed unequally over the country.[9] In addition to the long-standing lead of urban over rural areas, it is a stylized fact that the main drivers of China's recent economic growth were its coastal regions (see Holz (2008)). Figure 2.A.4 provides descriptive evidence about the rise of income inequality between geographical areas over time where CU, NCU, CR, and NCR represent trend lines for mean per capita income levels in coastal and non-coastal, and urban and rural areas.[10] By now there is a vast literature that aims to measure and explain the drift between inland and coastal provinces which basically escalated from the mid 1980s onwards (see Bramall (2009) for a summary). In a nutshell, two eminent determinants have been identified (Kanbur and Zhang (2005)). First, as a key component of its Coastal Development Strategy (started in 1984) the government followed a so-called "open-door" policy that aimed to attract foreign direct investments (FDI) and stimulate international trade. More specifically, this regime comprised the set-up of special economic zones in coastal areas to stimulate exports, preferential allocation of resources, improvements of the local infrastructure and favorable tax treatment for investors. There is strong evidence that these measures created an important momentum for economic growth to accelerate in coastal provinces while leaving non-coastal provinces unaffected (Bhalla et al. (2003)). Attracted by the benefits as well as by the low level of labour costs and geographic advantages numerous national and international firms invested in new businesses and production sites. These (mainly) trade related activities generated significant labour absorption from agriculture and heavy industries into manufacturing and light industries, thus leading to significant restructuring of the local labour market in coastal areas. Although the open-door was extended to provinces in inner China in the mid 1990s, it did not quiet reach the same effect (Yang (2002)).

Second, fiscal decentralization measures weakened the financial capacity of the Chinese

[8]One key element of the 1978 reforms was the liberalization of food production. Moreover, foreign investments were stimulated.

[9]Heilig (2006) uses global comparisons to document that in terms of regional GDP per capita some Chinese provinces are on a par with industrialized countries while others are still at levels of low-income countries - thus suggesting a high degree of heterogeneity in the level of economic development across Chinese provinces.

[10]In this paper the group of coastal provinces consist of Liaoning, Jiangsu and Shandong while non-coastal provinces are represented by Heilongjiang, Henan, Hubei, Hunan, Guangxi and Guizhou. See the data section for a corresponding map.

government and, thus, lowered its ability to redistribute tax money across provinces for equity purposes. If anything, starting in the early 1980s, government tax resources were incrementally channeled towards coastal provinces which augmented the subsequent geographic divide in economic growth. Moreover, to incentivize future non-agricultural investments, the agricultural sector was often subject to heavy taxation. Unlike expected, this often had a crippling effect on overall economic activity, especially in inland provinces where the agricultural sector is relatively larger (Hao and Wei (2010)). Of course, that heterogeneity is also closely mirrored by a comparison of work related physical activity levels across regions (see Figure 2.A.5). As a direct result of these dynamics, coastal and non-coastal provinces in China are nowadays significantly different, not only in terms of income levels, but also with respect to their labour market characteristics (i.e. more agriculture and heavy industry in non-coastal regions, more light industry and services in coastal regions). These differences should be kept in mind as major confounding factors for the later analyses.

2.2 The Evolution of Overweight in China

China's economic growth was also accompanied by negative health effects at the population level (Zhu and Jones (2010) and Van de Poel et al. (2009)). As predicted by the theory of Philipson and Posner (2003) technological progress came along with significant increases in average body weight which is now a major concern to policy makers. Alarmingly, Popkin et al. (2006) estimate the future health cost of the overweight epidemic (and direct consequences thereof) to reach 9% of China's GDP by 2025. A WHO report finds that China could loose $558 billion of national income till 2015 due to the spread of diabetes and heart disease - both illnesses are known to be closely related to excess body weight. The main drivers of this increase are reduced physical activity and changing lifestyles (Ng et al. (2009)) as well as shifts towards western diets (Du et al. (2004)). Table 2.1 provides descriptive insights about the process by showing gender-specific mean trends of body size measures, nutritional behaviours and physical activity levels.[11] As common in the obesity literature, this paper uses the body mass index (BMI) measure (i.e. body-weight relative to squared height (kg/m^2)) to describe individual body shapes. By WHO standards an individual is considered as overweight or obese if BMI exceeds 25 or 30, respectively.

Looking at broad averages the increase of body weights is inevitable. For both, females and males, average BMI have risen continuously throughout the 15 years under consideration - for the latest years, mean BMI was well above 23. While the averages are still about 3-4 points lower than those for developed countries (e.g. Germany or the U.S.) in absolute terms, BMI growth rates in China are comparably higher (see Popkin (2008)). Correspondingly, the percentages of overweight and obese females and males have risen too. In 2006, close

[11] The means are calculated based on observations of CHNS - a detailed description of the data is provided in section 3.

Table 2.1: Means of Selected Key Variables by Gender, 1991-2006

	Females 18-60						Males 18-60					
	1991	**1993**	**1997**	**2000**	**2004**	**2006**	**1991**	**1993**	**1997**	**2000**	**2004**	**2006**
BMI	22.16	22.19	22.55	22.93	22.95	22.92	21.58	21.79	22.20	22.69	22.95	23.13
Overweight (BMI>25)	0.16	0.16	0.20	0.24	0.25	0.24	0.10	0.12	0.16	0.22	0.24	0.26
Obese (BMI>30)	0.02	0.01	0.02	0.03	0.04	0.03	0.01	0.01	0.01	0.02	0.03	0.03
Calorie Intake												
kcal	2550 (700)	2473 (734)	2232 (626)	2224 (734)	2162 (698)	2069 (676)	2945 (829)	2838 (817)	2606 (714)	2590 (790)	2522 (810)	2458 (779)
fat (g)	63 (34)	65 (35)	65 (35)	71 (37)	70 (37)	67 (36)	69 (37)	72 (38)	73 (37)	80 (40)	79 (40)	76 (39)
protein (g)	73 (22)	72 (23)	65 (20)	64 (23)	64 (23)	64 (24)	84 (25)	83 (26)	76 (24)	75 (25)	74 (26)	75 (26)
carbohydrate (g)	419 (134)	394 (135)	345 (116)	322 (114)	313 (111)	302 (114)	471 (143)	445 (146)	399 (135)	373 (122)	361 (124)	353 (124)
Physical Activity												
Employed	0.85	0.84	0.80	0.75	0.63	0.64	0.94	0.92	0.90	0.86	0.78	0.78
Light Activtiy	0.28	0.28	0.35	0.39	0.43	0.45	0.24	0.22	0.28	0.29	0.33	0.35
Moderate Activity	0.15	0.15	0.15	0.13	0.18	0.15	0.19	0.22	0.20	0.20	0.22	0.22
Heavy Activity	0.55	0.56	0.48	0.47	0.37	0.38	0.55	0.55	0.50	0.48	0.43	0.41
No Activity	0.02	0.01	0.02	0.01	0.02	0.02	0.02	0.01	0.02	0.03	0.02	0.02
1st Income Tercile												
BMI	21.68	21.84	22.22	22.66	22.82	22.97	21.05	21.47	21.76	22.29	22.39	22.76
Overweight (BMI>25)	0.12	0.14	0.17	0.21	0.23	0.25	0.05	0.09	0.1	0.16	0.17	0.21
2nd Income Tercile												
BMI	22.2	22.14	22.49	22.94	23.06	23.03	21.59	21.69	22.11	22.47	22.79	22.98
Overweight (BMI>25)	0.18	0.15	0.2	0.24	0.26	0.26	0.11	0.1	0.16	0.19	0.22	0.23
3rd Income Tercile												
BMI	22.59	22.58	22.94	23.13	23.03	22.79	22.12	22.2	22.72	23.27	23.68	23.65
Overweight (BMI>25)	0.21	0.21	0.23	0.26	0.26	0.23	0.15	0.16	0.22	0.29	0.33	0.33
Non-Coastal Rural												
BMI	21.54	21.75	22.02	22.5	22.53	22.62	20.93	21.25	21.7	22.2	22.44	22.73
Overweight (BMI>25)	0.11	0.12	0.15	0.19	0.2	0.21	0.5	0.7	0.1	0.16	0.18	0.2
Non-Coastal Urban												
BMI	22.29	22.3	22.76	22.99	22.85	22.66	21.63	21.91	22.32	22.55	22.92	22.91
Overweight (BMI>25)	0.18	0.17	0.22	0.25	0.24	0.21	0.12	0.13	0.19	0.21	0.24	0.23
Coastal Rural												
BMI	22.75	22.54	23.21	23.38	23.61	23.63	22.28	22.29	22.81	23.22	23.52	23.73
Overweight (BMI>25)	0.21	0.2	0.27	0.28	0.32	0.32	0.15	0.15	0.22	0.26	0.3	0.34
Coastal Urban												
BMI	22.9	22.91	23.76	23.54	23.49	23.23	22.6	22.66	23.62	23.82	24.03	24.03
Overweight (BMI>25)	0.25	0.24	0.32	0.34	0.33	0.3	0.17	0.21	0.28	0.34	0.37	0.38

Note: Own calculations based on CHNS data. All trends shown in the table are age adjusted. Standard errors of measures related to caloric intake are presented in brackets.

to one-fourth of adults aged 18-60 were to be classified as overweight. Given the large population of China, overweight rates of this order translate into more than 200 million adults being affected by the epidemic at present.

Yet, rises in BMI wrongly suggest daily calorie intake to have increased over the same time span. Surprisingly, the numbers contrarily reveal that mean caloric intake (kcal) during the same time has actually *decreased*.[12] Roughly speaking, the data indicate that both sexes on average consume about 500 kcal less per day which should be interpreted as a result of changing caloric requirements.[13] Parallel to reductions in total caloric intake, the relative composition in individuals' diets have clearly shifted away from proteins and carbohydrates towards foods rich in fat. This is equally true for females and males, respectively. In fact, China has traditionally been regarded as a country with a relatively lean population that generally featured a healthy diet rich in cereals and vegetables. This, however, appears to be changing rapidly. The numbers suggest that the country is undergoing a nutritional transition at high speed in which its traditional diet is continuously giving way to more western diets characterized by animal foods and higher shares of edible oils.

Observing increasing body weights while actual caloric consumption dropped suggests that caloric expenditure (i.e. physical activity) must have decreased disproportionately.[14] In fact, this is what can be observed when examining indicators related to work related energy expenditure. The percentage of employed individuals has steadily decreased over time. Furthermore, employment shares are consistently lower for females than for males. This is in parts due to longer spells in education.[15] Besides education an increasing share of individuals reports to seek work or - especially in the case of females - do housework activities. Further insights can be gained by looking at indicators for physical activity at

[12]The numbers are based on food consumption records at both the household as well as individual level. Consumption was recorded for a period of three consecutive days for each household (start day was randomly allocated from Monday to Sunday). At household level, inventories of food were carefully weighed and recorded at the start of each day. All food purchases, home production, and snack preparations were recorded. Waste was weighed when possible and estimated when not. Inventories were also recorded at the end of the day, and the change in inventory over the day enabled estimation of household food consumption. In addition, individual food consumption was separately estimated for each of the same 3 days by 24-h recall, including both home consumption as well as food eaten away from home. The availability of household records as well as individual consumption allows for further checks of data quality and adjustment for errors. The CHNS dietary data are considered of high quality due to the rigor in their collection and the triangulation procedures followed. The 1991 Food Composition Table (FCT) for China was utilized to calculate macro-nutrient intake values for the dietary data of 2000 and previous years (Institute of Nutrition and Food Hygiene, 1991). A new 2000 version of the FCT (Institute of Nutrition and Food Hygiene, 2002) was used for 2004 and 2006 surveys, and updated for new foods each year. Note that the switch between FCTs may limit the comparability across years.

[13]This is in line with findings of Lakdawalla and Philipson (2009) for the U.S. who find constant (or even decreasing) levels of caloric intakes over a period of significant increases of BMI levels.

[14]Scholars are in agreement that genetic changes in humans play -at most- a marginal role in explaining recent upward trends in body weights around the globe (see e.g. Popkin (1994)). Thus, the only reasonable explanation when observing increasing body weights while caloric consumption has decreased is a disproportionate decrease in caloric expenditure.

[15]While only 8% of the 18-26 year old claimed to be in education in 1991 the corresponding number more than doubled to 17% in 2006 (not shown in table).

the workplace where three main categories are presented. Light activity includes jobs that are generally less physically demanding (i.e. sedentary jobs (e.g. office workers), teacher, laboratory technician, etc.). Moderate activity is defined to include occupations such as electrician, metal worker and salesman. Heavy activity, eventually, incorporates jobs which are physically demanding, e.g. farmers, steel workers, loaders, miners or stone-cutters. The time trends in these three categories are unambiguous: there is a marked shift away from heavy physical activities towards more light and sedentary ones. As with employment status, the shift is more pronounced for females where, by 2006, nearly half were classified to follow a light activity. In line with the findings of Popkin (1999) these changes should be interpreted as sectoral reallocation brought about by China's economic development, i.e. occupations are incrementally shifting from the agricultural and manufacturing sectors to the service sector (jobs here are often less physically demanding and sedentary). Hence, lower levels of work related physical activity appear to be the main driver of increasing weight levels. Following the arguments of Philipson and Posner (2003) this is most likely due to the fact that caloric expenditure has become relatively more expensive over time. Note, however, that these measures exclude any time of physical activity outside the workplace which is, of course, an important dimension of overall caloric expenditure.

Investigating mean body weight measures by income terciles reveals further interesting insights. Apparently, back in 1991 there was a rather large disparity in overweight rates between lower and higher income quintiles. Average BMI and shares of overweight increased across income quintiles and both measures were consistently higher for females at that time - overweight rates were roughly twice and three times higher in the highest tercile than in the lowest for females and males respectively. As mentioned, overall weight levels rose in the following years, but for females the differences in average BMI and overweight rates across income terciles vanished. By the end of the observation period overweight rates for females have been homogeneously distributed across income terciles. This resulted from two dynamics: weight levels rose at fast pace in lower income quintiles while they only grew slowly (or even stagnated and recently dropped) in the upper parts of the income distribution. Thus, while overall levels increased, there has been a shift of overweight concentration from higher income quintiles to lower ones for females. Naive extrapolation of that trend would suggest that in the future overweight levels will be highest for females in lower income quintiles. Yet, the situation is different for males. Here, too, all parts of the income distribution show considerable growth in average BMI and overweight over time. However, contrary to females, weight levels in males grew more heavily for higher income quintiles. By 2006 one-third of males at the top end of the income distribution were overweight - while in the lower quintiles this was only true for about 20%.

The lowermost panel of the table displays body weight statistics by geographical clusters which are -as discussed above - assumed to coarsely reflect the varying levels of economic development. As with different income quintiles, significant variations in mean weight levels are visible. Keep in mind, that these are confounded by differences in the labour market structures across regions as well as by differences in income levels. Throughout the time span under consideration coastal areas, and especially urban areas, show higher mean BMI and, likewise, higher levels of overweight than their non-coastal counterparts. Again, this is true for males and females. By the year 2000 (and beyond), overweight prevalence in coastal urban areas was >30% for both genders. Inevitably, there has also been a pronounced growth in BMI levels in non-coastal provinces, but their absolute levels in 2006 were just as high as those of coastal areas in the beginning of the 1990s. Yet, investigating *changes* rather than absolute levels, growth rates have lately been most pronounced the less developed areas. Overall, these patterns suggest that overweight and obesity, while they were mainly a problem of higher income groups and more developed areas, are continuously trickling down to lower income groups and less developed areas.

3 Data

To empirically test the reversal hypothesis this paper uses data from the China Health and Nutrition Survey (CHNS).[16] Unlike the naming of the survey suggests it is *not* nationally representative, but only covers nine provinces: Guangxi, Guizhou, Heilongjiang, Henan, Hubei, Hunan, Jiangsu, Liaoning, and Shandong (see Figure 2.A.1).[17] These provinces account for approximately 56% of China's total population. Within each of this provinces a random draw of 24 communities was selected to participate in the survey which provides a total of over 200 communities that were grasped overall. When possible these communities were constantly covered in subsequent CHNS rounds. The survey is of longitudinal nature and comprises questions at the individual, the household and the community level. Thus, households were traced throughout the years of the survey and repeatedly interviewed when possible. So far, seven rounds of the CHNS are available to researchers namely 1989, 1991, 1993, 1997, 2000, 2004 and 2006 - yet the 1989 wave is not used in this study as then data was only collected for pre-schoolers and adults aged 20-45.[18]

Survey questions asked to individuals and their households are, among others, related to general demographics, time use, health indicators, educational outcomes and economic indicators such as labor force participation and various types of earnings and incomes. With

[16] CHNS is an international collaborative project between the Carolina Population Center at the University of North Carolina at Chapel Hill and the National Institute of Nutrition and Food Safety at the Chinese Center for Disease Control and Prevention. For a broader discussion of the CHNS and the sampling design, see Liu (2008) and/or visit the official website at `http://www.cpc.unc.edu/projects/china`

[17] Note that Heilongjiang was first affiliated to the CHNS 1997. At the same time Liaoning did not participate in the 1997 wave, but returned for the consecutive waves. Unfortunately there are no weights (publicly) available to make these data representative of the provinces covered.

[18] The 2009 wave of the CHNS has been made available to researchers only recently. However, at the time when this paper was drafted the community variables of the 2009 wave were not yet available.

respect to the latter, aggregate household income measures are created based on individually earned incomes (i.e. wages from employment, income from agricultural activities as well as from other types of self-employment less of investments in that business), unearned incomes (i.e. income derived from assets) and state-transfers (e.g. subsidy payments) respectively. This paper uses deflated per capita household income instead of individual incomes as it is strongly believed that nutrition decisions are largely made at the household level.[19] With respect to the outcome variables of interest, individuals' anthropometrics are exactly measured by trained health workers largely ruling out measurement error[20] and reporting bias (Cawley (2004)). Moreover, the fact that individuals are repeatedly measured allows to calculate changes in BMI over time. These changes BMI are used to compute average yearly growth rates at the individual level, the main outcome used in the later analysis of this paper.

At the communal level, municipal officials are interviewed in each wave and a wide range of community related information on issues like infrastructure, local labour markets, service availability, etc. is collected. The panel structure allows to capture changes in community characteristics over time and, thus, to make judgments about developmental changes. Furthermore, a distinction between rural or urban communities has been made initially in the first wave of the survey where roughly two third of the communities where classified as rural. Yet, this variable has been held fixed throughout the later waves. As this study aims to exploit not only between variation, but also within variation of communal characteristics, the rural-urban indicator does not suffice for this purpose.

CHNS follow-up rates of individuals are generally high (i.e. being interviewed for the first time in a certain wave, there is on average a 89% chance that the same individual is observed in the following wave). Once covered, individuals are repeatedly interviewed in later waves as long as they remain in their original community - however individuals who leave their original community (e.g. due to migration) are not followed. Moreover, migration in China is known to be significant and may be worrisome for the analyses in this paper if migrant status is systematically correlated with individual BMI and its growth rate, respectively. Since migration is mainly from less developed to more developed areas, the direction of the induced bias would strongly depend on the correlation between migrant status and BMI. If individuals with higher (lower) BMI are more likely to migrate, we should expect to see a lower (higher) level of mean BMI in less developed regions than in the absence of migration. In turn, how migration would then affect the BMI distribution in more developed areas is difficult to know as it depends on the BMI level of migrants relative to the population of the destination province. The matter is complicated, because the CHNS is does not differentiate between reasons for individual attrition from the sample (migration being one of them). Reasons for attrition from the sample other than migration

[19] All incomes measures are inflated by the consumer price index using 2006 as the basis year.
[20] Remaining errors have been cleaned out during data processing.

are (obviously) death and intentional refusal to being interviewed . Hence it is not possible to distinguish between different types of attrition and how they relate to BMI.

Nevertheless, to check whether individual attrition from the panel is systematically correlated with BMI I run regressions within sub-groups of female and male adults aged 18-60 who are first interviewed in a given survey wave. More precisely, I regress the individual BMI measurement of the initial wave (i.e. the wave in which an individual is observed by the CHNS for the first time) on binary attrition status in the next wave - while also controlling for measures of individual age and residence. A summary of the results is provided in Table 2.2 which shows that for almost all subgroups the difference in initial BMI is not significantly different from zero. Only for the sub-group of male adults who were first observed in the 1991 wave, there is a marginally significant (α=0.10) difference in initial BMI between individuals who drop out and individuals who remain in the sample. Although the table is inconclusive as to whether individuals who drop out from the sample experience different growth paths in their BMI throughout later years, the fact that the difference in their initial BMI is mostly non significant is (at least) reassuring .

Table 2.2: Attrition and differences in BMI

	Year first observed				
	1991	**1993**	**1997**	**2000**	**2004**
			Female		
BMI difference	0.10	-0.09	-0.01	-0.13	-0.22
	(0.07)	(0.21)	(0.13)	(0.2)	(0.25)
P-value	0.18	0.66	0.92	0.5	0.38
			Male		
BMI difference	0.12	0.01	-0.11	-0.04	0.22
	(0.07)	(0.21)	(0.13)	(0.2)	(0.27)
P-value	0.09	0.97	0.41	0.85	0.42

Note: The table shows differences (age and residence adjusted) in initial BMI between individuals who remain in the panel and individuals who drop out. Standard deviations are displayed in parantheses.

The final estimation sample investigated in this paper consists of non-pregnant females and males aged 18-60 at the time of the survey - thus children, pregnant woman and elderly outside this age bracket are excluded from the beginning. Starting with 25155 and 24559 female and male observations, respectively, individuals with missing anthropometric data (N=6222 and N=7210) or unrealistic BMI values (i.e. smaller than 15 or larger than 40, N=15 and N=13) are dropped. Also excluded are adults with only one appearance in the CHNS, i.e. for whom no over time changes can be computed (N=4038 and N=3983). I finally exclude individuals for whom there is no or a negative[21] measurement of household income

[21]This is done although negative incomes are reasonable under certain circumstances. Negative incomes might be observed if individuals report more expenses than income which can be legitimate if, for example

(per capita), no information on community characteristics or missing values in key covariates that are used in the estimation. After applying these selection criteria the final sample consists of N=13980 and N=12408 female and male observations. Table 2.3 summarizes the details of the selection process together with a comparison of the mean values of key variables in order to provide further insights of how their distribution changes as the sample is formed. While the averages of BMI, overweight status and income remain relatively stable, mean age and the distribution of survey years is observed to change throughout the selection steps. Compared to the initial sample, individuals in the final sample are, on average, 3-4 years older at the time when they are interviewed. Moreover, observations in the final sample are more likely to be drawn from earlier survey waves. A careful examination of the CHNS data indeed confirms that the probability of missing values in BMI increases in later survey waves and decreases significantly with individual age. Moreover, community information is more more often missing in the later waves of the CHNS. Unfortunately, there is no readily-available explanation as to why we observe these patterns in missing values. Table 2.A.1 (appendix) provides additional summary statistics for the final estimation sample.

4 Development Index

4.1 An Index of Economic-Development at the Community Level

This section describes how CHNS community information is used to construct a one dimensional index of economic development which is used as a sample splitting device in the later analysis. Although economists often approximate stages of development by simply looking at average income (per capita) I argue that for the purpose of this study a purely monetary measure of regional development is overly simplistic. Rating a region's relative standing solely based on the mean income of its inhabitants does not allow to fully capture the structural differences between regions that potentially affect the relationship of interest. Variations in terms of infrastructure, job market characteristics and available services could all potentially influence the way individuals at different positions of the income distribution behave with respect to their caloric intake and expenditure behaviours respectively. In the fashion of Van de Poel et al. (2009) and Jones-Smith and Popkin (2010) this paper will therefore define development in a broader manner embracing various community characteristics which are believed to reasonably depict a regions relative stage of development. Advantageously, the CHNS provides a large battery of variables related to the surveyed communities that can be exploited for this purpose.[22] The final set of variables used to

if someone invests in a new small business that is not yet profitable at the time of the survey. I conducted all analyses without dropping these observation, the results remained very similar.

[22]Admittedly, the choice of the variables is to a certain extent driven by their availability throughout the community survey questionnaires. Potential candidate questions that were, however, not asked for at least one survey year have been excluded as their inclusion would have resulted in the loss of significant

Table 2.3: Details of sample selection process and mean key variables, by gender

		Selection criteria					
	Initial sample	**Missing BMI**	**Invalid BMI**	**Missing BMi Growth**	**Missing Income**	**Missing Community Data**	**Missing Covariates**
				Females			
N dropped	-	6222	15	4038	236	528	136
N remaining	25155	18933	18918	14880	14644	14116	13980
Age	36.37	38.79	38.79	40.30	40.30	40.36	40.42
BMI	22.52	22.52	22.52	22.63	22.62	22.55	22.55
Overweight	0.19	0.20	0.20	0.21	0.21	0.21	0.21
LogIncome	7.93	7.94	7.94	7.93	7.93	7.91	7.91
Year							
1991	19.1%	20.6%	20.6%	22.5%	22.7%	24.7%	25.0%
1993	18.3%	19.2%	19.2%	17.2%	17.3%	18.5%	18.8%
1997	19.8%	19.3%	19.3%	19.9%	19.8%	18.8%	18.2%
2000	21.7%	21.3%	21.3%	20.8%	20.7%	19.6%	19.5%
2004	21.1%	19.6%	19.6%	19.6%	19.5%	18.3%	18.6%
				Males			
N dropped	-	7210	13	3983	210	339	396
N remaining	24559	17349	17336	13353	13143	12804	12408
Age	36.59	38.65	38.65	39.78	39.77	39.84	39.93
BMI	22.22	22.22	22.23	22.26	22.26	22.16	22.16
Overweight	0.15	0.16	0.17	0.17	0.17	0.16	0.16
LogIncome	7.93	7.95	7.95	7.93	7.93	7.91	7.91
Year							
1991	18.8%	20.2%	20.2%	22.6%	22.9%	25.0%	25.4%
1993	18.2%	18.9%	18.9%	17.2%	17.3%	18.7%	18.9%
1997	19.9%	20.0%	20.0%	21.0%	20.9%	19.8%	19.0%
2000	21.9%	21.4%	21.4%	20.7%	20.6%	19.3%	19.2%
2004	21.1%	19.6%	19.5%	18.5%	18.3%	17.2%	17.5%

Note: The table shows details of the selection process as well as mean values of selected key variables. The initial sample contains only non-pregnant adults (aged 18-60). Covariates that are required to be non-missing are age, employment status, physical activity level, education, marital status, education of household head, household size, dependency ratio and share of males in household.

assess a community's relative level of development is shown in Table 2.4. [23]

The averages of the selected variables by survey year are shown in Table 2.5. As can be seen, all variables depict the expected trends over time and their accuracy is supported by their consistent patterns. Simply judging on the basis of these means, CHNS communities have become more developed over the years: infrastructure has improved, services have become more widely available, the labour markets are increasingly moving away from traditional agricultural jobs to ones characterized by less physical activity and regions have become more economically integrated between 1991 and 2006.

Yet, the choice to approximate the stage of economic development by multiple dimensions disallows a unique mapping without further assumptions. In order to make well-defined inter-communal comparisons of economic development feasible a single index measure is required. Following the examples of Van de Poel et al. (2009) and Van de Poel et al. (2012)[24] factor analysis is utilized (retaining only the first factor) using the pooled sample of communities (including the 2006 wave) to form a one-dimensional indicator of development based on the community variables.[25] The major advantage of factor analysis in this setting is that it allows to create a measure which - assuming that the underlying latent process that drives the overtime changes of the selected community variables is economic development- is not merely a product of weights *subjectively* assigned to each of the variables. Even though factor analysis is a well-accepted tool for the purpose of dimension reduction, its major drawback is that resulting factors do not have a meaningful interpretation in absolute terms. This is because assumptions of zero mean and unit variance of the underlying factor is rather arbitrary. Thus, the resulting index does not allow to make statements about the degree of inequality in the development of communities. Nevertheless, the approach allows to deduce an ordinal ranking which is sufficient for the purposes of this paper. In our case the resulting indicator (i.e. the first factor) allows both, cross sectional and over time comparisons of the *relative rank* of communities covered by the CHNS with respect to their stage of economic development. Using only community observations for which none of the selected variables are missing in a certain wave leaves us with 185 communities in 1991,

amounts of data. The inclusion of various other (partly related) measures has been carefully considered, but the selected ones have finally been chosen for mainly three reasons. First, the variables are believed to adequately capture the underlying development process as they cover various relevant areas. Second, while many other variables show a high degree of missing values, these variables are filled for most communities throughout all years of the survey. Third, carefully investigating within community variation of the variables, the selected candidates showed reasonable consistency over the years.

[23]Apparent inconsistencies have been cleaned in the data. For example, missing values of continuous variables have been imputed by taking the average of the previous and the consecutive survey waves where possible. Moreover, binary variables showing clear inconsistencies compared to previous and consecutive episodes have been adjusted.

[24]Notice that Van de Poel et al. (2009) and Van de Poel et al. (2012) do not investigate the phenomenon of economic development, but the degree of urbanicity. While these two concept might be similar to some extent, the final list of variables used in the factor analysis of this paper differs from that of their paper.

[25]The principal-factor method is used to analyze the correlation matrix of the variables presented in Table 2.4. Only the first factor has an eigenvalue significantly greater than 1.

Table 2.4: Community variables reflecting level of development

Category	**Variable**	**Description**
Community Infrastructure	Dirt Roads	Community has mainly dirt roads
	Stone roads	Community has mainly stone or gravel roads
	Paved roads	Community has mainly paved roads
	Bus	Community has bus stop (or long distance bus stop)
	Train station	Community has a train station
Available Services in Community	Telephone	Community has a convenient telephone service
	Post	Postal service available in community
	Newsletter	Community received provincial newspaper the day it is published
	Electric power cut	Average days per week community has electric power cut-off
Indicators of Occupational Environment	Farmland	Community has farmland
	Share non-agricultural	Share of non-agricultural workers in community
	Very light activity	Share of respondents in communtiy claiming very light physical activity
	Light activity	Share of respondents in communtiy claiming light physical activity
	Moderate activity	Share of respondents in communtiy claiming moderate physical activity
	Heavy activity	Share of respondents in communtiy claiming heavy physical activity
Economic Location	Open Trade Area	Community is near open trade area/city ($<$ 2hr by bus)

Table 2.5: Means of Community Variables by Year

Variable	Mean by Wave					
	1991	**1993**	**1997**	**2000**	**2004**	**2006**
Dirt Roads	0.25	0.23	0.19	0.13	0.05	0.03
Stone Roads	0.14	0.17	0.17	0.14	0.15	0.13
Paved Roads	0.47	0.55	0.59	0.68	0.77	0.84
Bus stop	0.54	0.59	0.67	0.77	0.78	0.83
Train station	0.14	0.18	0.24	0.21	0.23	0.28
Telephone	0.63	0.72	0.84	0.93	0.96	1
Post	0.85	0.9	0.95	0.96	0.96	0.97
Newsletter	0.32	0.4	0.47	0.51	0.63	0.76
Electric power cut	1.11	0.67	0.37	0.15	0.12	0.2
Farmland	0.65	0.63	0.58	0.54	0.53	0.5
Share non-agricultural	0.54	0.58	0.59	0.61	0.66	0.68
Very Light Activity	0.11	0.1	0.16	0.17	0.2	0.25
Light Avtivity	0.14	0.13	0.16	0.16	0.2	0.19
Moderate Activity	0.3	0.31	0.29	0.28	0.25	0.22
Heavy Activity	0.35	0.36	0.33	0.33	0.29	0.3
Near Open Trade Area	0.25	0.33	0.35	0.41	0.44	0.4
Average Income (in Yuan)	2561	2919	3673	4774	6403	7392

179 in 1993, 186 in 1997, 212 in 2000, 214 in 2004 and 213 in 2006 - altogether leaving 1189 community-year observations - that are used in the factor analysis.

The results reveal that about 61% of the common variance among the input variables is explained by the retained first factor which is in the following interpreted as the underlying level of economic development. The time constant factor loadings, i.e. the degree to which each of the community variables correlates with our factor of interest, are given in Table 2.A.2.[26] The loadings are bounded between -1 and 1 and their absolute values indicate the corresponding correlation with the underlying factor. All loadings show the expected sign (i.e. they are in line with theoretical arguments) fortifying the assumption that the underlying first factor indeed reflects the relative level of economic development of communities. Clearly, enhancements in infrastructure and transport, service availability and economic integration correlate positively, while labour markets characterized by more heavy physical occupations and a high concentration in the agricultural sector uni-vocally reflect relatively lower degrees of development.

In a further step, the retained factor is used to compute the rank of a community in the overall distribution (ranging from 1 for the lowest level level to 100 for the highest level of retained first factor). The resulting measure is used as a ranking device throughout the rest of the paper. Moreover, for the index to be used as a mean to stratify the data with regard to the later analysis, also a discrete measure is required. Therefore, an additional variable is created that is directly based on the terciles of the originally continuous measure.

[26]Note: Two community variables - availability of post services and train station - are excluded from the final construction of the development variable as their loadings were relatively low (i.e.< 0.3). As expected, their exclusion does not alter any of the subsequent results.

The resulting discrete variable classifies the development stage of a community as being in the categories from one to three with the highest value representing the highest level of development. The decision to choose exactly three classes of development stages is directed by two arguments. On the one hand there should be enough variation (i.e. there should be not too few strata) to allow the capturing of effect heterogeneities in the main analysis. On the other hand, a large number of strata reduces the number of observations in each of the strata and, thus, reduces the precision of our result.[27]

4.2 Assessment of Development Index

For the credibility of the subsequent analysis it is essential to assure trustworthiness of the development index. Following DeVellis (2003), particularly two properties of the index should be ensured. First, it should reflect the over-time development that China has experienced for the last two decades, i.e. the index should on average show higher values for later periods than for earlier ones. Second, as there is certainly heterogeneity between communities and regions, the index should be able to discriminate accordingly.

Plotting the average logarithmic incomes of communities against their value of the development index by waves provides Figure 2.A.2. Linear fit-lines are added. As should be expected the graph consistently shows a positive relationship between the mean income level and the development rank of communities. However it is also revealed that communities with similar values in their development rank have significant variation with respect to their income levels. This supports the assumption that apart from income also non-monetary variables should be considered in the assessment of communities development levels. Moreover, it can be seen that over time the scatter-plot is monotonously moving from the lower-left quadrant to the upper right one which is an indication of increasing levels of aggregate development over time. This point is further underscored by Figure 2.A.3 which draws the naive kernel density estimates of the distribution of the development index for each of the waves. Two things are noteworthy. First, the estimated distribution of the development index is clearly shifting rightward over time, once again supporting the hypothesis of overall developmental progress. Second, the densities of all years are bimodal in nature. This is probably due to persisting rural urban disparities. In fact, when following the definition of rural and urban areas as they are initially coded in the CHNS community surveys, the development index discriminates surprisingly well between them. In the group of communities that are assigned to be rural the average of the development index does not exceed the zero (mean) value in any of the waves while for the urban sub-group the average never falls below it.

Still, significant variation of the development index even within the sub-groups of rural

[27] Alternatively, all analyses have been conducted using quartiles and quintiles with the qualitative results remaining largely the same.

Table 2.6: Average Development Rank by Province and Year

	Year					
Province	**1991**	**1993**	**1997**	**2000**	**2004**	**2006**
Liaoning	44	50		55	60	68
Jiangsu	33	44	53	56	60	63
Shandong	49	51	53	60	66	69
Henan	25	37	38	43	50	55
Hubei	35	38	44	50	54	57
Hunan	32	42	47	56	62	64
Guangxi	39	47	53	55	60	62
Guizhou	32	42	48	51	57	61

Table 2.7: Average Development Rank by Geographical Clusters and Time

	Year					
Area	**1991**	**1993**	**1997**	**2000**	**2004**	**2006**
Non-Coastal Rural	27	33	37	42	48	52
Non-Coastal Urban	44	57	64	68	72	75
Coastal Rural	28	35	42	42	51	56
Coastal Urban	75	79	73	86	85	89

and urban communities indicates that the index is able to capture nuances beyond a crude rural urban definition. To further highlight how the index captures variation in the stage of development between communities, Tables 2.6 and 2.7 show unweighted averages of the constructed development index for each of the covered provinces and for geographical clusters (as defined in section 2) by year.

Again in line with the findings from above, time variation suggests that on the aggregate level all provinces exhibited a positive trend in their development paths. With the exception of Shandong in 1993 the development index universally increases over time.[28] Investigating differences across provinces it can be seen that the eastern coastal regions (especially Liaoning and Shandong) appear to be comparatively more developed throughout the entire time span under consideration. This point is further supported by looking at mean levels of the development index by geographical clusters. As discussed earlier, urban areas exhibit higher levels of development throughout the time span under consideration. Moreover, especially coastal regions are consistently at the forefront. At the same time regions in inner China consistently show lower levels in their values for all waves considered which supports the ability of the index to capture between variations.

Finally, Table 2.8 presents the communities' transition probabilities with respect to the development quintiles between two consecutive waves labeled as t (baseline period) and

[28]Note that in Shandong province in 1993 only 16 communities were covered by the CHNS, i.e. 7-8 communities fewer than throughout all other years and compared to other provinces respectively. Moreover, the missing communities in 1993 were mainly classified as urban in the previous wave. Hence, it is likely that the missing of these communities causes the development index to temporarily drop in 1993.

Table 2.8: Development Category Transition Probabilities in %

		DC t+1		
		1	**2**	**3**
	1	71.92	26.36	1.72
DC t	**2**	4.61	73.76	21.63
	3	0	7.69	92.31

$t + 1$ (future period).[29] For this, the discretized version of the development index variable is used. E.g. the entry in the second column of the first row gives the transition probability that anyone community was classified to belong to the lowest development tercile in one wave, but jumped to the next higher (i.e. the second lowest) tercile in the subsequent wave. Accordingly, the entry in the third column of the third row represents the probability that anyone community was classified to belong to the highest tercile and remained there also in the subsequent wave. Assuming that the majority of communities experienced positive economic development throughout the time under considerations one should expect to find that pattern to be reflected by the data. Indeed, the figures confirm that the majority of communities experienced constant or positive development levels, i.e. probabilities of remaining in a development tercile or jumping into a higher tercile are higher in absolute terms. Moreover, the higher the category in the initial period t, the more likely it is that a community stays there throughout the consecutive waves. At the same time it can be seen that there is a significant number of communities that - given their initial development stage- show a jump into a higher category in the consecutive period. Only few cases ($< 5\%$) experienced a drop in their classification.

In sum, the constructed index appears to reasonably reflect the *relative* stage of development. It shows some desirable properties in that it reenacts the overall positive development path that most areas in China have undergone during the last 20 years. Moreover, it consistently accentuates between community differentials in the level of development. In what follows, the created index is used to characterize the economic environment that individuals live in.

5 Regression Analyses

5.1 Set-Up

This section investigates the determinants of individual BMI growth with a special focus on its income gradient and whether or not it differs when less developed areas are compared with

[29] There are exceptions where communities are left out for one wave. In that case consecutive waves refers to the two closest waves in which the community is observed.

more developed areas in China.[30] The empirical specification that is estimated closely follows the theoretical model of weight management as formulated by Philipson and Posner (2003) and Lakdawalla and Philipson (2009). In their health production framework changes in body weight are first and foremost a result of changes in the ratio of caloric intake to expenditure. Economic development naturally affects this ratio through three channels. First, due to more efficient production technologies and increased supply, food prices (in real terms) should decrease with economic development which, ceteris paribus, should increase food consumption. Second, development leads to lower levels of average strenuousness in daily life. Jobs become more sedentary due to sectoral shifts away from agriculture and heavy industry towards light industry and services. Moreover, improvements in infrastructure lead to more efficient means of transportation. Third, technological advancements increase worker productivity which, through higher wages, allows for more consumption.

The first two channels, lower food prices and lower levels of average strenuousness are exogenous to individuals and should lead to monotonous growth in mean BMI - which is in line with the observation that developed areas in China show a higher mean level of BMI and overweight (compare section 2). In economic terms, developmental progress creates an overall upward pressure on individuals' body weights by making food relatively cheaper and physical activity relatively more expensive. However, complementary to utility from consumption of food and other goods, individuals also have a preference for ideal weight. In particular, individuals aim to achieve (or maintain) a weight level which they consider as ideal[31] whereas larger deviations (in either direction) from their ideal weight lead to more disutility. As a result, overall consumption choices are partly governed by the desire to minimize the deviation from that ideal weight. It follows that the relationship between income and BMI is ambiguous - more income should be associated with consumption choices that lead to weight gain for the underweight (i.e. by further increasing food consumption), but to behaviours associated with weight control (or even weight loss) for the overweight, ceteris paribus. Importantly, this also suggests that any dynamic model of weight management has to account for the base level of body weight. Popular examples of activities related to weight control are dieting, i.e. the substitution of high-caloric diets by low caloric (and often more costly) ones, or increased sports participation. Also, active weight control should be most commonly observed among higher income classes in more developed societies who are most likely to face both, sedentary jobs as well as food abundance.

Perhaps restricted by the use of cross-sectional data existing studies often model contemporaneous body weight as a function of income (i.e. income and body weight are measured at the same time) and either estimate conditional correlations between the two (e.g. Jolliffe

[30]Alternative analyses (not shown here) that stratified sub-samples according based on mean-income levels of communities as well as geographical clusters (non-coastal rural, non-coastal urban, coastal rural and coastal rural areas) instead of the development index yield qualitatively similar results.

[31]This ideal level is idiosyncratic to the individual, but usually gender specific and determined by cultural and social norms (Etil (2007)) as well as by health concerns (Kan and Tsai (2004)).

(2010), Lu and Goldman (2010) or Chou et al. (2004)) or exploit additional exogenous variation in income to estimate local average treatment effects for the sub-population of compliers (Akee et al. (2010) or Cawley et al. (2010)). On the one hand - despite controlling for all confounding factors - simultaneity of body weight and income makes coefficient estimates questionable and hard to interpret.[32] Moreover, using contemporaneous measurements of body weight neglects the dynamic nature of the problem. Body weight is a stock variable and income-related differences in individual behaviour need time to materialize. On the other hand, although overcoming biases due to (1) omitted variables in the estimated equation and (2) reverse causality, instrumental variable approaches make it difficult to reveal general patterns, because the effect is usually not generalizable to the population at large.[33] Using the longitudinal dimension of the CHNS, I follow Ng et al. (2012) and approximate changes in BMI based on a linear growth regression which explicitly recognizes the time dimension of weight management in the theoretical model:

$$BMI_{i,t+1}^{growth} = \beta_0 + \beta_1 logInc_{i,t} + \beta_2 BMI_{i,t} + \beta_3 PA_{i,t} + \beta_4 X_{i,t} + \beta_5 wave_t + \beta_6 geo_i + \varepsilon_{i,t} \quad (1)$$

Note that, from an econometric point of view, a growth specification reduces the bias that stems from the reverse relationship between income and BMI to the extent that it is difficult to think of a person's future growth in body weight to have any influence on his income level today. In the above specification the geometric mean growth rate of individual BMI (BMI_{t+1}^{growth})[34] is modeled as a linear function of its present level (BMI_t) (which we take as given), the log of deflated per capita household income ($logInc_t$), the present level of job related physical activity (PA_t), additional covariates (X_t) (such as age at the time of the survey and its square, education, marital status, household characteristics) as well as time and geography dummies. β_1 is the main coefficient of interest which should be informative with respect to the direction in which a higher level of income relates to weight growth. Present levels (t) in this context refer to measurements taken in the CHNS waves 1991, 1993, 1997, 2000 and 2004 - the 2006 wave is only used for the calculation of growth rates. As usual, $\varepsilon_{i,t}$ is the individual time specific error term. The model is estimated withing subsets of the data that are defined based on terciles of the development index.

Several remarks are in order. First, dynamic changes in body weight reflect a convergence process where adjustments are always made relative to the base level value. Moreover, the present level of body weight differs across income levels which makes controlling for BMI_t

[32] Shimokawa (2008) actually shows that (at least at higher levels of the body-weight distribution) there is a reverse and negative effect of weight on income. Thus, identification strategies that estimate simultaneous relationships are likely to fail even when only controlling for observable characteristics of individuals.

[33] It is well-known that instrumental variables only allow to identify the effect of interest for the sub-population of compliers, i.e. the part of the population that reacts to the instrument.

[34] Formally, the geometric mean growth rate of BMI is defined as $BMI_{i,t+1}^{growth} = (\left(\frac{BMI_{i,t+1}}{BMI_{i,t}}\right)^{\frac{1}{(t+1)-t}} - 1)$.

essential. Given a natural upper bound of body weight and based on the assumption of individual preferences for ideal weight I expect to find negative values for β_2, independent of the level of economic development - i.e. the probability of weight control is higher at higher levels of BMI. However, a key issue when estimating the model by OLS is the endogeneity of BMI_t. It is likely that BMI_t and BMI_{t+1}^{growth} are jointly influenced by unobserved individual specific factors (i.e. genetics and individual specific habits) which are captured in the error term. To the extent that these unobservable factors are positively correlated with both, the present BMI_t and BMI_{t+1}^{growth} , the estimated coefficient β_2 should be upward biased. With respect to $logInc_t$, the estimation bias introduced by the endogeneity of BMI_t might carry through and affect β_1. One popular solution to that problem would be to instrument BMI_t with its lagged value BMI_{t-1} . Unfortunately this not feasible in the present context, because it would shrink the estimation samples to individuals who are observed in three consecutive waves of the CHNS - the resulting sample sizes are simply too small. Another solution would be to rely on time varying parameters (e.g. such as prices) which could be interacted with household or community exogenous variables for identification. Although the CHNS provides the necessary variables, the first stage estimates of potential instruments on BMI_t turned out too weak to be used in the analysis.[35] Likewise, there is no credible instrument for income available in the present context. Income affects BMI through induced changes of caloric intake and/or caloric expenditure. Hence, for exogenous variations in income to affect BMI the effect of the instrument has to be strong and long-lasting. Several IV candidates were tested: income contributions by elderly individuals (>60) who live in the household, the existence of twins younger than sixteen in the household (which would, by definition, lead to a reduction in per capita income) and subsidy payments at the communal level. For all candidates the first stage estimates were simply to weak to be useful for the analyses. Hence, due to the lack of exogenous variation in BMI_t and $logInc_t$ the coefficient estimates in this setting lack a causal interpretation, but reflect conditional associations.

Finally note that the inverted U-shaped relationship is true conditional on work related physical activity PA_t, but it may not be unconditionally true. The theoretical framework suggests that unconditionally at higher incomes job related exercise could be lower and weight may be rising unconditionally. Thus, as long as less strenuous work is associated with higher income, not including the level of job related physical activity could result in an upward bias of β_1. The inclusion of PA_t should account for differences in BMI growth paths that exist due to different levels of job-related strenuousness. Therefore, β_1 should be interpreted as the mean association of income and BMI growth net of differences in growth

[35]Note that in the earlier version of this chapter I used the mean BMI level of adults (18-60) living in the same community as a instrument for individual BMI. While this provided very strong first stage results, the exclusion restriction could not be defended reasonably. Mean community BMI is (strongly) correlated with individual BMI due to shared environmental factors, social norms, food habits and lifestyles. There are good reasons to assume that these factors did not only affect BMI in the past, but will continue to affect future BMI growth paths of individuals. Hence I decided against the use of the instrument.

paths that are due to different levels of work related physical activity.

As the model equation suggests further covariates need to be controlled for. Due to metabolism affects age and its square are important factors to be taken into account. It is known that BMI is an increasing and concave function of age which, in turn, is likely to affect the income level of individuals. After controlling for employment status, and a battery of other individual and household characteristics, the estimated coefficient of $logInc_t$, ($\hat{\beta}_1$), is assumed to indicate the conditional association of household income and future growth in BMI. All analyses are conducted by gender and development tercile (measured in t), and include geography and time fixed effects.[36] Since the model is estimated on a pooled sample the standard errors are adjusted for clustering at the community level.

5.2 Results

Table 2.9 shows the coefficient estimates from the pooled OLS regressions for the determinants of individual BMI growth by gender and level of economic development under the inclusion of geography and year dummies. The coefficients of main interest are found in the first row and support the hypothesis that the association of income with growth in BMI differs across different levels of development. When looking at the sub-samples, in the least developed areas a higher level of (log) income is related to a significantly stronger growth of individual BMI. This is true for males and females likewise, and the coefficients are similar in size. After controlling for the other covariates in the model a 10% higher income level is, on average, related to a 0.015 p.p. (0.013 p.p. for males) increase in BMI_{t+1}^{growth}. However, β_1 turns insignificant for females and males when the second development category is considered (though the point estimates are still positive). It should be noted that here the coefficients are approximately half the size of their correspondents in the least developed categories. This suggests a flattening of the income gradient. The overall picture changes for the economically most developed areas where the income gradients switch sign from positive to negative. Here, for the sub-group of females, a 10% higher income level is associated with a significant 0.017 p.p. decrease in BMI_{t+1}^{growth}. Stated differently, after adjusting for differences in initial BMI levels and controlling for other covariates, in more developed areas the body weights of females located at higher levels of the income distribution (ceteris paribus) grow significantly slower than of those at lower levels - a pattern which is qualitatively similar to what can be in industrialized societies nowadays. Although the income gradient is also negative for males in the most developed areas it is not significantly different from zero. Although the estimates lack a causal interpretation, they suggest that additional weight growth is higher at higher income levels in less developed areas, but lower at higher income levels in more developed areas (at least for females). Not surprisingly, this is in line with the theoretical predictions, but also with the insights from section 2 where BMI was

[36] Geography dummies indicate non-coast versus coast, and rural versus urban location

found to have grown fastest in less developed areas and lower income groups, respectively.

At this point it is worth to investigate the other coefficient estimates. Like expected, there are significant convergence effects in growth paths as reflected by the negative coefficients of BMI_t across all sub-samples - thus, higher levels of BMI_t are associated with significantly lower future growth rates.[37] Furthermore, the model hints at a non-linear relationship between age and BMI growth for females in the least and second least developed areas where the growth rates increase up to the age of 40, but decrease afterwards . In the most developed areas growth reaches its peak somewhat later, i.e. approximately at the age of 50. For males, however, both age coefficients are insignificant. Interestingly, most of the coefficients that capture the effect of job related physical activity appear to be insignificant. This might be due to the fact that some variation in physical activity levels is already captured by the development index. A priory one would assume individuals with low levels of physical exhaustion (i.e. non-employed or with very light and light physical activity) to show higher levels of subsequent BMI growth. This, however, is not reflected by the estimates. A possible reason could be that individuals experience transitions into less demanding activities between t and $t+1$ (which is of course not captured in this setting). In contrast to other studies, the level of education plays virtually no role in the process of body weight accumulation after controlling for the other variables. The same is basically true for the education of the household head. This finding is not in line with the prediction of theoretical models which consider the decision to be overweight as an outcome of limited knowledge about potential health risks. Finally, the relationship between marital status and BMI_{gwt} is also rather weak.

6 Conclusion

Using data from the 1991-2006 waves of the China and Nutrition Survey this paper investigated the recent spread of overweight and obesity among adults in China. The period under consideration was not only characterized by impressive economic growth, but also by a marked divergence in the level of economic development across different regions. The concurrence of these two phenomena provide an interesting setup to study the linkages of economic development with population weight levels - overcoming the limitations of cross-country comparisons. Descriptive statistics confirm that weight levels are continuously rising over all Chinese provinces which are represented in the data. Moreover, while coastal and urban areas show significantly higher levels of overweight and obesity at the beginning of the observation period, that gap is narrowing continuously as non-coastal and less devel-

[37]Interestingly, the pattern of BMI convergence also hold when absolute annual changes of BMI (not shown here) are regressed on the same set of covariates. Hence, the convergence effects which are found are not only due to the way the dependent variable BMI_{t+1}^{growth} is constructed, namely as percentage growth relative to its base level BMI_t .

Table 2.9: Estimation Results by Gender and Development Level, Adults 18-60

	FEMALES			MALES		
Dep. Variable (in %)	Dev.Cat=1 BMI_{gwt}	Dev.Cat=2 BMI_{gwt}	Dev.Cat=3 BMI_{gwt}	Dev.Cat=1 BMI_{gwt}	Dev.Cat=2 BMI_{gwt}	Dev.Cat=3 BMI_{gwt}
$logInc_{t_0}$	0.1458***	0.0751	-0.1720**	0.1329***	0.0689	-0.0839
	(0.051)	(0.062)	(0.073)	(0.049)	(0.061)	(0.074)
BMI_{t_0}	-0.3302***	-0.3473***	-0.3990***	-0.4110***	-0.3018***	-0.3192***
	(0.021)	(0.026)	(0.023)	(0.031)	(0.022)	(0.021)
Age	0.1142***	0.1883***	0.1750***	-0.0093	-0.0579	-0.0205
	(0.039)	(0.043)	(0.053)	(0.035)	(0.040)	(0.054)
Age^2	-0.0014***	-0.0023***	-0.0017***	-0.0000	0.0004	0.0002
	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)
Not Working	reference	reference	reference	reference	reference	reference
Working	1.0848	0.3292	-0.1216	0.6246	0.6872	0.3238
	(0.681)	(0.330)	(0.218)	(0.817)	(0.464)	(0.270)
Very Light PA	reference	reference	reference	reference	reference	reference
Light PA	0.8175	0.6987**	-0.0995	0.1690	0.0246	0.3167
	(0.717)	(0.303)	(0.208)	(0.829)	(0.518)	(0.352)
Moderate PA	-0.1320	0.4563	0.3562	0.2709	0.3472	0.4672
	(0.732)	(0.353)	(0.333)	(0.908)	(0.507)	(0.378)
Heavy PA	0.4813	0.5367*	-0.6793	0.4226	0.3471	0.1861
	(0.663)	(0.294)	(0.760)	(0.808)	(0.504)	(0.722)
Never Married	reference	reference	reference	reference	reference	reference
Married	-0.1165	-0.1467	-0.3370	0.2186	0.4695**	0.1868
	(0.329)	(0.315)	(0.313)	(0.170)	(0.211)	(0.273)
Divorced/Widow	0.1450	-0.4114	-0.6780	-0.4685	0.2805	0.5630
	(0.454)	(0.422)	(0.430)	(0.321)	(0.401)	(0.402)
No Education	reference	reference	reference	reference	reference	reference
Primary	-0.0179	0.0562	0.1980	0.0828	-0.1420	0.2883
	(0.099)	(0.134)	(0.210)	(0.179)	(0.224)	(0.325)
Secondary	-0.0090	-0.0145	-0.0243	0.1363	-0.2530	0.4091
	(0.127)	(0.164)	(0.198)	(0.191)	(0.228)	(0.296)
Higher	-0.0930	-0.1232	-0.3383	0.1328	-0.3259	0.2460
	(0.182)	(0.198)	(0.217)	(0.269)	(0.259)	(0.301)
Constant	3.5272***	3.3334***	6.9035***	7.5023***	7.6474***	8.6822***
	(0.995)	(0.969)	(1.132)	(1.062)	(1.007)	(1.263)
Geography Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,903	4,330	3,747	5,314	3,823	3,271
R-square	0.091	0.109	0.146	0.116	0.093	0.118

Note: Robust standard errors in parentheses. Standard errors account for ID clusters.
*** p<0.01, ** p<0.05, * p<0.1
Table does not report coefficients and s.e. for the following variables: Interaction Terms of Working Status and Physical Activity Level, Education of Household Head, Household Size, Dependency Ratio, Share of Males in Househould

oped areas in China catching up. Likewise, lower income groups are increasingly affected by overweight. In line with findings of earlier studies, it appears that strong declines in physical activity are the main drivers of the epidemic. This appears reasonable as, due to ongoing structural changes in the economy, the are profound shifts across different segments of the labour market, i.e. individuals are moving from jobs in agriculture and heavy industries into light industry and service sector jobs.

Looking more deeply into the income gradient of obesity, the revealed dynamics are mostly in line with the hypotheses of the model by Philipson and Posner (2003) and Lakdawalla and Philipson (2009): as regions become more developed there is some evidence for a "trickling down" of overweight from higher to lower income groups - in particular for females. Results from growth regressions at the individual level show that while income levels are positively correlated with subsequent BMI growth in less developed areas the relationship is negative (and significant for females) in the most developed areas. Extrapolating from this pattern we may expect (as China continues to develop) that lower income groups are going to catch up further - and eventually overtake higher income groups - with respect to overweight prevalence in the future. As such, the ongoing changes are in line with the overall picture that emerges from existing meta studies which find a positive income gradient of overweight in developing countries, but a negative gradient in developed economies (Sobal and Stunkard (1989); Monteiro et al. (2004b)).

One striking insight from this paper (that coincides with existing findings in the literature) is the earlier reversal of the income gradient of body weight for females. Two aspects are likely to play a role in this respect. First, there is a known literature that investigates gender differences in the changes of labour force participation and work conditions as countries develop (see e.g. Mammen and Paxson (2000)). A common finding there is that male labour force participation is consistently high throughout all stages of development, but female labour force participation follows a U-shape trend. In poor societies both, males and females, are mostly engaged in subsistence agriculture or informal activities - both of which are often physically demanding and generate little income. As undernutrition and economic hardship is widespread, it is reasonable to expect body weights to be positively correlated with income levels for both genders. Gender differences are more pronounced in middle income countries, however, where female labour force participation rates are generally lower and women often find themselves in the household or participating in family activities (especially when the husband earns sufficient income). Moreover, those women who are in wage employment typically work in jobs that are on average physically less demanding. On the contrary, the relatively more males remain in agriculture or work in blue-collar jobs, both of which are characterized by a higher level of strenuousness. In that respect, economic development leads to a stronger decline in job-related physical activity for females than for males (as also seen in Table 2.1) or, stated differently, the increase in the cost of caloric expenditure is stronger for females than for males. Although less pronounced, these gender differences persist throughout industrialization. The second aspect relates to gender differences in the views of ideal body shapes. Males are traditionally the bread-winners of the family and larger body sizes are thought to reflect physical dominance and strength. On the contrary, females often face stronger societal pressures to conform to thin body shapes (determined by cultural norms and media images). Indeed, findings by Cawley (2004) show that males tend to overstate while females tend to understate their true weight level. The stronger upward pressure that comes with economic development paired with the gender differences in the perception and relative importance of ideal weight could explain why the turning point where weight control becomes important arrives at an earlier stage of development for females than for males.

Besides the non-causal interpretation of the findings, a further shortcoming of the paper is the inability to look more deeply into the pathways through which income translates into BMI changes at the individual level, namely adjustments of caloric expenditure and intake. While the CHNS contains data on physical activity related to occupations and homework since 1991, precise time-use information that covers leisure time activities (e.g. sports) was completely missing until 1997 and was only added fragmentary thereafter. Since leisure activities are essential for weight control, a closer look at changes in caloric expenditure for the entire period was, thus, not possible. Moreover, the CHNS collects dietary intake data since 1991 - yet a detailed investigation revealed that there is a significant amount of implausible within individual variability of macro-nutrient measurements over time (with the correlations between two adjacent measurements of any macro-nutrient, i.e. fat, proteins and carbohydrates, never exceeding 0.19). While the dietary data may still suffice to look at mean changes, it appears to be too noisy to investigate changes in dietary composition at the individual level.

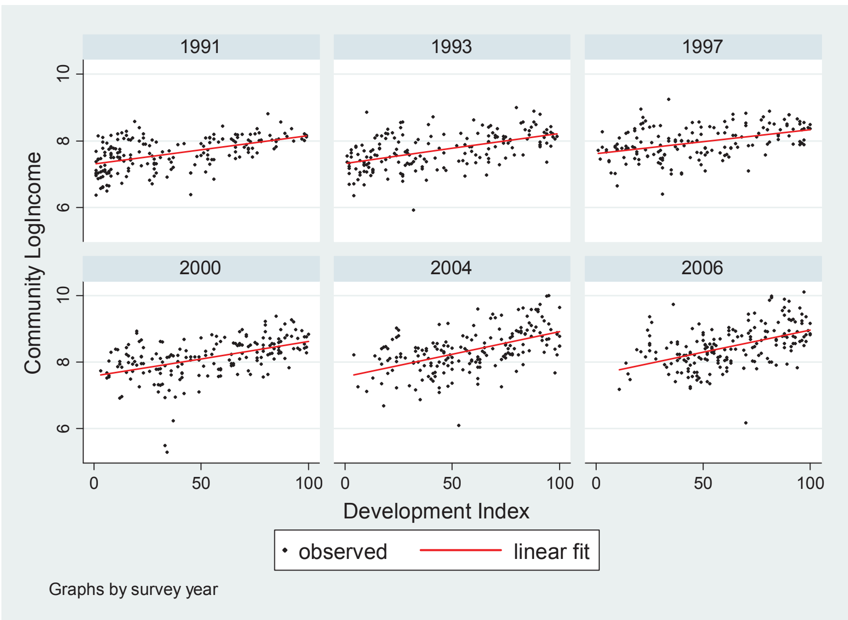
The findings of this study are - despite their descriptive character - relevant for policy makers and future research. First, they provide a comprehensive long-term picture about the evolution of overweight and obesity in China that closely links the phenomenon with the aspect of economic development. The result that overweight is not only on the rise, but also shifting towards less developed areas and poorer parts of the population has direct practical implications. Any policy measure that aims to contain the epidemic is ill-conceived if its restricted to higher income groups only. On the contrary, effective policies need to embrace lower income groups especially - as they are the ones most at risk - and inform individuals about the serious health impairments that can result from unhealthy diets, sedentary lifestyles and excess weight. Failure to do so will come at great expense for the public health system. Being a major determinant of numerous non-communicable diseases such as cardiovascular disease, diabetes, certain types of cancer, psychological problems, among others, a continuing reversal of the overweight gradient will trigger already existing health inequalities at the population level. While in no way conclusive, the findings of this paper should stimulate future research in the field. In particular, future studies should try to identify the causal effects of income on the key parameters of the weight management. This is a formidable task as long lasting and significant exogenous variations in income levels are hard to find in reality. Moreover, more detailed data on time use of individuals as well as higher frequency data on dietary schedules will help to deepen the understanding about the relative importance of caloric expenditure versus caloric intake in weight management.

7 Appendix

Figure 2.A.1: CHNS Regions

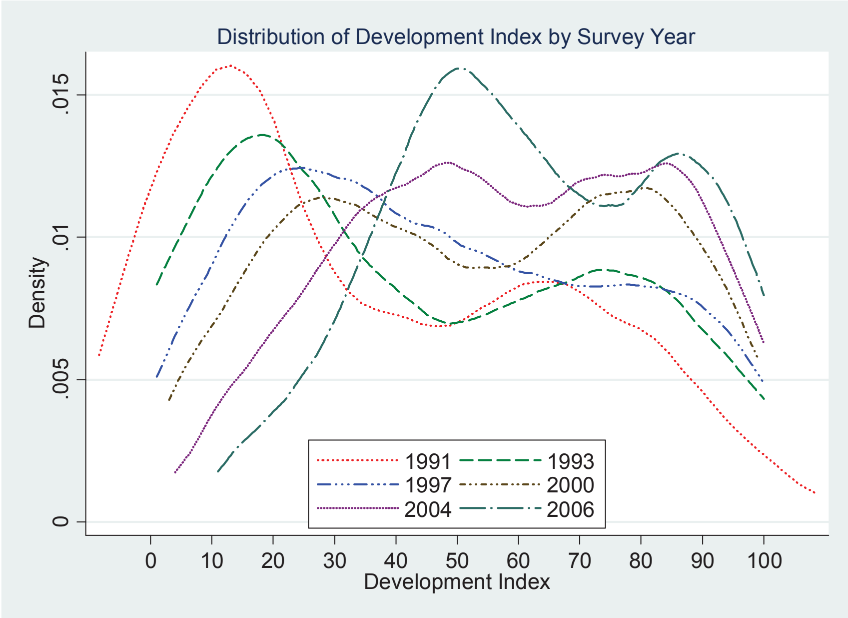
Note: The map shows the geographic location of provinces covered by the CHNS.

Figure 2.A.2: Logarithmic Income vs. Development Index by Wave



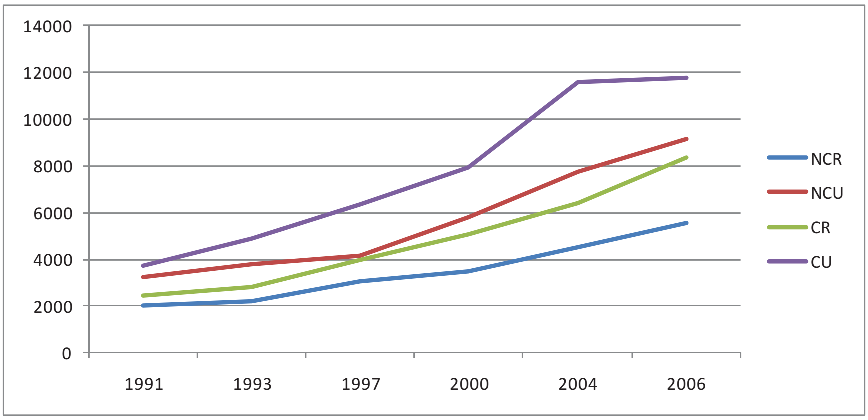
Note: The line represents the unconditional linear fit between the development index and the logarithm of mean income at the community level.

Figure 2.A.3: Distribution of Development Index by Survey Year



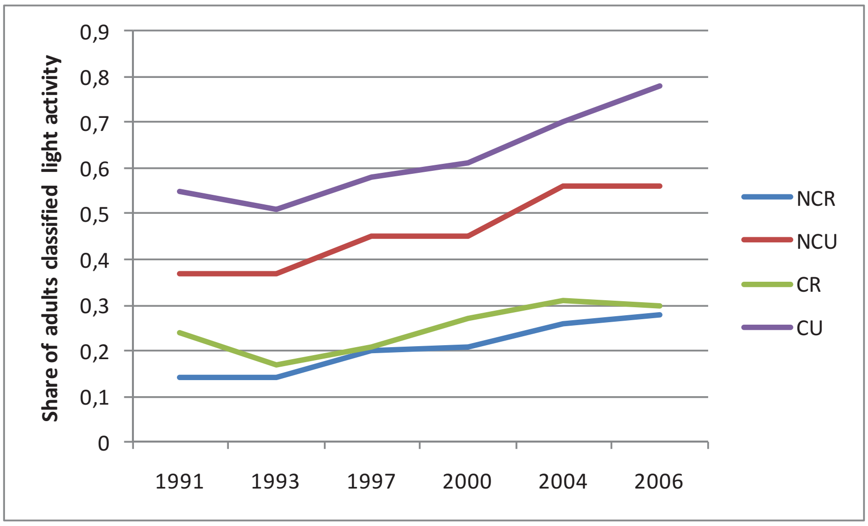
Note: The figure depicts kernel density estimates (by survey wave) of the development index at the community level

Figure 2.A.4: CHNS Trends: Per Capita Income over Time for Geographical Clusters



Note: Own calculations based on per capita household income data (deflated to 2006) as provided by the CHNS

Figure 2.A.5: CHNS Trends: Changes in Activity Levels over Time



Note: Own calculations based on occupation information as provided by the CHNS.

Table 2.A.1: Descriptive Statistics for estimation Sample, Adults 18-60

Variable	**Female (N=13980)**		**Male (N=12408)**		**Total (N=26388)**	
	mean	sd	mean	sd	mean	sd
BMI	22.55	3.18	22.16	2.88	22.37	3.05
Overweight	0.21	0.4	0.16	0.36	0.18	0.39
BMI growth rate (t+1)	0.5	3.4	0.55	3.07	0.52	3.25
LogIncome	7.91	0.93	7.91	0.93	7.91	0.93
Age	40.42	10.56	39.93	11.19	40.19	10.86
Activity Level						
Very light	0.12	0.32	0.12	0.33	0.12	0.32
Light	0.2	0.4	0.14	0.35	0.17	0.38
Moderate	0.14	0.35	0.2	0.4	0.17	0.38
Heavy	0.52	0.5	0.52	0.5	0.52	0.5
Education						
No education	0.34	0.47	0.13	0.34	0.24	0.43
Primary	0.23	0.42	0.24	0.43	0.24	0.43
Secondary	0.27	0.44	0.39	0.49	0.32	0.47
Higher	0.16	0.37	0.24	0.43	0.2	0.4
Marital status						
Never married	0.06	0.23	0.14	0.35	0.1	0.29
Married	0.9	0.29	0.84	0.37	0.87	0.33
Divorced/Seperated/Widowed	0.03	0.18	0.02	0.14	0.03	0.17
Year						
1991	0.25	0.43	0.25	0.44	0.25	0.43
1993	0.19	0.39	0.19	0.39	0.19	0.39
1997	0.18	0.39	0.19	0.39	0.19	0.39
2000	0.19	0.4	0.19	0.39	0.19	0.4
2004	0.19	0.39	0.18	0.38	0.18	0.38
Geography						
Rural non-coast	0.43	0.5	0.45	0.5	0.44	0.5
Urban non-coast	0.21	0.41	0.2	0.4	0.21	0.4
Rural coast	0.25	0.43	0.24	0.43	0.25	0.43
Urban coast	0.11	0.31	0.11	0.31	0.11	0.31
Other						
Share non-agricultural	56.4	35.11	55.62	35.4	56.03	35.25
Household size	4.26	1.43	4.25	1.42	4.25	1.42
Dependency ratio	0.28	0.21	0.26	0.21	0.27	0.21

Note: Summary statistics provided in the table describe the final estimation sample

Table 2.A.2: Factor Loadings of Community Variables

Variable	**Factor Loading**
Dirt Roads	-0.58
Stone Roads	-0.56
Paved Roads	0.88
Bus stop	0.33
Telephone	0.46
Newsletter	0.47
Electric power cut	-0.36
Farmland	-0.59
Share non-agricultural	0.67
Very Light Activity	0.74
Light Avtivity	0.72
Moderate Activity	0.32
Heavy Activity	-0.84
Open Trade Area	0.32

Part III

Economic Development and Socioeconomic Inequality in Female Body Weight

Abstract

The origin of the obesity epidemic in developing countries is still poorly understood. It has been prominently argued that economic development provides a natural interpretation of the growth in obesity. This paper tests the main aggregated predictions of the theoretical framework to analyze obesity: Average body weight is positively associated with economic development. In relatively poor countries, female obesity is a phenomenon of the socioeconomic elite. With economic development, obesity shifts towards individuals with lower SES.

1 Introduction

The obesity epidemic has attracted considerable attention in recent years. Much of the related research has focused on obesity in developed countries (see for example Baum (2009); Baum II and Ruhm (2009); Chou et al. (2008); Gruber and Frakes (2006) among others). On the contrary, the emerging epidemic in the context of developing countries has received less attention in the economics literature (see Abdulai (2010); Doak and Popkin (2008); Asfaw (2007)). A theoretical framework to analyze the obesity epidemic was first provided by Philipson and co-authors (Philipson and Posner (2003);Filmer and Pritchett (2001); Lakdawalla and Philipson (2009); Lakdawalla et al. (2005)). These researchers argue that economic development (and, closely related, technological progress) provides a natural interpretation for the emerging obesity epidemic as it lowers the price of caloric intake relative to the price of caloric expenditure. Using a utility maximization framework, they also suggest that on the individual level, a non-monotonic relationship between income and weight arises, with obesity concentrated among the socio-economic elite in poor countries and among the poor in more developed countries. The non-linearity occurs when overweight individuals use additional resources to loose weight (i.e. substituting unhealthy food with healthy food, or reducing working hours in a sedentary job in favour of additional time spent in sport activities).

So far, the existing empirical evidence supports these main aggregated predictions. Sahn (2009) confirms that mean body weight levels are globally on the rise. Also, based on literature reviews, there is evidence for a rotating gradient of body weight as countries develop - i.e. obesity in developing countries was initially a disease of the affluent (Sobal and Stunkard (1989)), but shifted towards groups of lower socioeconomic status (SES) more recently (Monteiro et al. (2004b); McLaren (2007)). However, while they provide a good starting point, results from these literature reviews are of limited value, because the number of reviewed studies is relatively small, and because included studies often use very different indicators to measure socioeconomic status (such as education, income, wealth, etc.) and/or a great variety of study populations (general population studies, immigrants, selected areas, etc.) - which limits the comparability across countries. A notable exception is the paper by Tafreschi (2011) which investigates changes in the BMI gradient in China. Since that paper, however, focuses on a single case study its results can not be generalized.

Additional insights about the association between socio-economic level and overweight status within countries is very relevant for policy makers in order to predict the dynamics of the disease and to design counter-measures in favour of population sub-groups who are most affected by the problem. Moreover, it is known that socio-economic inequalities in health measures other than overweight exist in many developing countries (Deaton (2003);Wagstaff et al. (2003);Grimm (2011)) where public health systems are weak and economically disadvantaged subgroups in the population are usually over-proportionally affected by acute conditions and infectious diseases. As countries develop, these sub-groups might face a double burden - and, thus, a increase in health inequality - if overweight and obesity concentrations are shifting towards them as a consequence of economic development.

This note is the first one to empirically investigate the SES-obesity gradient in developing countries by (1) using comparable micro-data (Demographic Health Surveys), (2) with identical indicators for SES-rank across surveys, (3) from more than 50 developing countries and (4) matching it with macro level data. It is to be mentioned from the start that the aim of this paper is to unveil associations. We test two main hypotheses which are derived directly from predictions of the underlying theoretical model of Philipson and Posner (2003). Being aware of endogeneity issues in our econometric analyses we do not estimate the causal effect of economic development on either mean body weight at the country level or its gradient in our analyses, but instead investigate whether the patterns (i.e. correlations) found in the data are consistent with the theoretical predictions of Philipson and co-authors - and it turns out that they are. Extrapolating from our findings we suggest that body weight levels will further increase in developing countries and, in addition, add to the existing health gradients in other diseases as the poorer population segments are increasingly affected by the epidemic..

2 Theoretical Background

Philipson and co-authors (Philipson and Posner (2003); Lakdawalla and Philipson (2009); Lakdawalla et al. (2005)) analyzed the long-run rise in obesity in a standard micro-economic framework, where body weight is a commodity produced with chosen inputs (calorie consumption and physical exercise). In their theoretical framework, body weight is influenced by three exogenous factors, i.e. (1) the relative food price, (2) the calories expended per hour of work, and (3) the individual wage rate. Economic development is likely to affect all three factors. First, with economic development, relative food prices decline. Second, technological change alters the industry structure. Fewer people are needed for food production, while other sectors, particularly the service sector, become more relevant. Work becomes more sedentary and, hence, individuals need fewer calories to perform in their jobs. Third, technological progress increases productivity, and thus, wages. All these three factors would lead to a higher demand for calories and higher body weight, particularly among the better-off. On the other hand, individuals have preferences for an ideal weight. If an ideal weight is a normal good, the marginal disutility of deviating from this ideal weight is higher for higher income groups. This results in a negative association between socio-economic status and body weight. Whether or not we observe a positive or negative income gradient of overweight in a population thus depends largely on the level of economic development. In relatively poor countries, the first effect dominates and we expect a positive relationship between SES and body weight. In more advanced countries, the second (disutility) effect dominates for at least some proportion of the population and the relation between SES and weight may be inverted U-shaped. Ideally one would like to empirically test this theory by estimating the causal impact of absolute SES on body weight, either on different parts of the income distribution to analyze effect heterogeneity within a country, or across different countries to compare effects. This is, however, difficult for two reasons: (1) Measuring absolute SES (for example income or consumption) is problematic (Pyatt (2003)) and many household surveys include measures for relative SES-rank but not for absolute SES. (2) A creditable identifying strategy is needed to isolate the impact of SES on body weight, but exogenous variations of SES are rare. We therefore test for two predictions on the aggregate level that follow from the micro-economic predictions. Note that this exercise provides empirical evidence consistent with the theoretical model, but does not directly test for it.

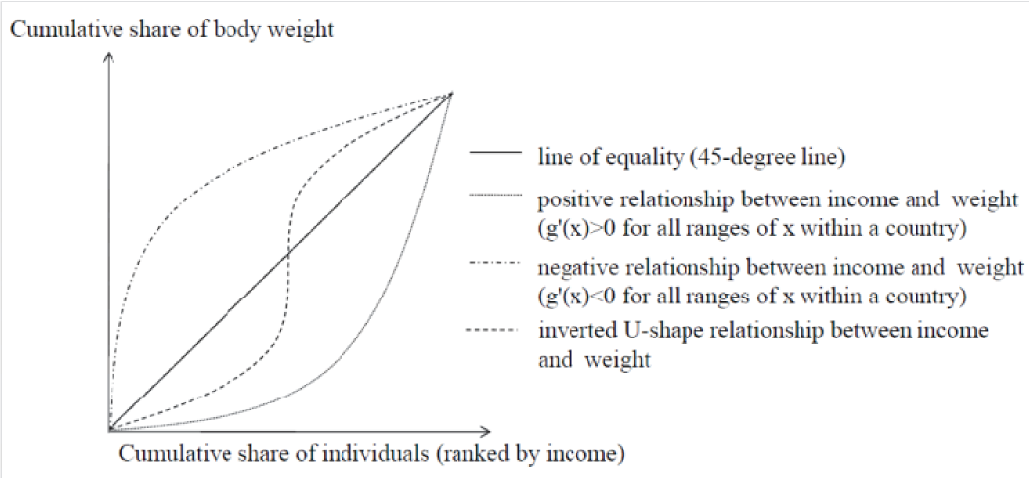
Proposition 1: Average body weight is associated with economic development.

Denote the relationship between body weight and income with $W = g(x)$ and the probability density function of income with $f(x)$. Economic development shifts the income distribution function to $f^*(x)$ such that the average income increases $\int xf(x)dx < \int xf^*(x)dx$. As long as $g(x)$ is a strictly positive transformation (or in other words, if higher incomes lead to higher body weight), economic development should be associated with higher average weight (i.e. $\int g(x)f(x)dx < \int g(x)f^*(x)dx$). The disease is self-limiting, however, if the disutility effect being overweight dominates for at least a some share of the population. If this is the case, economic development can be associated with constant or even declining average weight (i.e. $\int g(x)f(x)dx \geq \int g(x)f\int^*(x)dx$). We, thus, expect a non-linear (and maybe even non-monotonic) relationship between average weight and economic development.

Proposition 2: SES-related health inequality is associated with economic development.

Denote the cumulative proportion of body weight with $q_W(x) = \frac{1}{(E[g(X)])}\int_0^x g(X)f(X)dX$ and the cumulative income distribution with $p(x)$. The concentration curve$L_W(p)$ denotes the relationship between $q_W(x)$ and $p(x)$, and indicates the proportion of weight in individuals with incomes less than or equal to x. The concentration index is twice the area between the concentration curve and the line of equality (Wagstaff et al. (1991)). If the relationship between body weight and income is differentiable, the resulting second derivative of the concentration curve is equal to (Podder and Tran-Nam (1994)): $L''_W(p) = \frac{(g'(x))}{(E[g(X)]f(x))}$ Suppose that on the individual level, higher incomes lead to higher body weight (i.e. $g'(x) > 0$ for all plausible incomes), the concentration curve is convex (the concentration index takes positive values). If in contrast, higher incomes would lead people to lose weight (i.e. $g'(x) < 0$ for all plausible incomes), the concentration curve is concave (the concentration index takes negative values). In the intermediate case, where the relationship between income and weight is inverted U-shaped, the concentration curve crosses the equality line (see Figure 3.1).[38] We would thus expect a negative association between the concentration index and economic development.

Figure 3.1: Concentration curves



3 Empirical Analysis

3.1 Data

To test the two propositions outlined above, we use data from the Demographic and Health Surveys (DHS), which are nationally representative household surveys.[39] These cross-sectional surveys typically include 5,000 to 30,000 households, and provide anthropometric measurements (weight and height) for a selected sample of women. We use DHS data from 52 different countries from 1991 to 2008 (in total 115 different surveys, including information from 943'605 women) to construct aggregate measures at the country-year level. To test proposition 1, average Body Mass Index ($BMI = kg/m^2$), proportion of the population being overweight or obese ($BMI \geq 25$) or obese ($BMI \geq 30$) are used as dependent variables. The concentration index (Kakwani et al. (1997)) based on the DHS wealth index (Rutstein (2008); Filmer and Pritchett (2001)) is used as the dependent variable to test proposition 2. We are only able to compute the concentration indices for 100 surveys, while for the remaining 15 we do not observe the DHS wealth index. Our main independent variable approximating economic development is the level of per capita GDP - this information, as

[38] The concentration index is often used to order countries. Here, concentration curves that cross the line of equality make it very difficult to judge which country has a more unequal body weight distribution. In our application, however, concentration indices are used to test the implication of a theoretical model, where crossing concentration curves are part of this theory.

[39] For further information, see http://www.measuredhs.com/.

Table 3.1: Sample Descriptives

	Obs	Mean	Std. Dev.	Min	Max
Dependent vairables					
Average BMI	115	23.067	2.164	18.92	29.846
Concentration index (BMI)	100	0.0197	0.0107	-0.0064	0.0388
% Overweight	115	0.256	0.1809	0.0195	0.7899
Concentration index (overweight)	100	0.2192	0.1431	-0.0222	0.6611
% Obese	115	0.0811	0.0881	0.001	0.4469
Concentration index (obese)	100	0.35	0.2061	-0.0409	0.8332
Independent Variables					
Time (year)	115	2000	5	1991	2008
Per capita GDP ($)	115	930.96	1020.5	124.85	5115.1
Urbanization (rates)	115	38.51	18.64	11.42	81.11
Population (thousands)	115	37,500	106,000	497	1,090,000
Consumer Price Index	108	73.634	32.38366	0.0501	143.11

well as total population sizes, urbanization rates and consumer price indices are added using the World Development Indicators (2011) database.

Descriptive statistics of the dependent and independent variables are shown in Table 3.1 - and listed in more detail by country and survey year (including original sample sizes) in Table 3.A.1. Our data covers DHS samples of adult females aged 15-49 from most parts of the world ranging from low income countries with a minimum GDP per capita of US$ 125 (Ethiopia (2000)) to higher middle income countries such as Brazil or Turkey with a maximum GDP per capita of US$ 5'155 (Brazil (1996)). While mean BMI at the country level is around 23 (i.e. normal weight), there is significant variation in weight levels across populations. For example, females from the Middle East (e.g. Egypt or Jordan at the upper end) appear to be significantly heavier on average than their counterparts from (South-) Asian countries (e.g. Nepal, India or Bangladesh at the lower end) - which, of course, also translates into significant differences in the shares of overweight and obese females, respectively. Note that a close look at Table 3.A.1 already suggests a positive correlation between mean weight levels and GDP. Moreover, the mean values of the concentration indices for overweight (0.22) and obesity (0.35) indicate that, within countries, the epidemic is mostly concentrated among the relatively wealthier parts of the populations that we investigate. Looking at within country changes over time, however, it can be seen that overweight and obesity appears to be shifting towards the lower ranks of the wealth distribution in some countries (e.g. in Egypt, Ghana or Peru), while it mostly remains (or even expands) at higher ranks in other countries (e.g. in Tanzania or Uganda).

Our data two major advantages over other existing studies. First, individual income as a measure of SES is usually very difficult to measure and (even more difficult) to compare across countries - especially in developing countries where a major share of income is generated in the informal sector. We are less affected by this difficulty as we approximate the overweight gradient using a relative rank measure that is based on household assets. This measure is known to approximate the within country SES rank of individuals reasonably well (seeFilmer and Pritchett (2001)) and is also sufficiently comparable across countries. Second, unlike other sources DHS data do not suffer from measurement error due to self reporting bias in height and weight. However, a major disadvantage of our data is that we do not observe measures of caloric intake or caloric expenditure - which limits our ability to provide empirical evidence about possible mechanisms that underlie our findings.

3.2 Results

We run a regression of per capita GDP on average BMI to test for proposition 1. Alternatively, we run a regression of per capita GDP on the proportion of the population with overweight and the proportion of the population with obesity. A squared term of per capita GDP is included to capture a potential non-linear relationship. Controlling for a linear time trend Table 3.2 (column 1) shows that average BMI, as well as overweight and obesity significantly increase with economic development. The results are robust to controlling for confounding variables on the aggregated level (see columns 2-4).

To test proposition 2, we a run a regression of per capita GDP, its squared term and a linear time trend on the concentration indices for BMI. Alternatively, this model is also estimated for the concentration indices for overweight, and obesity. Again, we gradually extend the models with a set of control variables (Table 3.3) to account for major sources of confounding. As predicted by the theory, our results indicate a significantly negative association between the concentration indices for overweight and obesity and per capita GDP. The association between the concentration indices for BMI and per capita GDP is also negative, but not significant on standard levels ($p = 0.15$). The results are robust to including control variables on the aggregated level (see columns 2-4). This sustains to the second hypothesis that obesity shifts from the socioeconomic elite to people with a relatively low socioeconomic status with ongoing economic development.

Several sensitivity checks are performed to support our main findings. The results are not reported, but available from the authors upon request. First, we perform a demographic standardization (Kakwani et al. (1997)) to account for the heterogeneity in the population structure. Included variable in this standardization are age, pregnancy status, marital status, type of residence, number of children aged below 5 in the household and total number of children ever born. All these factors are very likely associated with body weight for women, and may be also associated with the development status of the country. Second, DHS surveys for more than one year are available for 35 countries allowing estimating time fixed effects panel models. This controls for any time-fixed confounders, such as for cultural differences. Third, we address for the fact that overweight and obesity are binary variables. This is particularly a problem since concentration indices are bounded by the mean of the health variable with binary variables. Since the theory predicts increasing

Table 3.2: Estimation Results for Proposition 1 (per capita GDP measured in thousands)

Dependent variable		Model 1		Model 2		Model 3		Model 4	
Mean BMI	GDP	4.1377	***	3.6763	***	3.6173	***	3.8382	***
		(0.4646)		(0.5019)		(0.5204)		(0.5364)	
	GDP SQ	-0.00074	***	-0.00065	***	-0.00063	***	-0.00068	***
		(0.00012)		(0.00012)		(0.00013)		(0.00013)	
% Overweight	GDP	0.3566	***	0.3136	***	0.3146	***	0.3314	***
		(0.0372)		(0.0397)		(0.041)		(0.0416)	
	GDP SQ	-0.00006	***	-0.00005	***	-0.00005	***	-0.00006	***
		(0.00001)		(0.00001)		(0.00001)		(0.00001)	
% Obesity	GDP	0.1544	***	0.1327	***	0.1334	***	0.1366	***
		(0.0232)		(0.0239)		(0.0254)		(0.0262)	
	GDP SQ	-0.00003	***	-0.00002	***	-0.00002	***	-0.00002	***
		(0.00001)		(0.00001)		(0.00001)		(0.00001)	
Controls		Time trend		Time trend Urbanization		Time trend Urbanization Population		Time trend Urbanization Population CPI	
Obs.		115		115		115		108	

Note: Mean values of BMI, overweight and obesity (i.e. the dependent variables) are computed at the country-year level Robust standard errors clustered at the country level in brackets.
Stars denote significance levels at (*) 10%, (**) 5% and (***) 1%.

Table 3.3: Estimation Results for Proposition 2 (per capita GDP measured in thousands)

Dependent variable		Model 1		Model 2		Model 3		Model 4	
Concentration	GDP	-0.00432		-0.0015		-0.0007		-0.00287	
Index BMI		(0.00273)		(0.00325)		(0.00318)		(0.00335)	
	GDP SQ	0.00000		0.00000		0.00000		0.00000	
		(0.00000)		(0.00000)		(0.00000)		(0.00000)	
Concentration	GDP	-0.2134	***	-0.1701	***	-0.1594	***	-0.1888	***
Index Overweight		(0.0279)		(0.0343)		(0.0328)		(0.0336)	
	GDP SQ	0.00003	***	0.00003	***	0.00002	***	0.00003	***
		(0.00001)		(0.00001)		(0.00001)		(0.00001)	
Concentration	GDP	-0.3161	***	-0.2341	***	-0.2252	***	-0.2601	***
Index Obesity		(0.0385)		(0.0444)		(0.0439)		(0.0457)	
	GDP SQ	0.00005	***	0.00003	***	0.00003	***	0.00004	***
		(0.00001)		(0.00001)		(0.00001)		(0.00001)	
Controls		Time trend		Time trend Urbanization		Time trend Urbanization Population		Time trend Urbanization Population CPI	
Obs.		100		100		100		93	

Note: Concentration indices of BMI, overweight and obesity (i.e. the dependent variables) are computed at the country-year level. Robust standard errors clustered at the country level in brackets.
Stars denote significance levels at (*) 10%, (**) 5% and (***) 1%.

average obesity levels with ongoing economic development (at least for developing countries), lower concentration indices could simply reflect higher average obesity levels and may not portray lower SES-related inequality. We therefore normalize the concentration index by dividing it by its feasible minimum or maximum (Wagstaff (2011a)). And finally, we use alternative measures for the relative SES-rank to adjust for the fact that our results may be sensitive to the choice or our measure for SES (Wagstaff and Watanabe (2003)). The DHS wealth index does not use the same asset types in all countries. We thus construct an alternative asset index that uses the same types of asset. This, however, does not solve the incomparability problem (i.e. the ownership of certain assets may not correspond to a similar SES-rank in different countries). We also use the highest education of the household head as an indicator for SES-rank (even in culturally and economically diverse countries, higher education should lead to higher incomes and thus higher SES-rank). Our results are robust to all these sensitivity checks.

4 Concluding Discussion

This note investigated the recent rise of overweight and its socio-economic gradient in developing countries. Using 115 samples of females aged 15-49 from 52 low and middle income countries who were surveyed during the period 1991 to 2008, we find a positive and robust association between the level of economic development and a country´s mean levels of BMI, overweight and obesity. Moreover, economic development seems to determine the distribution of overweight and obesity within countries. While overweight appears to be mainly an "upper-class phenomenon" in low income countries, lower socio-economic classes are increasingly affected by the disease in middle income countries. Both empirical findings are in line with the patterns predicted by the theory of Philipson and Posner (2003) which argues that technological advancements are a natural explanation for the observed obesity epidemic as they lower the price of caloric intake (be in in monetary units or due to reduced time needed for food preparation) relative to the price of caloric expenditure. As economies undergo sectoral shifts towards more sedentary jobs, physical activity levels are continuously reduced (see Ng et al. (2009)). This trend is also fueled by the rapid urbanization observed around the globe. At the same time, rising incomes will increase the demand for calories in the population - particularly in sub-groups that experienced food-shortages in the past (see Asfaw (2007)). These dynamics can explain the positive association of mean body weight and GDP observed in the data. Additionally assuming that individuals have a latent desire for healthy weight, Philipson and Posner (2003) predict the socio-economic gradient of overweight rotate from positive to negative as an economy develops.

We add useful evidence to the ongoing discussion about the nutrition transition that is happening in many parts of the world (Popkin (1999)). By computing distribution moments based on comparable sets of nationally representative individual level data, we provide a comprehensive snapshot of the epidemic with a high degree of external validity. The patterns that we find are worrisome from a policy perspective. Our findings suggest that, with future economic development, the obesity epidemic is going to further burden health systems and to add to the socio-economic health inequalities which already exist in many developing countries. Thus, policy makers are advised to implement measures that tackle the obesity epidemic. This is not only the case for middle income countries such as e.g. Brazil, Egypt or Turkey where overweight levels are already high at present, but also for low income countries where the disease is only at its onset.

Surely, designing effective responses to tackle the disease is not an easy task for policy makers, especially as it is not uncommon in developing countries that under- and overweight coexist (sometimes even within the same household (Sahn and Younger (2009))). However, with obesity being a major risk factor for non-communicable diseases such as cardiovascular disease, type 2 diabetes, certain types of cancer and others, neglecting the problem will come at the price of significant health related costs in the future and trigger existing health inequalities - the US example supports this prediction (Cutler et al. (2003)). A recent article by Gortmaker et al. (2011) provides a rich discussion about possibilities of cost effective interventions in this context. Discussing the complexity of the phenomenon, the authors argue that in order to reverse the recent global trends that we observe in the data, many sustained interventions at several levels (i.e. governments, international organizations and the private sector) are required. Adequately designed infrastructure projects have the potential to stimulate physical activity at the individual level (e.g. construction of bicycle lanes or parks in urban areas for recreational activities). Subsidizing "healthy food" or, alternatively, taxing unhealthy food can lead to desirable changes of individual consumption patterns (Schroeter et al. (2008)) - especially when paired with information campaigns that explain the negative health consequences of unhealthy diets to consumers (Kan and Tsai (2004)) and change preferences towards more healthy alternatives. At best, policies should be tailored to specific sub-groups to gain maximum efficiency. In addition, more resources should be made available for monitoring systems and research on the topic.

Finally, it should be noted that our empirical findings have important limitations that should be tackled by future research on the topic. First, using data from the DHS, we can not make any statements with respect to male populations, children or adult females older than 49. Second, unobserved confounding factors (e.g. cross-country differences in cultures and attitudes towards overweight) and potential reverse causality (e.g. the negative effect of obesity on worker productivity) are likely to bias our coefficient estimates. Thus, despite the inclusion of important observed confounders and country fixed effects in the analyses as an attempt to reduce the bias, our results have no causal interpretation. The robustness of the associations we find, however, makes our results (at least) a relevant benchmark for future studies on the topic. Third, due to data limitations we do not investigate the two main channels, i.e. individual´s caloric intake and expenditure, and how they change as countries become richer. For sure, knowledge about behavioral changes with respect to these variables is crucial to gain a deeper understanding of why we observe the gradient to rotate - instead of relying on theoretical arguments only. Unfortunately, DHS data is rather limited in that respect and would not allow us to make any statements in that direction. Finally, one could argue that using GDP as a proxy for economic development is overly simplistic. While this critique is certainly legitimate, we experimented with alternative measures of economic development (e.g. share of labour force in agriculture or child mortality), but the qualitative patterns remained.

5 Appendix

Table 3.A.1: Estimation Sample - Detailed Descriptive Statistics

Survey	# Observations	Mean BMI	CI Mean	Overweight (%)	CI Overweight	Obese (%)	CI obese	Per Capita GDP ($)
Armenia(2000)	6166	24.92	0	0.41	0.01	0.14	0	621.48
Armenia(2005)	6301	24.98	0	0.42	-0.02	0.15	0	1598.88
Azerbaijan(2006)	8147	25.4	0.01	0.47	0.08	0.18	0.12	2472.99
Bangladesh(1996)	4409	18.92	.	0.03	.	0.01	.	311.35
Bangladesh(1999)	5073	19.36	0.03	0.04	0.66	0.01	0.34	330.55
Bangladesh(2004)	11306	20.2	0.03	0.09	0.51	0.01	0.63	375.25
Bangladesh(2007)	10836	20.69	0.04	0.12	0.47	0.02	0.66	433.69
Benin(1996)	2603	21.3	0.02	0.1	0.3	0.02	0.59	373.85
Benin(2001)	6168	22.47	0.03	0.19	0.34	0.06	0.52	344.78
Benin(2006)	16717	22.64	0.03	0.19	0.32	0.06	0.49	568.81
Bolivia(1994)	2635	24.48	0.01	0.36	0.12	0.08	0.29	817.62
Bolivia(1998)	4604	25.43	0.01	0.48	0.05	0.12	0.2	1064.91
Bolivia(2003)	17277	25.52	0.01	0.47	0.03	0.15	0.11	914.79
Brazil(1996)	3448	24.1	0.01	0.36	0.07	0.1	0.12	5115.14
Burkina Faso(1993)	3931	21.14	0.01	0.07	0.31	0.01	0.64	240.71
Burkina Faso(1998)	3836	21.03	0.02	0.06	0.48	0.01	0.83	250.61
Burkina Faso(2003)	12253	20.98	0.03	0.09	0.54	0.02	0.74	326.43
Cambodia(2000)	7467	20.65	0.01	0.06	0.27	0.01	0.48	293.1
Cambodia(2005)	8350	20.96	0.02	0.1	0.28	0.01	0.42	462.51
Cameroon(1998)	1853	22.79	0.03	0.22	0.31	0.04	0.48	636.47
Cameroon(2004)	5181	23.66	0.03	0.3	0.23	0.08	0.38	906.14
Chad(1996)	4443	20.69	0.01	0.06	0.16	0.01	0.23	217.58
Chad(2004)	3582	20.88	0.02	0.08	0.34	0.02	0.49	450.04
Colombia(1995)	3882	24.63	0.01	0.42	0.05	0.09	0.06	2534.81
Colombia(2000)	3784	24.86	0.01	0.43	0.04	0.11	0.11	2364.27
Colombia(2005)	38457	24.35	0.01	0.39	0.02	0.12	0.04	3371.04
Comoros(1996)	888	22.64	0.01	0.22	0.1	0.05	0.25	464.17
Congo D.R.(2007)	4726	21.4	0.02	0.11	0.33	0.02	0.41	159.45
Cote d'Ivoire(1994)	3512	22.15	0.02	0.14	0.34	0.03	0.49	573.28
Cote d'Ivoire(1998)	3011	22.59	0.03	0.19	0.25	0.05	0.35	779.41
Dominican Republic(1991)	2454	23.28	.	0.27	.	0.07	.	1307.71
Dominican Republic(1996)	7945	24.36	0.01	0.38	0.06	0.12	0.09	2222.26
Egypt(1992)	5475	26.82	.	0.58	.	0.23	.	694.06
Egypt(1995)	7748	26.32	0.04	0.53	0.16	0.2	0.32	942.08
Egypt(2000)	15353	29.17	0.03	0.77	0.08	0.39	0.14	1422.73
Egypt(2005)	19280	29.85	0.03	0.79	0.06	0.45	0.12	1162.42
Egypt(2008)	16410	28.96	0.02	0.77	0.05	0.38	0.08	1997.1
Ethiopia(2000)	15211	19.95	.	0.03	.	0	.	124.85
Ethiopia(2005)	6643	20.3	.	0.04	.	0.01	.	164.81
Gabon(2000)	2795	23.6	0.02	0.29	0.19	0.08	0.29	4108.8
Ghana(1993)	1931	21.82	0.02	0.13	0.36	0.03	0.54	365.63
Ghana(1998)	2321	22.16	0.03	0.16	0.37	0.05	0.53	401.61
Ghana(2003)	5346	23.12	0.04	0.25	0.32	0.08	0.45	363.84
Ghana(2008)	4813	23.65	0.04	0.3	0.27	0.09	0.42	690.48
Guatemala(1995)	5801	24.22	0.02	0.35	0.19	0.08	0.38	1465.14
Guinea(1999)	3803	21.79	0.02	0.13	0.32	0.02	0.53	420.95
Guinea(2005)	3962	21.85	0.02	0.14	0.31	0.03	0.43	353.61

Survey	# Observations	Mean BMI	CI Mean	Overweight (%)	CI Overweight	Obese (%)	CI Obese	Per Capita GDP ($)
Haiti(1994)	2237	21.36	0.03	0.12	0.4	0.03	0.56	309.35
Haiti(2000)	9965	22.98	0.03	0.26	0.24	0.08	0.42	448.93
Haiti(2005)	5254	22.41	0.03	0.21	0.28	0.06	0.54	463.81
Honduras(2005)	19246	25.57	0.02	0.47	0.08	0.19	0.14	1415.1
India(2005)	118734	20.49	0.04	0.12	0.44	0.03	0.62	740.15
Jordan(1997)	3676	27.38	0.01	0.63	0.03	0.28	0.01	1625.44
Jordan(2002)	5671	28.45	0.01	0.7	0.02	0.36	0.03	1902.39
Jordan(2007)	5196	27.87	0	0.66	0.02	0.3	0.02	2890.87
Kazakhstan(1995)	3684	24.78	0.01	0.39	0.05	0.16	0.07	1288.24
Kazakhstan(1999)	2283	24.07	0	0.32	0.03	0.12	-0.01	1130.11
Kenya(1993)	3784	22.06	0.02	0.14	0.28	0.03	0.4	222.83
Kenya(1998)	3689	22.04	0.03	0.16	0.32	0.03	0.47	474.87
Kenya(2003)	7705	22.68	0.04	0.23	0.29	0.06	0.4	441.2
Kyrgyz Republic(1997)	3783	23.49	0.01	0.28	0.08	0.08	0.11	374.15
Lesotho(2004)	3414	25.16	0.03	0.43	0.14	0.16	0.25	656.1
Liberia(2007)	6920	22.55	0.02	0.2	0.23	0.05	0.35	202.61
Madagascar(1997)	3006	20.52	0.01	0.04	0.24	0	0.43	254.11
Madagascar(2003)	7806	20.87	0.02	0.07	0.36	0.01	0.59	328.63
Malawi(1992)	2753	21.78	0.01	0.11	0.19	0.01	0.47	183.73
Malawi(2000)	13014	22.04	0.01	0.13	0.21	0.02	0.42	150
Malawi(2004)	11127	22.12	0.01	0.14	0.19	0.02	0.41	203.6
Mali(1995)	5031	21.22	0.01	0.09	0.36	0.01	0.76	282.3
Mali(2001)	12077	22.02	0.02	0.15	0.32	0.03	0.52	255.36
Mali(2006)	14271	22.19	.	0.18	.	0.05	.	490.13
Moldova(2005)	7245	25.12	-0.01	0.42	-0.02	0.18	-0.04	794.87
Morocco(1992)	3314	24.27	0.03	0.34	0.21	0.11	0.37	1133.71
Morocco(2003)	16642	24.26	0.02	0.37	0.12	0.11	0.22	1687.73
Mozambique(1997)	3642	21.68	0.02	0.11	0.39	0.02	0.5	222.29
Mozambique(2003)	11705	22.17	0.03	0.14	0.37	0.04	0.65	237.95
Namibia(1992)	2535	22.62	0.04	0.21	0.32	0.07	0.44	1867.54
Namibia(2006)	9533	23.22	0.04	0.28	0.24	0.12	0.29	3898.69
Nepal(1996)	3778	19.96	.	0.02	.	0	.	203.53
Nepal(2001)	8694	20.42	.	0.06	.	0.01	.	224.2
Nepal(2006)	10738	20.66	.	0.09	.	0.01	.	328.29
Nicaragua(1998)	13059	24.94	0.01	0.43	0.09	0.13	0.19	722.63
Nicaragua(2001)	12575	25.66	0.02	0.48	0.09	0.18	0.16	791.39
Niger(1992)	3948	20.85	.	0.07	.	0.01	.	280.45
Niger(1998)	4019	20.76	.	0.07	.	0.01	.	200.65
Niger(2006)	4542	21.44	0.02	0.12	0.44	0.03	0.61	265.41
Nigeria(1999)	2481	22.92	.	0.24	.	0.08	.	286.12
Nigeria(2003)	7444	22.42	0.03	0.21	0.26	0.06	0.38	502.42
Nigeria(2008)	32358	22.71	0.03	0.23	0.27	0.06	0.4	1401.54
Peru(1991)	5789	24.86	0.01	0.41	0.1	0.09	0.32	1554.8
Peru(1996)	11735	25.12	0.02	0.46	0.12	0.1	0.26	2293.09
Peru(2000)	26699	25.43	0.01	0.47	0.07	0.13	0.16	2049.3
Rwanda(2000)	10071	22.15	0.01	0.14	0.19	0.01	0.43	218.02
Rwanda(2005)	5642	21.98	0.01	0.12	0.2	0.01	0.43	264.58
Senegal(1992)	3444	21.9	.	0.16	.	0.04	.	753.01

Survey	# Observations	Mean BMI	CI Mean	Overweight (%)	CI Overweight	Obese (%)	CI Obese	Per Capita GDP ($)
Senegal(2005)	4580	22.33	0.02	0.22	0.22	0.07	0.32	770.09
Swaziland(2006)	4856	26.4	0.02	0.51	0.09	0.23	0.17	2348.59
Tanzania(1991)	5221	21.84	.	0.12	.	0.02	.	188.36
Tanzania(1996)	4282	22.08	0.02	0.14	0.3	0.02	0.48	210.76
Tanzania(2004)	10232	22.41	0.03	0.18	0.31	0.04	0.57	299.15
Togo(1998)	3721	21.74	0.02	0.12	0.35	0.02	0.52	323.38
Turkey(1993)	2655	25.84	0.01	0.51	0.05	0.19	0.05	3036.61
Turkey(1998)	2564	26.05	0	0.54	0.01	0.19	0	4177.43
Turkey(2003)	3288	26.57	-0.01	0.58	-0.01	0.23	-0.02	4393.43
Uganda(1995)	3868	21.64	0.01	0.09	0.27	0.01	0.42	274.69
Uganda(2000)	6626	22.03	0.03	0.14	0.35	0.03	0.56	253.48
Uganda(2006)	2860	22.23	0.03	0.17	0.34	0.04	0.63	335.78
Uzbekistan(1996)	4375	22.74	0.01	0.22	0.09	0.05	0.17	600.6
Zambia(1992)	3896	21.96	.	0.15	.	0.02	.	380.06
Zambia(1996)	4557	22.11	0.02	0.14	0.25	0.02	0.43	348.93
Zambia(2001)	7524	21.65	0.03	0.13	0.37	0.03	0.57	339.15
Zambia(2007)	7046	22.63	0.03	0.2	0.3	0.05	0.52	926.6
Zimbabwe(1994)	2171	23.18	0.02	0.24	0.25	0.06	0.4	599.24
Zimbabwe(1999)	5613	23.57	0.02	0.28	0.17	0.07	0.31	482.13
Zimbabwe(2005)	8717	23.13	0.03	0.25	0.21	0.07	0.32	273.99
Overall Mean (unweighted)		23.07	0.02	0.26	0.22	0.08	0.35	930.96

Note: The table shows descriptive statistics of all DHS surveys used in the analyses.

Part IV

Retention Effects in Higher Education

Abstract

Retention policies are commonly used to maintain student quality at educational institutions. Their effectiveness, however, is debated in the literature. Exisiting papers investigate the effect of retention on student outcomes in primary and secondary education - results for higher education are non-existent. This paper complements the literature as it analyses the effects of retention during the first year at the university level. To establish causality a binding minimum requirement of the first year is utilized in a regression discontinuity framework. Administrative data from the University of St. Gallen, Switzerland, is used to estimate causal effects of retention on subsequent drop-out probabilities of students, the choice of major studies, their study speed and grade performance. While the effects of retention on immediate drop-out and subsequent study speed are rather modest, significant improvements in grades are found.

1 Introduction

Equal access to tertiary education is a major goal of education policies in many OECD countries.[40] As a result, student numbers have increased sharply over the last decade. For example, in the US student full-time enrollment in tertiary education has increased by 32% between 1996 and 2006 (OECD (2008)). While this trend seems positive in terms of inequality concerns, increasing student numbers can harm the ability of universities and colleges to meet accountability standards. As a direct consequence they might have to apply stricter selection criteria at different stages throughout college or university education which, in turn, might lead to higher student dropout rates (OECD (2008)).

At the same time, integrating students from various backgrounds might challenge universities from a different angle. Students from families with low academic background might struggle more while trying to adapt to the university environment. Due to their slower adjustment pace, these students might also drop out, voluntarily or involuntarily. As emphasized by Tinto (1975), an inverse relation exists between the quality of the student-environment fit and his dropout probability.

Some OECD countries, e.g. the EU/EFTA countries, have been facing the above-mentioned problems caused by free access to university for many years. Tuition fees in these countries have been traditionally low or even in-existent. Due to legal restrictions public universities are very often not allowed to select their students in advance. Therefore, selection has to take place during college/university. Furthermore, students tend to drop out in late stages of their studies due to their preferences and ability.

Consequently. universities have developed idiosyncratic strategies of coping with the tension between equal access, accountability standards, and differential paces of students in adjusting to a new environment. So far, no systematic evaluation of the different potential systems exists that would provide guidelines to policy makers. This study is the first to evaluate one particular selection approach, taking the University of St. Gallen in Switzerland as an example. Due to nationwide rules, the University of St. Gallen is not allowed to select students in advance. Therefore, students are selected throughout their first year, the so-called Assessment Year (ASY). Students who fail this year according to a strictly enforced rule are not promoted to the Bachelor level. At the same time, the university follows a remedial approach, i.e. failing first year students are given the chance to repeat the full first year.

This paper examines lessons that can be learned from this selection and remedial approach for the design of selection systems at the university level. In particular, it focuses on three research questions. First, how does retention in such a system affect drop-out

[40] In his State of the Union Address on January 24, 2012, Barack Obama emphasized the importance of this goal: "When kids do graduate, the most daunting challenge can be the cost of college. (...) Higher education can't be a luxury - it's an economic imperative that every family in America should be able to afford."

behaviour, i.e. are students willing to take the chance of repeating, or are they scared away by the outlook of repeating a full year? Second, given that they repeat, how is the subsequent choice of major studies affected by it? Third, does repeating the first year help students to successfully proceed throughout university, and what do they gain in terms of educational outcomes? Both questions center around the benefits of the retention approach for the individual student.

2 Literature

This paper largely builds on the literature on retention in primary and secondary education. There exists no formal theory that would clearly predict the effects of retention per se. Instead, it is mostly intuitively argued that retention can have both positive and negative effects on student outcomes. As summarized in Manacorda (2012), positive effects might include learning gains from repetition as well as improved confidence as students can cope better with their performance requirements. Negative effects might occur through retarded learning, low aspirations, stigmatization and necessary adjustments to a new classroom environment. Therefore, determining the net effect of retention on student outcomes ultimately remains an empirical question.

The empirical literature on retention effects in education has vastly developed during the last decade. The majority of early studies from the pedagogical and psychological literature have shown negative (conditional) correlations between retention and subsequent academic outcomes, as summarized in Jimerson (2001). Newer attempts to measure causal effects of retention include instrumental variable methods (Eide and Showalter (2001)), selection models (Lorence and Dworkin (2006)), difference-in-differences approaches, or most prominently, regression discontinuity designs (Jacob and Lefgren (2004), Roderick and Nagaoka (2005), Jacob and Lefgren (2009), Greene and Winters (2007), Manacorda (2012)). The results of these studies are less negative than earlier studies would suggest.

Results regarding test scores are mixed, but there is a tendency towards positive effects of retention in lower grades. An unequivocally positive effect on test scores seems to exist for retained 3rd graders in the US. Three studies independently find a positive effect for Chicago (Jacob and Lefgren (2004)), Texas (Lorence and Dworkin (2006)), and Florida (Greene and Winters (2007)). These results, however, do not translate to 6th graders as shown by Jacob and Lefgren (2004). Roderick and Nagaoka (2005) find even negative effects on test scores of 6th graders in Chicago. All outcomes examined in these studies are short-term outcomes, i.e. measured 1-3 years after retention

With respect to dropout, existing studies confirm the intuition that retention leads to higher dropout rates as the costs of finishing a degree increase for retained students. This result has been confirmed by Jacob and Lefgren (2009) for 6th graders in Chicago, by Ou

(2010) for 9th graders in New Jersey, and by Manacorda (2012) for 7th to 9th graders in Uruguay. As emphasized by Ou (2010), dropout is higher for minority students, which is in line with Tinto's theory on the inverse relation between the quality of the student-environment match and his dropout probability (Tinto (1975)).

So far only one study has looked at post-education outcomes, i.e.10 years after being retained. Eide and Showalter (2001) find no effect of retention during high school on earnings. This results is not surprising as all other studies did not find any immediate effects of high school retention either. Although retention effects in primary and secondary education have been studied extensively, there is no study in the economics literature that investigates retention effects in higher education. Identification in a higher education setting might be difficult due to a lack of clear-cut retention rules in many institutions. Yet, it might be worthwhile to investigate retention effects for students in higher education for various reasons.[41] First, negative effects of retention might be less pronounced in post-secondary education. More mature students are expected to display stronger self-confidence that can less easily be harmed. Furthermore, detachment from initial cohort members is probably less harmful in a university environment where interaction does not only take place within a single classroom. Second, the effect on dropout might be higher for university students. On the one hand, university education is voluntary, so that dropouts from university do not have to face any sanctions. On the other hand, outside options of university dropouts are certainly more valuable than outside options of high school dropouts. Third, university students might benefit especially from repeating the first year. Quickly adapting to the university environment might be especially difficult for students from families with low academic background. Furthermore, some students need additional time for developing new study habits, e.g. self-guided learning.

Therefore, this paper aims to contribute to the literature by examining retention effects in higher education on subsequent academic outcomes, especially on dropout, grades, and major choice. Methodologically, our design is closest to the routes taken by Jacob and Lefgren (2004) who exploit a strictly enforced and clear-cut test score cutoff in order to identify retention effects.

3 Institutional Setup

The University of St. Gallen is a Swiss public institution. As a traditional business school it offers degree courses in Business Administration, Economics, International Affairs, and Legal Studies. It is the largest college of its kind at the national level when measured by the number of students in Business and Economics. The significance of the institution as

[41] There is a literature that investigates the effectiveness of remedial education programs (e.g. Martorell and McFarlin Jr (2011)) - which is, however, substantially different as remedial courses are meant to support student achievements in major courses.

such is reflected in Table 4.A.1 which shows the number of graduates in Switzerland and in particular at the University of St. Gallen over the last few years - the institution accounts for roughly 30% of all graduates in Economics and Business Administration in Switzerland.

For legal matters the university has no direct control about the number of new entrants. By federal law it is committed to accept every student with a Swiss university entrance license, i.e. a Swiss high school certificate (so-called matura). For students with foreign high school degrees there exists a pre-defined admittance rate which varies by year. Foreign students' admittance is based on an entrance test. The unrestricted admittance of Swiss high school graduates is reflected by the continuous rise in student numbers the university experienced over the last decade. While the number of first-year students amounted to about 800 students in 2006, 10 years earlier only about 600 students entered (Table 4.A.2).

In order to maintain a high quality of education and degrees, the so-called assessment year (ASY) was introduced in 2001. The primary goal of the ASY is to select first-year students into the Bachelor level. Students are allowed to proceed to the Bachelors level when they meet the requirements as stated by the ASY regulations. Over the years 2001 - 2006, the university admitted only approximately two thirds of all first-year students to the Bachelor directly - however non-admittance can be due to both voluntary and non-voluntary dropout.

The ASY requires identical core subjects and test criteria for all students. By the end of the ASY, students must have chosen their Bachelor specialization (Business, Economics, International Affairs, Law and Economics (not considered in this paper)). There are two sub-groups of students for which the ASY differs. First, students who intend to specialize in legal studies follow a different curriculum during the ASY. Second, students of non-German mother tongue can chose to complete the assessment courses within two years instead of one (so-called extended track). Due to these special terms both groups are excluded from all analyses in this paper. All other students follow a strictly defined standard curriculum (Table 4.A.3) - henceforth denoted as the Business/Economics track - and form the population of interest for this study.

The core curriculum of the Business/Economics track comprises courses in *Business Administration, Economics, Legal Studies and Mathematics*. These subjects are tested after the first and second semester during predetermined examination weeks (so-called central examination periods). Moreover, students have to proof sufficient foreign language skills which are also examined during the second central examination period. In addition to the core subjects, students have to take courses in *leadership skills* and *critical thinking*. Both are either evaluated on the basis of group presentations or written essays. Finally, students have to submit a major essay in one of the core subjects to be handed in before registering to the second central examination period.

All courses are compulsory and each course in the ASY is graded and weighted by a number of pre-defined credit points. The overall grading of all courses, in turn, leads to the final decision on whether the student passes or fails the ASY. The whole curriculum of the ASY comprises a total of 60 credits (55 in 2001). The grading scale in the Swiss education system is defined from 1 to 6 in steps of 0.5 - with 4 being the worst passing grade and increasing values indicating better performance. The grading process makes sure that students cannot pass or fail only due to one evaluation. Instead, a student passes or fails according to a *performance measure based on all course grades*.

The ASY is designed so as to make selection into the retention-treatment as objective and non-manipulable as possible. This is ensured by the following four steps. First, all courses are compulsory, and examination dates are fixed by the university. Second, examination dates are blocked in short time periods. In particular, all core subjects are tested within a central examination period of five weeks in both the fall and the spring term. This time pattern leaves only limited space for students to strategically adjust their learning behavior during these periods. Third, grade disclosure takes place exclusively at the end of each semester. All course grades are jointly disclosed at the same day by mail. Notice that exams taken in the final central examination period take place in calendar week 25-29 and account for 25.5 credits (Table 4.A.3). Students receive no information on their performance in these courses before calendar week 35. Thus, students receive no information about pivotal grades during this examination period and therefore there exists no meaningful strategic behavior when exams are taken. Fourth, students have basically no chance to enforce a revision of grades.[42]

Once all compulsory courses are completed, the ultimate criterion for passing the ASY is threefold. First, individuals have to accumulate 240 (220 in 2001) credit-grade-points, which corresponds to an overall average grade of 4.0. Second, individuals must not accumulate more than 12 so-called minus credits (MC), a rule that we will elaborated on below. Third, individuals have to submit a proof of sufficient accounting skills.

The further analysis will concentrate on the second criterion, for the following reasons. The first criterion is rarely violated if the other criteria are fulfilled, i.e. only 3 individuals in the cohorts of 2001 - 2006 failed the ASY because of an insufficient number of credit-grade-points. These individuals are excluded from the analysis. Furthermore, violating the third criterion means neither passing nor failing the ASY. Once the other two criteria are fulfilled, violating the third criterion allows for conditional acceptance into the Bachelor. Yet, students have to pause and submit a proof of accounting skills in the meantime. As this group is particular and rather small, we will also exclude them from the analysis (see section 4).

[42]Only in the case of obvious mistakes during grading, grade revision is unequivocally granted. In all other cases, the individuals have to file a case (so-called recourse). Yet, in the data for 2001 - 2006, we observe 2 cases with insufficient performance according to the data who are still observed in the Bachelor afterwards. These individuals might have won a case for grade revision.

Thus, the decisive rule for passing vs. retention is given by a strict threshold of 12 minus credits which will later be exploited as a treatment rule. For any course, a student receives minus credits if he obtains a grade below 4. Minus credits in the respective course are then defined as the difference between the actual grade and a grade of 4, multiplied by the number of credits for this course. E.g., if a students receives a grade of 3.5 in a course with 4 credits, he obtains 2 minus credits for this course. To describe the accumulation of minus credits more formally, suppose that the overall number of compulsory subjects in the ASY is S. Let G_s be the grade obtained in a course, and C_s the number of credit points associated with the subject. The total sum of minus credits (MC) is then calculated as

$$MC = \sum_{s=1}^{S} (4 - G_s) * C_s * \mathbf{1}(G_s < 4) \qquad (1)$$

Note that minus credits cannot be compensated for by grades greater than 4 in other subjects. If $MC > 12$, the respective student fails the ASY and cannot directly proceed to the Bachelor level, i.e. the student is retained. Yet, he is allowed to repeat the *full* ASY. If the ASY is successfully passed in the second attempt, the student is admitted to the Bachelor level and can proceed. In case of failing again, the student is coercively exmatriculated. Enforcement of this rule is strict (see figure 4.A.1).

Prior to entering the Bachelor level, individuals choose their Bachelor specialization (Business, Economics, Law and Economics, or International Affairs). During the Bachelor phase, they complete a number of compulsory courses and electives as well as a Bachelor thesis, but are free to set the pace of degree completion themselves. On-time graduation follows after 4 Bachelor semesters, but only a minor fraction of students manages to complete their Bachelor degree within this time frame. Yet, once individuals have entered the Bachelor level, more than 84% graduate within at most 6 Bachelor semesters.

4 Data

4.1 General description

The analyses in this paper are all based on administrative records from the University of St. Gallen and are, thus, free of measurement error. The data consist of enrollment and course data at the student level and cover the population of all students entering the ASY between 2001 and 2010. Yet, students entering the ASY after 2006 are excluded from the analysis as long-term outcomes are unobserved for this group (i.e. for the latest cohorts we only have incomplete information about their performance at the Bachelor level).

The enrollment data are on a half year basis (by semester) and contain the following information. First, we can observe whether the ASY has been passed successfully or not. Second, for each student who fails the first year, we observe whether he repeats the ASY or drops out, respectively. Third, the data contain information about major choice (mostly Business or Economics) at the Bachelor level. Forth, we observe individual characteristics at the data of university entry, i.e. age, sex, nationality, mother tongue, country of origin as well as region of origin for students from Switzerland, type of high school degree and in which country or region it has been obtained, as well as the date of high school graduation.

The course file contains information on individual performance at the course level for each semester, i.e. grades and credits for each completed course. This information is crucial at different points in the analysis. First, on the basis of course information, we restrict the estimation sample to first-year students who have completed all compulsory courses as required by the curriculum. Second, it allows us to compute the precise number of minus credits obtained by each student. Third, we use grades and credit points at the Bachelor level as measures of academic performance. In order to capture the pace of degree completion, we use credits obtained by the end of each Bachelor semester. As a measure of the quality of performance, we use standardized grades. In order to account for grade inflation, grades are standardized at the level of Bachelor entry cohort.

Thus, the initial sample consists of all entering first-year students with German mother tongue who start the Business/Economics track at the University of St. Gallen (n = 3762, see Table 4.A.4). This sample is homogenous in the sense that first, all Business/Economics students have to complete the same courses during their freshman year, and second, these students face exactly the same exam conditions. The latter is not the case for students with foreign mother tongue as they might have longer exam durations. Table 4.A.4 shows that the compositions of the student cohorts remain approximately stable over the years in terms of background characteristics. The large majority of students is male (73%). Moreover, most students enter the ASY when they are 20 or 21 years old. Foreign students account for 22% on average. This number is low due to admission rules: The fraction of students with neither high school diploma from Switzerland nor Swiss citizenship is restricted to at most 20% of the student body. As mentioned, these students are selected on the basis of an entrance test.

For the main analyses in this paper we only consider students who have completed all first-year courses. For all other freshman students, the assumption of random assignment to retention is problematic, since course non-completion might relate to unobserved characteristics which also affect subsequent outcomes. Therefore, the following types of students are dropped:

- students who have not completed all mandatory first semester courses (type 1)
- students who have completed all mandatory first semester courses, but have exceeded the threshold of 12 MC already in the first semester (type 2)
- students who have passed the first semester but have dropped out voluntarily after

Table 4.1: Descriptive statistics: Entering first-year students, by type

	(1)	(2)	(3)	(4)	(5)	(6)
	1st sem: Not all exams	1st sem: MC > 12	1st sem: Voluntary dropout	2nd sem: Not all exams	2nd sem: Accounting failed	2nd sem: All exams
Background Characteristics						
Age $\leq$ 19	10%	15%	12%	17%	15%	12%
Age 20/21	52%	53%	62%	51%	56%	68%
Age $\geq$ 22	38%	32%	27%	32%	29%	21%
Nationality foreign	20%	16%	0%	19%	37%	23%
Entrance test	5%	5%	0%	12%	33%	18%
Matura in SG	20%	16%	12%	14%	23%	15%
Minus Credits						
# MC in 1st sem.	-	19.09	8.38	7.53	4.05	1.60
# MC in first year	-	-	-	-	13.85	4.26
Failed: MC > 12	-	-	-	-	53%	11%
Course Performance						
Business 1 taken	28%	-	-	-	-	-
Grade Business 1	2.90	3.09	3.60	3.71	3.96	4.24
Math 1 taken	5%	-	-	-	-	-
Grade Math 1	3.27	2.83	3.69	3.58	3.99	4.54
Econ 1 taken	26%	-	-	-	-	-
Grade Econ 1	2.98	3.15	3.68	3.91	4.15	4.60
Business 2 taken	-	-	-	34%	-	-
Grade Business 2	-	-	-	3.07	3.77	4.34
Math 2 taken	-	-	-	26%	-	-
Grade Math 2	-	-	-	2.76	3.86	4.56
Econ 2 taken	-	-	-	34%	-	-
Grade Econ 2	-	-	-	3.27	3.76	4.46
Term paper taken	-	-	-	64%	-	-
Grade Term paper	-	-	-	4.36	4.77	4.97
Bachelor						
Bachelor started	10%	26%	-	43%	56%	97%
# obs	225	365	26	90	73	2983

Note: The sample includes all first-year students with German mother tongue entering between 2001-2006 into the Business/Economics track.

the first semester (type 3)

- students who have not completed all courses in their second semester (type 4)
- students who fail the mandatory accounting test and therefore cannot be promoted (type 5)

All other students (denoted as type 6) are included in the final estimation sample (n = 2'983). The estimation sample is therefore a selected sample of all entering first-year students and -as such - endogenous. We describe the different types of students in Section 4.2 in order to deepen our understanding about ongoing selection processes throughout the first year.

4.2 What happens during the first year?

Our final estimation sample contains only students who complete all first-year exams (see Table 4.2). These account for 79% of all freshman with German mother tongue who enter the Business/Economics track. Accordingly, 21% of all freshman miss at least one requirement of the ASY. In order to understand the selection dynamics throughout the first year we examine several pathways that lead to the exclusion from our final estimation sample. For this purpose, we classify students according to 6 types as mentioned in Section 4.1. Classification is based on three criteria: first, whether all main exams have been taken (first semester: Business 1, Econ 1, Math 1, second semester: Business 2, Econ 2, Math 2, Term paper), second, whether students have exceeded the threshold of 12 MC in the first or second semester, respectively, and third, the point in time when they drop out. An additional criterion is passing the mandatory accounting exam by the end of the first year. Students who have not passed this exam are "blocked", i.e. must not take any courses for at least one semester. Table 4.2 shows the relative size for each of the type sub-groups.

It can be seen that dropout is particularly high in the beginning of the first year. Among the first-semester dropouts, we can distinguish between three groups. First, 6% of students do not complete all main first-semester exams (type 1). These students might have been discouraged already in the beginning. Second, 10% of students take all required exams, but exceed the threshold of 12 MC already by the end of the first semester (type 2). As a result these students are not allowed to enter the second semester. Third, a minor fraction of 1% decides to dropout despite successfully passing his first-semester courses (type 3). Among the three groups, type-1-students appear as the lowest performing students (see Figure 4.1). Considering only the exams these students have taken, their median performance is slightly lower than the performance of the median type-2-student. Unsurprisingly, late voluntary dropouts (type 3) have better grades than early voluntary dropouts (type 1). Yet, their performance is still very often below the passing grade of 4. Overall, a visible descriptive relationship between performance and dropout exists during the first semester. It is however unclear whether students drop out due to their low expected performance, which would be

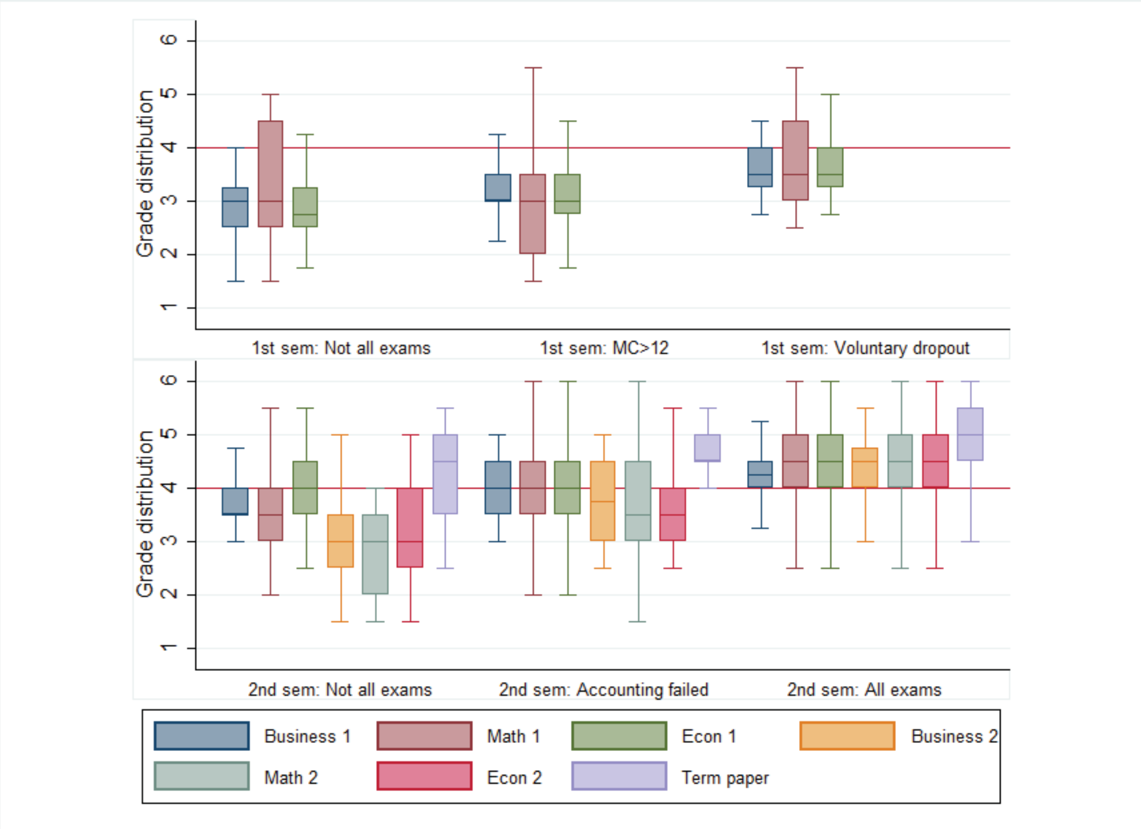
Table 4.2: Description of student types

Type	Main exams taken	MC 1st sem.	Dropout	# Obs.	% of sample
1	Max. 2 in 1st sem.	-	Voluntary dropout in 1st sem.	225	6%
2	All in 1st sem.	> 12	Blocked for 2nd sem.	365	10%
3	All in 1st sem.	≤ 12	Voluntary dropout after 1st sem.	26	1%
4	All in 1st, but not in 2nd sem.	≤ 12	Voluntary dropout in 2nd sem.	90	2%
5	All in both sem.	≤ 12	Delay (missing accounting exam)	73	2%
6	All in both sem.	≤ 12	No dropout / Voluntary dropout	2983	79%
Total				3762	100%

Note: The sample includes all first-year students with German mother tongue entering in 2001 - 2006 into the Business/Economics track. Type definitions are presented in Section 4.3.

in line with the idea of "schooling as experimentation" (Manski (1989)), or whether dropout and performance are confounded by other unobserved factors, or both.

Figure 4.1: Grades, by type



Note: Box-plots of the grade distributions, by type. The sample includes all first-year students with German mother tongue entering in 2001 - 2006 into the Business/Economics track. Type definitions are presented in Section 4.3.

Dropout during the second semester is low, i.e. only 4% stop after having entered the second semester. Half of them drop out voluntarily as they do not take all required exams in the second semester (type 4), and the other half is delayed due to a missing accounting exam. Again, better performing students tend to drop out at a later stage, or stay (Figure 4.1).

Table 4.1 also gives some indication about the extent that types might be confounded by observed characteristics. With respect to age, no clear pattern exists. In contrast, foreign students and students with an entrance test tend to stay longer. This observation is also in line with the considerations of Manski (1989) as initial costs of taking up studies in St. Gallen is higher for foreign students (e.g. studying for the entrance exam, moving to a foreign country, higher student fees for foreign students), and thus their initial performance expectations must also be higher in order to have positive expected gains of taking up a degree. Similarly, students who completed their high school diploma in the canton of St. Gallen are overrepresented among the first three types as they might have had lower costs in the beginning.

4.3 Estimation sample

Table 4.3 describes the final estimation sample (type 6). It can be seen that their cohort characteristics are relatively stable over the years, except for slight changes in the age distribution. Moreover, the estimation sample is comparable to the initial overall sample of entering first-year students in terms of student background characteristics (see Tables 4.A.4 and 4.3). Nevertheless, we can not rule out that the sample is selected, i.e. it differs from the overall freshman population in terms of their unobservable characteristics (see Section 4.2). Regarding the educational outcomes of our estimation sample we observe the following: First, the fraction of students who do not accumulate any minus credits amounts to 56% on average. Moreover, students tend to accumulate larger amounts of minus credits in the second semester. In total, 11% fail the ASY, and 10% repeat, implying that dropout rates after retention are low. The fact that 97% of all freshman are observed to enter the Bachelor level implies that, indeed, most of the repeaters must have repeated successfully.

Treatment assignment is random in the specified sample at the cut-off of 12 minus credits according to our argument outlined in Section. Table 4.4 provides descriptive support for this point. First, mean background characteristics are approximately stable at the cutoff value. Age is slightly lower at the cutoff, but on both sides of the cutoff, which does not threaten the analysis. The most worrisome observation is that the fraction of individuals with an entrance degree from St. Gallen is considerably higher among students just below the cutoff. Second, mean grades in the main subjects are also smoothly distributed. We use non-standardized grades in order to illustrate this point. Given that grades are distributed smoothly, manipulation on the part of graders seems unlikely. Third, in terms of outcomes, dropout rates after the first year increase sharply at the threshold, but they are still very low for retained students. This fact is also reflected by the high number of students starting the Bachelor degree to both sides of the cutoff.

5 Identification Strategy and Estimation

The passing requirements for first year students provide a strictly enforced rule that allows us to obtain a local estimate of the causal effects of retention and repetition on future educational outcomes, i.e. we use the threshold value (c) of 12 minus credits for identification. Just at the threshold retention is quasi-random, and we exploit this feature in the following. Descriptive evidence (shown later) supports this assumption.

In the potential outcome framework, Y_i^0 and Y_i^1 are the outcomes of an individual in a state without and with retention, respectively. Note that, however, for each student, only one state is observed at any moment in time, i.e.

$$Y_i = Y_i^1 * R_i + Y_i^0 * (1 - R_i) \tag{2}$$

Table 4.3: Descriptive statistics: Estimation sample, by year

	Observations	2001 - 2006	2001	2002	2003	2004	2005	2006
Characteristics								
Male	2983	74%	74%	74%	75%	75%	75%	72%
Age < 20	2983	12%	8%	8%	13%	15%	12%	17%
Age 20/21	2983	68%	68%	67%	69%	69%	70%	63%
Age > 21	2983	21%	24%	24%	18%	16%	18%	20%
Foreign nationality	2983	23%	19%	24%	25%	27%	24%	21%
Entrance degree from SG	2983	15%	13%	14%	16%	18%	14%	16%
Entrancetest	2983	18%	15%	18%	21%	21%	17%	16%
Grades								
Business 1	2983	4.24	4.24	4.24	4.20	4.23	4.26	4.28
Math 1	2983	4.54	4.43	4.62	4.49	4.64	4.55	4.56
Econ 1	2983	4.60	4.60	4.50	4.56	4.81	4.63	4.55
Business 2	2983	4.34	4.32	4.14	4.46	4.79	4.04	4.44
Math 2	2983	4.56	4.53	4.47	4.61	4.89	4.56	4.40
Econ 2	2983	4.46	4.41	4.30	4.26	4.44	4.54	4.77
Term paper	2983	4.97	4.90	5.00	4.92	4.98	5.00	5.03
Minus Credits								
MC > 0 in first year	2983	56%	63%	59%	56%	51%	55%	50%
# MC in first semester	2983	1.60	1.51	1.54	1.85	1.47	1.69	1.56
# MC in first year	2983	4.26	4.16	4.88	4.14	3.78	4.94	3.62
Fail Total	2983	11%	12%	12%	10%	9%	15%	9%
Retention								
Repeater	2983	10%	10%	11%	9%	9%	14%	8%
Bachelor								
Bachelor started	2983	97%	95%	96%	98%	98%	96%	97%
Bachelor ≤ 4 semesters	2886	37%	47%	50%	39%	33%	26%	23%
Bachelor ≤ 5 semesters	2886	61%	69%	71%	64%	60%	55%	49%
Bachelor ≤ 6 semesters	2886	84%	87%	89%	88%	86%	80%	77%
Observations		2983	609	498	426	380	511	559

The sample includes all first-year students with German mother tongue entering in 2001 - 2006 into the Business/Economics track who have completed all compulsory courses. The last three rows (Bachelor duration $\leq 4/5/6$ semesters) are only specified for students starting a Bachelor degree.

Table 4.4: Descriptive statistics: Estimation sample, by minus credits

	Observations	Total	0-4	4-8	8-10	10-12	12-14	14-16	>16
Background Characteristics									
Male	2983	74%	74%	73%	75%	73%	71%	73%	75%
Age $<$ 20	2983	12%	13%	10%	17%	6%	9%	5%	13%
Age 20/21	2983	68%	70%	65%	56%	69%	63%	58%	63%
Age $>$ 21	2983	21%	18%	25%	26%	24%	29%	38%	24%
Foreign nationality	2983	23%	26%	16%	17%	15%	14%	13%	18%
Entrance degree from SG	2983	15%	14%	16%	10%	23%	13%	17%	18%
Entrancetest	2983	18%	23%	10%	10%	7%	4%	6%	7%
Grades									
Business 1	2983	4.24	4.44	3.99	3.96	3.76	3.70	3.73	3.64
Math 1	2983	4.54	4.83	4.20	4.00	3.92	3.89	3.76	3.61
Econ 1	2983	4.60	4.88	4.24	4.04	3.95	3.89	3.88	3.82
Business 2	2983	4.34	4.58	4.10	4.03	3.92	3.71	3.64	3.42
Math 2	2983	4.56	4.89	4.17	3.93	3.97	3.95	3.67	3.39
Econ 2	2983	4.46	4.77	4.17	3.94	3.80	3.70	3.64	3.33
Term paper	2983	4.97	5.10	4.83	4.71	4.70	4.75	4.52	4.56
Minus Credits									
# MC in first semester	2983	1.60	0.27	2.37	3.52	4.59	5.35	5.41	7.38
# MC in first year	2983	4.26	0.69	6.00	9.09	11.02	13.08	15.05	21.78
Retention and Dropout									
Repeater	2983	10%	-	-	-	-	91%	92%	88%
Dropout after 2 sem	2983	2%	1%	0%	0%	2%	9%	8%	13%
Bachelor									
Bachelor started	2983	97%	99%	99%	100%	98%	83%	83%	74%
Bachelor $\leq$ 4 semesters	2886	37%	41%	31%	27%	21%	21%	25%	22%
Bachelor $\leq$ 5 semesters	2886	61%	68%	52%	45%	37%	45%	49%	44%
Bachelor $\leq$ 6 semesters	2886	84%	90%	76%	73%	74%	66%	70%	63%
Observations		2983	1'975	432	126	108	70	64	208

The sample includes all first-year students with German mother tongue entering in 2001 - 2006 into the Business/Economics track who have completed all compulsory courses. The last three rows (Bachelor duration $\leq$ 4/5/6 semesters) are only specified for students starting a Bachelor degree.

with $R_i = 1$ if the individual is retained, and $R_i = 0$ if the individual is not retained. In general, we are interested in the following difference which represents the mean effect of retention on future outcomes

$$\tau = E[Y^1|R] - E[Y^0|R] \tag{3}$$

where Y is the educational outcome of interest and R the binary retention or repetition (the treatment) status which jumps from zero to one at the cut-off value of 12 minus credits.

Ultimately we are interested in getting a consistent estimate of the difference in educational outcomes of students that were retained after the first year (and repeated) and the outcomes of the same students had they passed, *ceteris paribus*. However, this effect cannot be revealed from the data without further assumptions. This is because retention status is non-random. A conditional mean comparison (conditional only on observable characteristics X) between all students for which $R = 1$ and all students for which $R = 0$ would reveal the treatment effect of interest only if unobservable characteristics U were identically distributed in both groups. Since, however, U is likely to contain factors (e.g. motivation or ambition) that are systematically different across the two groups (e.g. students that passed are also more ambitious) conditional mean comparison between the overall groups of students who failed and students who did not fail will lead to a biased estimate of the retention effect as these groups are generally not comparable. Thus, we can not assume that $Cov(R, U|X) = 0$ in our setting.

Yet, students who *just* passed are expected to be fairly similar regarding their distribution of U compared to students who *just* failed the first year. Consequently, being retained is assumed to be quasi-random around the threshold. Following this logic, we restrict the identification of the retention effect to this local sub-population.

In the related literature, this assumption is known as the *local continuity assumption* (Imbens and Lemieux (2008)). The assumption implies that, had the individuals within a small window around the threshold been exposed to the same policy, they would have achieved on average the same outcome. Hence, as a consequence of student's inability to precisely control the number of minus credits achieved, in the local neighbourhood around the threshold retention is as good as random and, thus, allows us to identify the effect of interest around the discontinuity point, i.e.

$$\tau_{RD} = E[Y^1 - Y^0|MC = c] \tag{4}$$

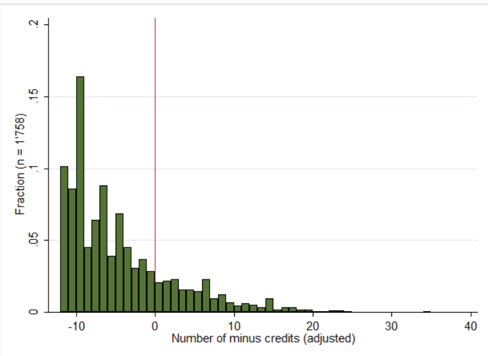
While we can test whether the observed covariates X are continuously distributed around the threshold, the assumption that this is also true for unobserved characteristics U is an identifying assumption that cannot be tested.

The *local continuity assumption* further implies that, for our identification strategy to be valid, we have to ensure that students who end up close to the critical threshold could not perfectly anticipate on which side of the threshold they will be placed. This assumption seems very reasonable in our setting. While students are aware of the threshold *ex-ante* (as it is announced in the rules of the ASY) it is unlikely that they are able to sort themselves just above or just below the threshold *once they have taken all exams*. Another argument against strategic sorting is that grades and minus credits are not perfectly predictable from the perspective of the student. Grading schemes are often designed after the exams are taken and the students do not have any control about that process. Moreover, grading schemes are solely decided upon by the teachers.[43]

In addition to these arguments, we examine the assumption of *local continuity* by investigation of the density of minus credits achieved on either side of the cut-off value (as proposed by Lee and Lemieux (2010)). If density plots are smooth at the threshold, we can be confident that no sorting took place.

Continuity is depicted by Figure 4.2, which shows a histogram of the (recentered) number of minus credits. The aggregated bins depict a smoothed version of the minus credits distribution. However, bin-width is chosen in order not to smooth over the threshold-value. The vertical line represents the cut-off value. Overall, it can be seen that the distribution of minus credits is heavily right skewed. At the same time there is no visual indication of sorting around the threshold. Likewise, the McCrary test (McCrary (2008)) does not reject the null-hypothesis of continuity of the running variable at the cut-off c (log difference in height=0.05, p-value=0.75).

Figure 4.2: Histogram of the assignment variable



Note: The assignment variable is defined as the amount of minus credits accumulated during the ASY. Minus credits are adjusted by subtracting the cut-off value (12.25 minus credits) from the actual amount of minus credits. The sample consists of all individuals in the estimation sample (cohorts 2001 - 2006) who have accumulated at least 0.25 minus credits (n = 1669).

Further using the threshold value in our identification strategy, we examine the effects of retention and repetition on various outcomes: drop-out probability after first year at

[43]Hence, the only way to purposely achieve a position just above the threshold is to apply for a revision of grades. However, from administrative sources we know that the number of individuals who manage to shift themselves below the critical cut-off value as a result of a revision process is, if anything, marginal (in most years even non-existent), as explained in Section3.

university (binary), whether a student is ever observed at the Bachelor level (binary), choice of major studies (binary) as well as continuous educational outcomes (credits and the grades) over the subsequent semesters at the Bachelor level.

In addition to simple mean outcome comparisons at the threshold, we estimate models of the following form:

$$Y = \alpha + \beta * 1(MC \geq 0) + \sum_{k=1}^{K} \gamma_k * MC^k + \sum_{k=1}^{K} \nu_k * MC^k * 1(MC \geq 0) + \varepsilon \tag{5}$$

where Y represents the various educational outcomes, MC corresponds to the recentered (minus 12.25) number of minus credits collected by the student at the end of the first year and k is a flexibility parameter. In all specifications the coefficient of interest is β - it represents the causal effect of retention on the outcomes. We further use varying windows of data around the threshold to assess the robustness of our findings. By using higher order polynomials and interaction terms we allow for a non-linear relationship as well as different slopes on both sides of the cut-off values. We also provide nonparametric estimates based on the guidelines provided by Imbens and Lemieux (2008). These are based on kernel methods (where the optimal bandwidth is computed by cross-validation) using local linear regressions to estimate the boundary points on each side of the threshold. As with our parametric specifications, the effects of interest are identified by the differences in the expected means of the outcomes on either side of the threshold. The results for our preferred specifications are reported in the result section, while various other specifications are to be found in the appendix.

Using the same model specifications we also investigate the local continuity of predetermined covariates around the threshold to debilitate concerns related to strategic sorting and, thus, support the validity of our approach. Table 4.5 presents the coefficient estimates. While there is no evidence for gender or age-related sorting around the threshold, there is some indication for discontinuities with respect to the origin of students - i.e. we find some evidence that students who just fail are less likely to be from "nearby St. Gallen" (as already mentioned in the data section). To account for such differences, we estimate all models for educational outcomes with and without covariates (covariates included correspond to the ones in Table 4.5) to check the robustness of our findings.

Our identification strategy faces two further difficulties. First, we are interested in performing same-grade comparisons, i.e. educational outcomes of retained students are compared to the same outcomes of non-retained students. By the nature of the problem, however, the outcomes of the two groups are measured one year apart (the outcome of retained students is typically measures one year later). However, it might well be that outcome distributions are not *per se* comparable across years. To account for potential problems that are due to changes in grade distributions over time we standardize on the level of the Bachelor level entry cohort.

Second, the problem of non-random dropout might bias the results when we investigate future educational outcomes other than dropout decisions that occur right after the first year (i.e. after the second semester). As discussed, retention might induce individuals to leave university (instead of repeating the first year). The identification results presented so far are only valid under the assumption of random dropout. If dropout, however, is selective our estimates would be biased. Assume that the less able students are more likely to drop out as a result of failing the first year - i.e. only the more able remain and repeat. In such a setting, one might estimate a positive effect of repeating the first year on subsequent outcomes (while the true effect is zero) just because of selective dropout. We clearly acknowledge such concerns. Yet, out of the 342 students in our estimation sample who are retained only 37 students (or 11%) drop out. Moreover, this number is smaller at the threshold (around 7%). We do not consider this share to be substantial and, thus, abstain from extending the analyses towards partial identification strategies that would only allow to estimate bounds for the effects of interest.

Table 4.5: RDD Estimates: Pre-determined Characteristics

Column	(1)	(2)	(3)	(4)
Male	-0.04	-0.14	-0.17	-0.05
	(0.09)	(0.10)	(0.15)	(0.14)
Age 17/19	0.02	0.04	-0.04	-0.02
	(0.10)	(0.11)	(0.15)	(0.15)
Age 20/21	-0.04	-0.06	-0.10	-0.07
	(0.10)	(0.11)	(0.16)	(0.17)
Age 22+	0.02	0.01	0.14	0.05
	(0.05)	(0.06)	(0.09)	(0.09)
Gap year	-0.06	-0.10	-0.14	0.02
	(0.10)	(0.12)	(0.16)	(0.17)
Entrance degree from SG	-0.13	-0.15*	-0.21*	-0.24*
	(0.08)	(0.09)	(0.13)	(0.14)
Foreign citizen	0.00	0.01	0.04	-0.02
	(0.07)	(0.08)	(0.11)	(0.11)
Entrance test	-0.02	0.00	-0.01	-0.03
	(0.05)	(0.05)	(0.07)	(0.06)
Estimation Window	[1;1]	[8;8]	[12;12]	[12;12]
Polynomial order	0	2	4	NP
Observations	94	908	1615	1615

Note: RDD estimates of Dropout behaviour and Major choice. Columns (1) - (4) display different specifications. The parametric specifications (1)-(3) are estimated using a linear probability model. Following Imbens and Lemieux (2008) the bandwidth for the local-linear nonparametric specification (NP) is determined by cross-validation. The respective estimation window for each specification is reported as the minus credit range on each side of the threshold.
* Significant at 10%- level, ** Significant at 5%- level, *** Significant at 1%- level. Standard errors in parentheses.

6 Results

This section provides visual evidence as well as regression estimates of the effects of retention on academic outcomes based on the discontinuity design (i.e. for students close to the cut-off value of 12 minus credits) as explained before. At first, we investigate how retention affects the dropout decisions of students after their first year, i.e. before enrollment into their third semester. To which extent causes retention students to drop out, immediately or throughout their second attempt of the ASY, respectively? After that, the effect of retention on the choice of the Major at the Bachelor level is discussed. Finally, we compare the academic outcomes of repeaters vs. non-repeaters at the Bachelor level. In particular, we focus on the number of credits accumulated in subsequent semesters as well as the corresponding grade point averages (GPA) by the end of each Bachelor semester to see if student retention at the college level can lead to improved learning and academic performance at later stages throughout the Bachelor.

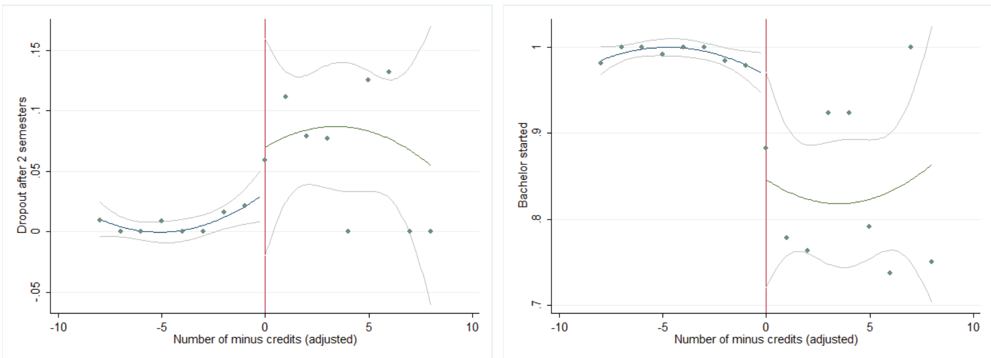
6.1 The effect of retention on dropout

Interpreting dropout decisions as an utility maximization problem Manski (1989) pointed out that students weigh their expected utility from dropping out against their continuation utility. Within the institutional framework we explained earlier, retained students incur higher continuation costs than non-retained students. They face not only one additional year of foregone earnings, but also the risks of failing the ASY for a second time. In addition, being separated from their entry cohort or being stigmatized as a repeater might be associated with additional (psychological) costs. These considerations, ceteris paribus, suggest that dropout rates are supposedly higher for retained students.

Figure 4.3 shows students' mean dropout rates (left panel) as well as the probability to ever be observed at the Bachelor level (right panel) as a function of the minus credits they accumulate by the end of their first year. The threshold value is depicted by the vertical line together with quadratic regression lines (on each side separately) for illustrative purposes. Looking at the mean values it can be seen that the vast majority (>97%) of students who meet the passing requirements of the first year are, indeed, not observed to drop out, and proceed with their Bachelor studies. On the contrary, the share of retained students who drop out immediately after their first year is larger (approximately 7% for students close to the threshold value) - and is also found to have a larger variance. Unsurprisingly, the probability to be observed at the Bachelor study at some point is also lower for retained student. The figure indicates that of those students who just failed their first attempt, 87% are observed at the Bachelor level later on - the regression line slightly underestimates that share. Given that we see large variances in both outcomes for the sub-sample of retained students we can expect regression estimates to heavily depend on the flexibility of the underlying model as well as the chosen window that they are based on.

In fact, this is found in the upper panel of Table 4.6 which presents various regression estimates of the effects of retention on dropout immediately after the first year (i.e. after the second semester[44]) We ran numerous specifications of the model, but only show the results based on our preferred model specifications. In order to account for the tradeoff between bias and precision, the specifications grow more flexible as the sample size (i.e. the estimation window) increases (Lee and Lemieux (2010)). Therefore it should come as no surprise that the standard errors of the coefficient estimates eventually increase - despite the larger bandwidth - as a result of the higher order polynomials. We also enrich the specification to include covariates in order to balance potential differences that are due to observable characteristics. In addition to parametric models we provide non-parametric estimates using the approach suggested by Imbens and Lemieux (2008).[45]

Figure 4.3: RDD estimates: Probability of immediate dropout and starting a Bachelor degree



Note: The panels above provide a graphical illustration of the RDD estimates of the probabilities of dropout after 2 semesters and starting a Bachelor degree. The green dots represent the mean outcomes within each minus credit category. The green lines display a quadratic fit to either side of the cutoff (95% confidence intervals in gray). The sample consists of all individuals in the sample within a range of 8 minus credits to either side of the cutoff (n = 908).

In line with the visual evidence most (i.e. all parametric) point estimates are positive. Moreover, standard errors generally increase despite growing bandwidths as a consequence of the added model flexibility. The non-parametric estimate in column (5) is negative, but also suffers from a larger standard error.[46] Overall, none of the estimates with respect to immediate student drop-out is statistically significant - which is supposedly caused by the relatively large variance in dropout rates for students above the cut-off point. The largest point estimate is found by the simple mean comparison in column (1) which shows a 6 percentage points higher dropout rate for the retained students which we still consider to be

[44]Dropouts are defined as students who are not observed to enroll for the third semester.

[45]We only consider results based on samples that exclude students with zero minus credits, as they are uninformative for our purposes.

[46]This property is a general problem of the local-linear estimator when applied to binary outcomes (Frölich (2006)).

Table 4.6: RDD estimates: Dropout and Major Choice

Column	(1)	(2)	(3)	(4)	(5)
		Selection			
Dropout after	0.06	0.04	0.00	0.01	-0.01
2nd semester	(0.05)	(0.05)	(0.06)	(0.06)	(0.10)
Bachelor Started	-0.15***	-0.12**	-0.04	-0.05	-0.05
	(0.06)	(0.06)	(0.08)	(0.08)	(0.11)
Estimation Windows	[1;1]	[8;8]	[12;12]	[12;12]	[12;12]
Polynomial order	0	2	4	4	NP
Observations	94	908	1615	1615	1615
Covariates	No	No	No	Yes	No
		Major Choice			
Economics major	0.10*	0.18***	0.09	0.11	0.10*
	(0.05)	(0.06)	(0.08)	(0.08)	(0.07)
Buisness major	-0.12	-0.22	-0.20	-0.21	-0.16
	(0.10)	(0.11)	(0.14)	(0.14)	(0.13)
Estimation Window	[1;1]	[8;8]	[12;12]	[12;12]	[12;12]
Polynomial order	0	2	4	4	NP
Observations	82	815	1484	1484	1484
Covariates	No	No	No	Yes	No

Note: The table shows RDD estimates of the the effect of student retention on dropout behaviour and major choice. Columns (1) - (5) display different specifications. The parametric specifications (1)-(4) are estimated using a linear probability model. Following Imbens and Lemieux (2008) the bandwidth for the local-linear nonparametric specification (NP) is determined by cross-validation. The respective estimation window for each specification is reported as the minus credit range on each side of the threshold. Covariates include the following indicator variables: Cohort dummies, Male, Younger than 20 by the start of the ASY, Older than 21 by the start of the ASY, Non-Swiss nationality, Non-German mother tongue, Entrance degree from St. Gallen, Entrance test participation, Gap year after finishing high school.
* Significant at 10%-level, ** Significant at 5%-level, *** Significant at 1%-level. Standard errors in parentheses.

modest. Hence, we can conclude that despite the substantial costs associated with retention, its effect on immediate dropout seems negligible.

In contrast to immediate student dropout, the negative effect of retention on the probability to be observed at the Bachelor level is more pronounced. Columns (1) and (2) suggest that the effect ranges from -12 to -15 percentage points and is statistically significant. The estimates become less precise and insignificant, however, when the larger observation windows with the more flexible models are considered. The reasons for the disparity between immediate dropout and later dropout (i.e. not being observed at the Bachelor level) are threefold. First, students might enroll into the third semester and profit from their student status while looking for outside opportunities, but are not observed to be enrolled in the fourth semester, or, second, students might update on the costs only after having started their second attempt to pass the ASY (together both account for 13 cases). Third, students might fail the ASY for a second time (25 cases). However, once failed, the chance of going through in their second attempt is relatively high (approximately 90%).

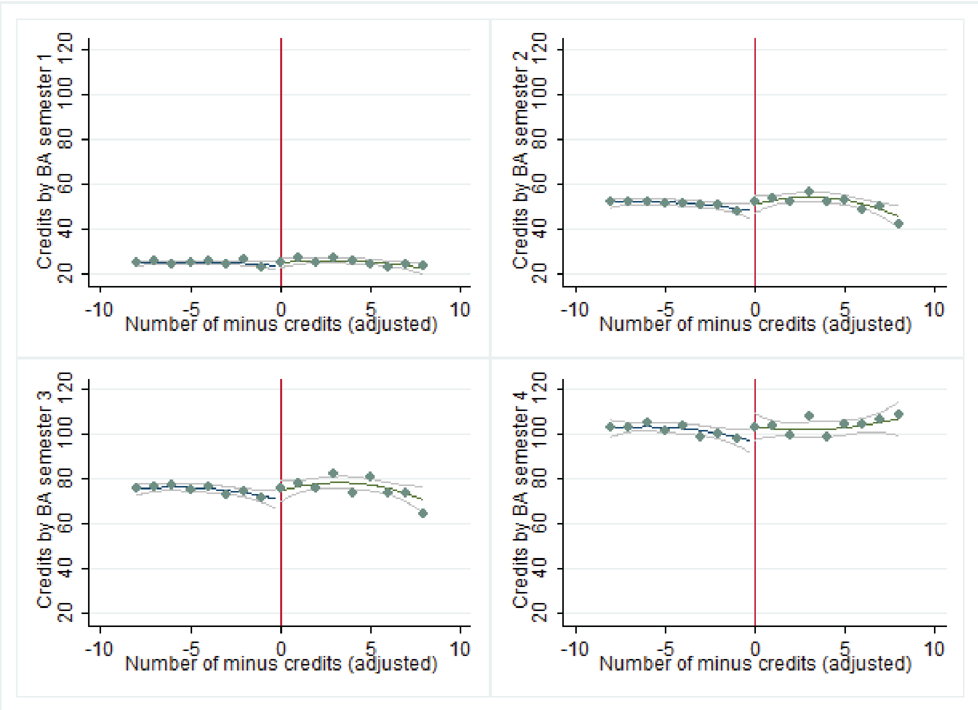
Overall, the analysis of the dropout effects suggests that, first, immediate dropout rates are *ceteris paribus* not significantly higher for retained students. Moreover, given that they have to complete the ASY again and face the additional risk of a second failure, the share of retained students close to the cut-off which is lost at some point during their second attempt in the ASY - and hence not observed at the Bachelor level - can be considered as moderate. These findings are suggestive of the high utility that students have from staying enrolled despite their failed first attempt. One reason for this observation might be that the expected benefit from obtaining a degree from the University of St. Gallen is high, especially as the University is a renowned Business School which grants good earning prospects to their graduates.

6.2 The effect of repeating on academic outcomes

Before we discuss the effects of retention on future academic outcomes, note that all related model estimates are based on the sample of students who are observed at the Bachelor level. We start by examining major choice - which is the first decision students face after successful completion of the ASY. Major choice is a relevant outcome to look at for two reasons. First, major choice can determine a student's human capital formation and future earnings. Second, differences in choices of major between individuals in the treated and control group should be accounted for in the further analysis as students' performance (grades, credits) might differ *ceteris paribus* across majors.

The lower panel of Table 4.6 investigates the effects based on various specifications of RDD models. The table shows significant differences in major choice for some specifications. Looking at columns (1) and (2) as well as the non-parametric estimate in (5) it appears that retained students are on average 10-18 percentage points more likely to favour Economics as their major. The other point estimates are of similar magnitude, but statistically insignificant (again, supposedly, caused by the added model flexibility). In contrast, no significant effect is found when choosing Business major as the outcome.[47] Thus, it seems that retention has an effect on the subsequent choice of the study path. Since we neither observe students' attitudes, nor changes thereof, we can not say much about the underlying reasons. However, we would interpret the finding as a result of a continuous learning process related to the courses taken throughout the ASY where students apparently reevaluate their interests.

Figure 4.4: RDD estimates: Credits accumulated by the end of each Bachelor semester



Note: The panels above provide a graphical illustration of the RDD estimates of the number of credits accumulated for the Bachelor's degree by the end of each of the first four Bachelor semesters, respectively. The green dots represent the mean outcomes within each minus credit category. The green lines display a quadratic fit to either side of the cutoff (95% confidence intervals in gray), respectively. The sample consists of all individuals in the estimation sample within a range of 8 minus credits to either side of the cutoff (n = 819).

We next turn to the outcomes that measure academic performance. The patterns of accumulated credits as well as grade performance over the first four semesters at the Bachelor level are illustrated in Figures 4.4 and 4.5. Again, quadratic regression estimates for retained and non-retained students are depicted in all figures. Figure 4.4 shows that - unlike in the case for the previous outcomes - the number of accumulated credits is a flat and smooth function of the number of minus credits accumulated in the ASY. Given that we interpret this outcome as a measure of study speed, retained and promoted students seem to proceed through their major at a rather similar pace. By and large this also holds when only the

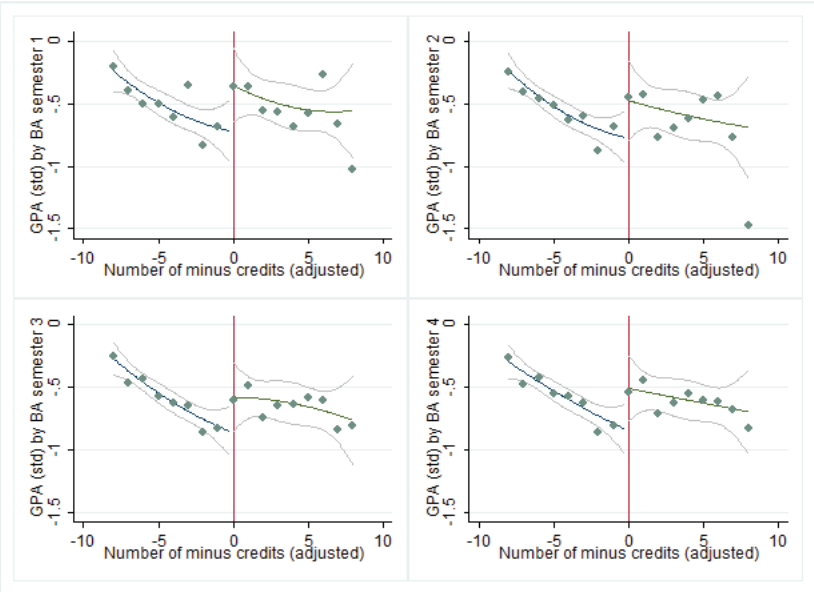
[47]This pattern can occur, because the majors International Affairs as well as Economics and Law are omitted from the estimation table.

Table 4.7: RDD estimates of Bachelor outcomes

Column	(1)	(2)	(3)	(4)	(5)
Credits (std) after 1st BA semester	2.86**	1.50	2.88	2.64	1.47
	(1.42)	(1.42)	(1.42)	(1.42)	(1.42)
Credits (std) after 2nd BA semester	4.81**	3.16	3.92	3.48	3.58
	(2.42)	(2.54)	(3.35)	(3.11)	(2.28)
Credits (std) after 3rd BA semester	5.19	4.29	5.89	5.19	4.53*
	(3.24)	(3.27)	(4.26)	(4.15)	(2.71)
Credits (std) after 4th BA semester	4.33	7.01*	5.00	4.45	4.11
	(3.99)	(3.98)	(5.17)	(5.30)	(3.19)
GPA (std) after 1st BA semester	0.35**	0.37**	0.41	0.41*	0.46**
	(0.15)	(0.19)	(0.26)	(0.25)	(0.21)
GPA (std) after 2nd BA semester	0.29**	0.30*	0.36	0.37	0.34*
	(0.15)	(0.18)	(0.25)	(0.25)	(0.22)
GPA (std) after 3rd BA semester	0.32***	0.29*	0.29	0.31	0.31*
	(0.13)	(0.16)	(0.22)	(0.23)	(0.19)
GPA (std) after 4th BA semester	0.37***	0.34**	0.36*	0.38*	0.33*
	(0.13)	(0.16)	(0.23)	(0.23)	(0.21)
Window	[1;1]	[8;8]	[12;12]	[12;12]	[12;12]
Polynomial order	0	2	4	4	NP
Observations	82	819	1488	1488	1488
Covariates	No	No	No	Yes	No

Note: RDD estimates of Dropout behaviour and Major choice. Columns (1) - (5) display different specifications. The parametric specifications (1)-(4) are estimated using a linear probability model. Following Imbens and Lemieux (2008) the bandwidth for the local-linear nonparametric specification (NP) is determined by cross-validation. The respective estimation window for each specification is reported as the minus credit range on each side of the threshold. Covariates include the following indicator variables: Cohort dummies, Male, Younger than 20 by the start of the ASY, Older than 21 by the start of the ASY, Non-Swiss nationality, Non-German mother tongue, Entrance degree from St. Gallen, Entrance test participation, Gap year after finishing high school. All specifications control for the choice of major studies. * Significant at 10%- level, ** Significant at 5%- level, *** Significant at 1%- level. Standard errors in parentheses.

Figure 4.5: RDD estimates: Grade point averages (GPA) by the end of each Bachelor semester (standardized)



Note: The panels above provide a graphical illustration of the RDD estimates of standardized grade point averages (GPAs) by the end of each of the first four Bachelor semesters, respectively. GPAs are standardized at the level of all individuals who have started their Bachelor degree in the same semester. The green dots represent the mean outcomes within each minus credit category. The green lines display a quadratic fit to either side of the cutoff (95% confidence intervals in gray), respectively. The sample consists of all individuals in the estimation sample within a range of 8 minus credits to either side of the cutoff (n = 819).

cut-off area is investigated. If anything, retained students close to the cut-off accumulate slightly more credit points by the end of their Bachelor studies. This is also supported by the results in Table 4.7 where the effect of retention on accumulated credits is mostly insignificant and the estimated effects become less precise as the estimated models grow more flexible. Based on these estimates there is some weak indication that just retained students have accumulated marginally more credits points already after their first Bachelor semester, and that they extend that lead until the end of the Bachelor studies - however, the effects are rather small and do not compensate for the time lost due to repeating the ASY.

Figure 4.5 shows that the relationship between the number of accumulated minus credits during the ASY and GPAs at all stages of the Bachelor is rather negative when taking each side of the cuf-off value separately. Hence, performance in the ASY appears to have some predictive power with respect to the grades achieved later on. However, the figures also indicate that there is a structural break just at the cut-off value which shows that just retained students appear to achieve better grades in their Bachelor studies than students who just passed the ASY. The quadratic regression lines perform reasonably well in estimating the jump at the discontinuity point. The regression estimates in Table 4.7 clearly support this finding. Taking all estimates together, there is sufficiently strong evidence to state that repeating the ASY leads to significant GPA improvements in the range of 0.29 to 0.46

standard deviations. The point estimates are mostly robust, but at times somewhat too imprecise to be significant when the largest window with the fourth order polynomial is used. Nevertheless, we interpret these findings as a positive causal effect of retention on educational achievements for students who were just retained. Given that students have to repeat all courses before they proceed to the Bachelor level, this finding is perhaps not surprising and in line with the idea of the ASY. However, the fact that the positive GPA effect persists over the entire duration of Bachelor studies is good news for policy makers - it is exactly in line with the original intention of the ASY rules.

Although we find that both, major choice and academic performance after the ASY, are significantly influenced by student retention, we are not able to make any statements about how retention affected academic performance had major choice not been influenced. If retained students only perform better on average because they are more likely to study Economics and GPAs are higher in Economics, the estimated positive effects of retention on GPAs would be upward biased. To overcome such concerns, we simply compared mean GPAs for all Bachelor semesters between the two subgroups of Business and Economics students (not shown here). It turns out that after controlling for minus credits in the ASY, if anything, GPAs of Economics students are lower than GPAs of Business students throughout all Bachelor semesters. Therefore we conclude that improvements in GPAs are not an artifact of major choice, but a direct effect of retention and repetition of the ASY.

Altogether we can summarize that retention and subsequent repeating appears to have a beneficial effect on the grades of students at the Bachelor level. Moreover, this effect is persistent as it lasts throughout the entire observation period. However, there is little indication of a catch up effect in terms of study duration. If anything, just retained students accumulate marginally more credit points. However, these clearly do not compensate for the additional year that they have to spend in the ASY. As such, they "loose" one year, and thus incur considerable opportunity cost when they decide to go for a second attempt.

7 Conclusion

An ever-increasing number of incoming college students is putting existing institutions of higher education in OECD countries under pressure to provide tertiary education in larger quantities while at the same time aiming to maintain their level of quality. Where law prevents these institutions from autonomous ex-ante selection of their incoming students, assessing them in the course of a "probation year" (i.e. the first year) is a feasible alternative. In particular, students are required to meet certain academic standards by the end of their first year - only then are they allowed to proceed - while non-compliance leads to retention. A growing number of institutions (especially in European countries) are nowadays applying comparable frameworks.

Using registry data from the Swiss University of St. Gallen this paper provides empirical evidence on the dynamics and outcomes of such a system. Analyzing six freshmen cohorts from 2001-2006, we find that roughly one fourth of freshman students drops out already before the end of their first year. This happens at different stages of the first year and the reasons are supposedly heterogeneous. Yet, we find evidence for potential deterrent effects that lead some weakly performing students to drop-out before their actual assessment. The remaining three-quarter are observed to take all the required first year exams and form our main estimation sample (a selected sample). Accounting for the endogeneity of students' retention status in our sample by using a regression discontinuity design we argue and show that students who perform just under the retention threshold are sufficiently comparable to students who just pass the first year. Within this selected group, we locally estimate the causal effects of being retained (and subsequently having to repeat the full year) on the subsequent drop-out probability, the choice of major studies and subsequent educational outcomes measured up to four semesters of Bachelor studies.

Given our rather limited sample size (especially around the retention threshold), the clear patterns that we find speak a strong language. Visual presentations confirm that retention increases immediate drop-out of students after the first year - however the regression results suffer from relatively large standard errors and are, thus, not significant. Beyond that, retained students are significantly less likely to ever be observed at the Bachelor level which reflects the combined effect of immediate drop-out as well as forced drop-out due to failing the ASY a second time. In addition, retention tends to influence the choice of major studies in favour of economics. Irregardless of that choice, the effects of retention on subsequent academic performance seem favorable for the policy and persist throughout the Bachelor studies: by the end of the fourth Bachelor semester, retained students show on average significantly higher GPAs than their non-retained comparison group. At the same time, however, we do not find much evidence for increased study speed, i.e. catch up effects can not be detected. Thus, from a policy perspective, retention in higher education appears to be a reasonable measure to improve academic performance - at least when the focus of the policy is on the better performing among the retained - yet, it comes comes at the cost of an additional year that students spend in education.

Admittedly, this study has several limitations. First, the local nature of our identification strategy limits the validity of our results to students who perform close to the minimum passing requirements as set by the university. The vast majority of students in our sample performs significantly better than required, and we cannot make any statements about the effects that retention would have in this group. Nevertheless, we argue that the sub-group that we investigate is the most relevant one from a policy perspective. Retention policies are exactly made to improve the academic performance of students with academic deficiencies that can presumably be straightened out. The relatively low drop-out rate also confirms, that most students are indeed willing undergo a second attempt, i.e. they accept the chance that retention provides them with respect to continuation of their academic career. Second, the limited number of observations does not allow us to look at effect heterogeneities, e.g. across males and females or younger and older students while such analyses would certainly provide additional insights about heterogeneities in learning behaviours (Tinto (1975)). Third, we study retention effects in a particular setting. As mentioned, institutions of higher education are very heterogeneous in terms of the subjects that they offer, the type of students that they attract as well as their specific rules of student assessment. The University of St. Gallen is known to be a high quality business school where graduates are said to have good future job and earning prospects. Hence, we can expect students to accept higher costs (in monetary terms as well as in terms of effort) before they decide to drop-out. Other institutions could attract different types of students where retention might have a stronger (or weaker) effect on motivation and academic improvement, respectively. In this light, further studies from other institutions are needed to improve our knowledge about retention effects in higher education. Ideally, these studies can also go deeper in that they investigate the pathways through which retention affects drop-out behaviours as well as further educational performance.

8 Appendix

Table 4.A.1: Graduation statistics Switzerland

Graduation	Average	2007	2008	2009	2010
Total (N=82233)	20558	17797	20205	21230	23001
Women [%]	51.8	49.0	51.0	52.8	54.4
Foreigners [%]	16.8	16.1	17.4	17.0	16.5
Business Admin. or Econ. (N=12258)	3065	2904	2963	3009	3382
in % of Total	15.0	16.3	14.7	14.2	14.7
Women [%]	31.2	30.4	30.3	31.2	32.8
at University of St. Gallen (N=3640)	910	903	903	881	953
in % of Total	**4.5**	**5.1**	**4.5**	**4.1**	**4.1**
in % of Business Admin. or Econ.	**29.8**	**31.1**	**30.5**	**29.3**	**28.2**
Women [%]	19.0	18.8	18.8	19.1	19.2
Foreigners [%]	16.8	16.5	16.5	16.7	17.6

Note: Graduation consists of Licentiate, Bachelor or Master in Switzerland. All Percentages are rounded to one decimal place.
Source: Federal Administration of Switzerland.

Table 4.A.2: Capacity constraints at the university due to high amount of entering students.

Year	No. of Students	ASY students
1990	3908	582
...	...	...
2000	4701	843
2001	4938	971
2002	4917	953
2003	4852	900
2004	4569	789
2005	4508	954
2006	4915	1022

Figure 4.A.1: Time line: Institutional setup

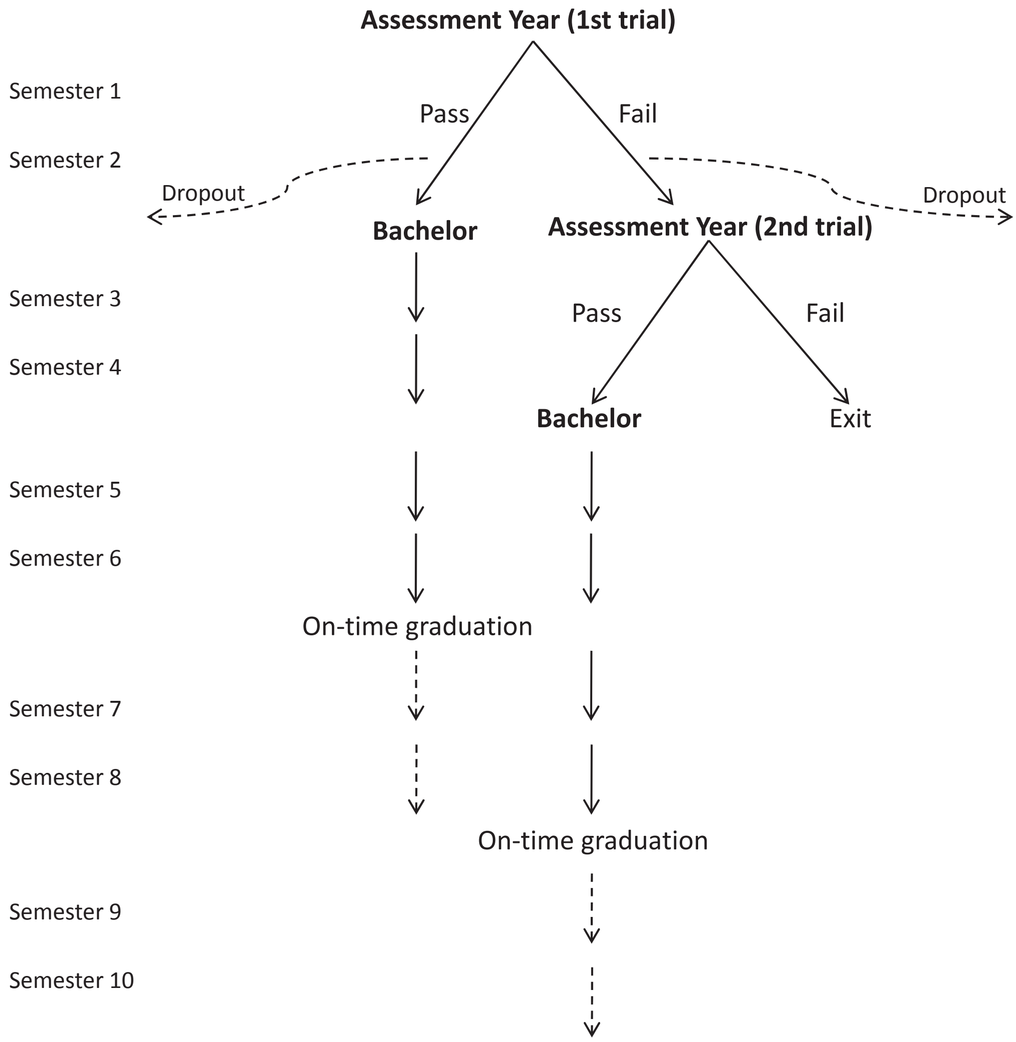


Table 4.A.3: Typical Curriculum of a Student

Calendar Week	38	45-46	46		50	51	3-7						10
1st Semester	Start	register to exams	Assignments			End	A Exams						get grades
Exam			LS		CT		Econ	BA	Law	Math			
Credits			3		2		5.5	5	5.5	3.5			

Calendar Week	8	15-16	50-15	16	16-21	21	25-29						35
2nd Semester	Start	register to exams	Assignments			End	B Exams						get grades
Exam			Essay	BA*	LS		Econ	BA	Law	Math	CT	FL	
Credits			5	2	3		5.5	5	5.5	3.5	2	4	

LS	Leadership skills
CT	Critical Thinking
BA	Business Administration
BA*	Business Administration: Case study
FL	Foreign Language
Assignments	Essays or oral assignments during the semester
Exams	Written exams during the exam period

Table 4.A.4: Descriptive statistics: Entering first-year students, by year

	# obs	2001 - 2006	2001	2002	2003	2004	2005	2006
Background Characteristics								
Male	3762	73%	75%	74%	75%	74%	72%	71%
Age $<$ 20	3762	12%	8%	9%	14%	14%	13%	17%
Age 20/21	3762	65%	64%	64%	65%	68%	66%	61%
Age $>$ 21	3762	23%	29%	26%	21%	18%	20%	22%
Foreign nationality	3762	22%	18%	24%	25%	26%	23%	20%
Entrance degree from SG	3762	15%	14%	14%	17%	17%	15%	16%
Entrancetest	3762	16%	13%	17%	21%	18%	14%	14%
Types								
1st sem: Not all exams	3762	6%	13%	4%	4%	4%	5%	3%
1st sem: MC $>$ 12	3762	10%	13%	7%	13%	13%	11%	11%
1st sem: Voluntary dropout	3762	1%	2%	0%	0%	1%	0%	0%
2nd sem: Not all exams	3762	2%	2%	2%	3%	2%	3%	1%
2nd sem: Accounting failed	3762	2%	1%	4%	5%	1%	1%	1%
2nd sem: All exams	3762	79%	77%	82%	75%	80%	79%	83%
Minus Credits								
MC $>$ 0 in first year	3762	63%	69%	65%	62%	59%	63%	58%
# of MC in first semester	3762	5.12	5.04	4.71	4.57	5.24	5.70	5.41
# of MC in first year	3762	8.43	8.02	8.59	6.91	8.87	10.58	7.67
Retention								
Fail: MC $>$ 12	3762	10%	10%	13%	8%	8%	13%	8%
Fail: Credits $<$ 60	3762	3%	7%	1%	4%	2%	2%	1%
Fail: Both reasons	3762	16%	16%	12%	16%	17%	18%	15%
Fail: Total	3762	29%	32%	27%	28%	28%	33%	24%
Repetition								
Repeater	3762	17%	15%	17%	16%	17%	21%	15%
Bachelor								
Bachelor started	3762	82%	75%	83%	83%	84%	82%	86%
Bachelor $\leq$ 4 semesters	3083	36%	47%	49%	38%	31%	25%	23%
Bachelor $\leq$ 5 semesters	3083	60%	68%	71%	62%	58%	54%	47%
Bachelor $\leq$ 6 semesters	3083	83%	87%	88%	86%	84%	80%	74%
# obs		3762	794	604	569	477	646	672

Note: The sample includes all first-year students with German mother tongue entering in 2001 - 2006 into the Business/Economics track. The last three rows (Bachelor $\leq$ 4/5/6 semesters) are only specified for students starting a Bachelor degree.

Part V

Labor Market Training and Occupational Mobility

Abstract

This paper evaluates the effectiveness of further vocational training (conducted between 2000 and 2004) on future labor market prospects of unemployed workers in Germany. It departs from existing papers in that it explicitly considers occupational mobility of workers throughout the analysis - i.e. (1) unemployed workers who intend to change their occupation are more likely to select into training and (2) training itself may induce occupation switches. We obtain each individuals intention to change occupation, determined at the beginning of unemployment, exploiting a unique feature of the available administrative data. The empirical analysis is conducted separately for unemployed with and without switching intentions. We estimate training effects for both sub-groups using matching methods and compare the respective outcome differences. Our results show that, overall, training appears to benefit its participants (at least in the long run). However, little is found with respect to effect heterogeneity across the groups of occupation switchers and stayers.

1 Introduction

Until today, training programs for unemployed workers are a major component of Active Labor Market Policy (ALMP) in many OECD countries. They intend to improve job perspectives of participants by maintaining or enhancing their human capital and, thus, reemployment chances. At the same time, training programs are costly - the Netherlands spend about 1.8% of GDP on ALMP measures in 2002, Sweden 1.4%, and Germany 1.2% (OECD (2004)) - motivating a large literature to analyze the effectiveness of such policy measures. As summarized by the comprehensive meta-analysis by Card et al. (2010) most training programs exhibit “modestly positive effects”. Nevertheless many of the respective studies provide mixed results so that it remains valuable for policy makers to improve the understanding about the mechanisms at work generating different effects. In particular, given scarce financial resources, knowledge about effect heterogeneities can support policy makers to better target the measures towards individuals most in need of training.

Following the discussion of Kambourov et al. (2012) we agree that the existing literature has largely overlooked a key characteristic of unemployed workers: occupational mobility. Traditionally, evaluation studies use econometric reweighting methods together with rich observational data to estimate the effects of program participation on individuals' labor market outcomes. Greatly simplified, they do so by comparing the post training outcomes of participants to those of 'quasi-identical' non-participants. When establishing evidence, however, evaluation studies usually neglect that participants and non-participants are often characterized by differences in occupational mobility where the former have higher rates than the latter. There are two explanations for this phenomenon which are, however, not mutually exclusive. First, it could be caused by poorer job perspectives and/or poorer occupational match quality for the group of participants. This is, for example, also suggested in the study by Wunsch and Lechner (2008) who show that a-priori employment prospects have an influence on program success. Labor market training, especially further vocational training (FVT), could support occupational mobility as it fosters the acquisition of skills not possessed so far, eventually qualifying the unemployed for alternative occupations. Thus, we can expect that self-selection into programs is (among other things) influenced by the individual's ex-ante desire to move into a new occupation. As such, differences in occupational mobility need to be accounted for explicitly when the outcomes of participants and non-participants are compared. Here, we account for it by analyzing program effect heterogeneity only and abstaining from explicitly investigating potential selection bias that could arise due to neglecting occupational mobility as a possible confounding factor. Second, training participation itself may induce occupational mobility, in particular for those who did not intend to change in the first place – and we provide some evidence that this is indeed the case. This could be due to learning in the course of the training program and should, thus, be rather seen as an outcome of training (not as a determinant). As such, enhancing the knowledge about the role of occupational mobility in the context of unemployment and labor market programs is valuable from a policy perspective as it can help to improve the allocation of training measures to participants who are expected to benefit the most. In addition, while analyzing the effectiveness of training measures, differences in occupational mobility should be kept in mind when interpreting the results.

To correct for differential mobility patterns between participants and non-participants Kambourov et al. (2012) use actual occupation changes (which are only observed after training participation and only for those who find a job) as a proxy for the ex-ante propensity of occupational mobility. Yet, as long as training induces occupation switching their strategy raises reasonable doubts.[48] In contrast to Kambourov et al. (2012), the analyses in this paper are based on administrative data which provides unique information about each individual's target occupation measured at the beginning of each new unemployment spell and, hence, before any decision about training participation is made. Using this feature we can determine the intention to switch occupation without conditioning on being employed and, thus, split occupation switching into an ex-ante (exogenous to training) and an ex-post (endogenous to training) part.

The purpose of this paper is to investigate mainly two things. First, we ask whether training programs are differently effective for unemployed who want to change their occupation compared to those who do not want to. This is relevant since any significant differences between the two groups would justify (ongoing) reforms towards a more intensive profiling of job seekers with a special focus capturing switching intentions of unemployed workers. Secondly, we analyze the effect of training participation on future patterns of occupational mobility. Another advantage of our data is the possibility to investigate a relatively long outcome horizon. Early studies as summarized in Heckman et al. (1999) or in Kluve and Schmidt (2002) could rarely detect any effects other than the well-known lock-in effect (Van Ours (2004)) due to short outcome horizons. Recent evaluations provide a more positive view on program impacts. They show that in the medium- and long-term training programs can increase the employability of participants implying that the post-treatment observation period should be sufficiently long to obtain convincing results.[49] To the best of our knowledge we are the first ones to explicitly exploit this feature of the data analyzing effect heterogeneity of government sponsored training programs with respect to occupational mobility. In addition, we consider medium-term impacts and investigate mobility patterns resulting from FVT.

For the analyses a large and informative administrative database for Germany is used. In particular, we investigate the effects of FVT, i.e. human capital enhancing programs with

[48]In fact, they argue on and test the conditional independence of training and occupation change. They claim, that once they control for factors influencing the ex-ante propensity to switch "the null hypothesis that training and occupational mobility are conditionally independent cannot be rejected."

[49]Recent studies supporting this view comprise e.g. Sianesi (2004); Lechner and Wunsch (2011); Lechner et al. (2011); Hotz et al. (2006); Fitzenberger and Speckesser (2006); Fitzenberger et al. (2008).

durations of 6 to 12 months for a sample of workers who entered unemployment between the years 2000-2004. Using microeconometric matching methods we estimate impacts on employment, successful occupation changes and earnings for up to 4 years after training assignment for those who intend to switch occupation and those who do not want to switch separately and compare the respective outcome differences across sub-groups.

The estimated effects show positive impacts of training on future unsubsidized employment for those who intend to switch occupation as well as for those who do not intend to do so. We observe higher employment probabilities from about 3 years after program start onwards, whereas over the entire 4 years program participation does not result in more months in employment than no participation. With respect to successful occupation switching we find that those who intend to move to a new occupation are more likely to be employed in a different occupation then their last one due to training participation. This effect of training is even more pronounced for participants without initial switching intention as they are about 3 months more employed in a new occupation than without training. With respect to heterogeneity, however, we do not find much evidence for differential effects across the two sub groups. Our findings show rather small differences in employment probabilities over time as well as in accumulated months in unsubsidized employment over the 4 year observation period. Thus, we would conclude that unemployed most in need of human capital adjustments, i.e. individuals who intend to change their occupation, do not benefit more from training than unemployed workers who look for a job in their old occupation. These conclusions and the interpretations of our findings are, however, limited as the differences cannot be estimated very precisely.

The remainder of the paper is structured as follows. Section 2 provides additional theoretical arguments about the relation between (intended) occupation switches, associated human capital loss and the role of publicly sponsored training programs. Details on FVT programs in Germany and their assignment process are given in section 3. Subsequently, in section 4, we describe the data and how we measure the intention to change occupation, before section 5 explains our identification strategy and the estimation procedure. In section 6 we present our results, and provide concluding remarks in section 7.

2 Occupation switching, human capital and training

Occupation Switching and loss of Human Capital

Since Becker (1962) it is widely accepted, that productive human capital has an accumulative component to it. Workers usually acquire and improve their knowledge and skills on a given job with tenure. It is an empirical fact that with longer duration spent on a job –ceteris paribus- workers are observed to have higher wages and positive path dependence in their job duration. This effect is also often labeled “experience”. It is further argued that workplace related human capital has mainly three major components: i.e. industry specific, firm specific and occupation specific human capital. The relative importance of each of these components is subject to intense debate.

The more recent literature ranges from studies of Jacobson et al. (1993) and Kletzer (1998) that attribute human capital enhancements mostly to firm tenure; to Neal (1995) and Parent (2000) who discuss the importance of industry specific human capital; toKambourov and Manovskii (2009) who set firm and industry human capital in contrast to occupation specific human capital. Based on samples of laid-off workers these studies usually use variation in industry, firm and occupation tenure at the workers level to estimate the relative importance of each component in wage determination. Using data from the US (PSID) Kambourov and Manovskii (2009) show that occupation change is the main channel through which unemployment spells affect wage losses of workers, i.e. particularly occupation specific human capital matters. The negative effects of occupation changes are most prevalent in the short-run and tend to fade out when longer outcome horizons are considered as workers tend to close their skill deficits over time. Moreover, there is consensus that involuntary unemployment in rigid labor markets is likely to increase the probability of occupation switching which, in turn, might result in a larger amount of underutilized human capital.

The importance of occupation tenure is also discussed in a more general framework by Lazear (2009) who models the job match of workers as a function of their specific skill portfolio.[50] Hence, while the firm and industry dimensions clearly matter, workers´ productivity levels are apparently most bound to occupation specific human capital. This is also supported by the empirical study of Gathmann and Schönberg (2010) who find that workers who move between occupations with relatively dissimilar skill requirements are on average followed by larger wage drops. They further illustrate that the likelihood of such moves is lower for workers with longer times spent in a certain occupation.

In this study we follow the related literature in two major respects: first, while employed in one occupation workers accumulate occupation specific human capital over time so that it can be seen, among others, as one key determinant of labor market outcomes. Second, in-voluntary unemployment has a detrimental effect on worker´s future labor market outcomes – more so if part of the accumulated occupation specific human capital is lost due to occupation change (Lazear (2009)). Thus, we expect especially workers with skills that are less demanded by the labor market to face worse future outcomes when job loss occurs, ceteris paribus, as their likelihood of finding a job is comparatively lower.

[50] The skill portfolio in Lazear (2009) is formulated in a very general framework using the relative shares (or weights) of two-different skills which are each possessed by every worker, yet with varying ratios.

Occupation Switch and Vocational Training

Building upon human capital improvements active labor market policies are expected to reintegrate unemployed workers into employment. Their effectiveness is at the core of numerous empirical studies (Card et al. (2010)). In particular, government-sponsored vocational training programs aim at adjusting the skills of their participants towards the demands of employers. Given their expected low marginal productivity training is, thus, especially relevant for unemployed individuals who would - given their skill portfolio and in absence of training - face rather unfavorable labor market perspectives. Traditional industries (e.g. textile industry) are obvious examples where the introduction of new technologies and cheap imports often obviates the need for the respective skills and, hence, for certain types of workers. Other examples are occupations like carpenter or shoemaker. Here, too, the replacement of routine labor by machinery has been documented in the literature (Kambourov and Manovskii (2009)). Structural changes in these labor market segments often lead to crowding out effects with respect to the affiliated occupations. It follows that skill adjustments can be very valuable for workers who are affected by those structural changes as they might enable them to be more productive in occupations other than their original one. On the contrary, one might expect training programs to be relatively less effective in cases where unemployed workers have good chances of successful reintegration anyway (i.e. skill enhancements/adjustments should be of minor importance for workers whose skills are, in general, heavily demanded by the labor market). Vocational training programs should, thus, support, (or induce) occupational mobility. Under this assumption, we would expect to see a larger share of occupational switchers in the pool of training participants compared to non-participants. In fact, the recent empirical study by Kambourov et al. (2012) confirms this pattern for US data. The authors argue that the intention of occupation switch bears the incentive to participate in vocational training.

Whether vocational training is more effective for workers who are occupational switchers compared to occupational stayers is ultimately a question that can only be answered by empirical evidence.At the same time, however, the inter-linkage between occupational mobility and the effectiveness of vocational training could also be investigated with respect to selection bias. So far most program evaluation studies have largely neglected the issue of occupational mobility in that context.[51] This bears the implicit assumption that occupation switching is entirely endogenous to training (not necessarily if it were measured prior to training) and should be seen as a subsequent outcome of training participation rather than a confounding factor. However, the study of Kambourov et al. (2012) raises reasonable

[51] Empirical studies that use the sample of non-participants as a (reweighted) control group usually do not account for differences in the propensity of occupation switch. Notable exceptions are the studies by Lechner and Wunsch (2009); Lechner et al. (2011) who control for previous occupations, as well as job looked for. However, only broad occupational categories are controlled for which, we assume, do not fully account for potential human capital losses that result from occupation switching.

doubts about the endogeneity of occupation switch. If the decision to change occupation is made before any training participation is considered, estimates of the impact of training on labor market outcomes might be biased if the propensity of occupation change is not adequately taken into account. The estimated effect of training would then comprise both, the effect of human capital improvements as well as the effect of occupation change. Given that the latter is expected to be negative, positive training effects might be disguised if occupation switching is not controlled for. Accordingly, an obvious strategy would be to explicitly investigate this potential selection bias and relate it to the findings of the existing literature. However, the purpose of this paper is not to quantify this potential bias and obtain 'improved' estimates of training effects. Instead, we focus on effect heterogeneity which in turn allows policy advice with respect to the program allocation procedure. In other words, we do not investigate how effective FVT actually is for their participants but whether it is differently effective related to occupational mobility. Thus, the above discussion is included to provide a rather comprehensive picture of the relation between occupational mobility and training, and to highlight that there are at least two possible dimensions and ways to analyze it.

Kambourov et al. (2012) attempt to address the issue of occupational mobility when they provide estimates of vocational training programs in the US. Their strategy is such that they control for occupation switches that are only observed ex-post, i.e. after training participation. Contrary to the existing literature, this approach implicitly assumes that although observed after training, occupation switches are conditionally independent of training. This is, however, problematic if occupation switching is (partly) determined by training participation. In fact, this approach is problematic if employment is influenced by training participation as the observability of occupation switching depends highly on employment. If this is the case, controlling for (or splitting the sample on the basis of) ex-post switching is an invalid strategy as it basically means controlling for part of the outcome.

We acknowledge this possibility and, thus, deviate from the existing literature in the way we think about occupational mobility of unemployed workers. We assume there are two components to the propensity of occupation switching – an exogenous part and an endogenous part. The first (exogenous) part can be thought of as being determined before any decision about training participation is made. We call this the ex-ante intention to switch. It reflects the workers propensity to start a job in an occupation other than the last occupation in the absence of training. This ex-ante decision results from the latent desire of the unemployed as well as the counseling of the caseworker in first consultations. We claim that this is only related to the potential employment prospects in the old occupation.[52] We

[52]Note that we cannot make any general statements about the relative importance of caseworkers in such decisions. Likewise, this paper will remain silent about normative questions like whether the caseworker should (or should not) convince the unemployed to change her occupation. Yet, what is crucial for our later identification strategy to be valid is that caseworkers do not influence the occupation switching decision of the unemployed because of training assignment.

will develop an empirical measure for this exogenous part and use it as a sample-splitting criterion in the analysis. Note that we do not intend to analyze potential selection bias in training effects by including this variable as an additional covariate in the estimation. Instead, we divide our sample based on the intention to switch variable and thus only indirectly account for selection bias. Our approach - assuming an exogenous component in the propensity to switch occupation - allows measuring the causal effects of training for each sub-group separately. The second (endogenous) part is the one that is directly influenced by training participation and is, thus, rather to be seen (and treated) as an outcome. This basically refers to the effect of training on occupational mobility. In fact, participating in training could encourage or discourage individuals to change their occupation, thereby updating their initial (exogenous) switching intention. For analyzing these effects of training one could investigate the change in intended occupational mobility as well as the actual change in occupation. While the impact on intended mobility might be an interesting policy parameter its measurement is problematic. Since we do not observe a second value for the target occupation after training it is simply not possible to analyze the effect on the intention to switch. Thus, we will only investigate the impact of training on realized occupational mobility. Note that this outcome reflects actual occupation switching merely in a limited way, as it is observed only for individuals who are employed after (non-)training. Yet, at least for individuals who find a job within our 4-year outcome horizon we are able to measure the endogenous part defining an outcome measure that combines employment and occupational mobility. Although this is a selected group we believe to learn something about the relation between FVT and occupational mobility.

3 Further Vocational Training Programs in Germany (2000-2004)

ALMP has a rather long tradition in Germany and comprises various measures that intend to (re-)integrate unemployed individuals in the labor market.[53] In this analysis we focus on human capital enhancing programs instead of measures that provide job search assistance, improve job finding abilities, or support minor improvements of skills. Government sponsored investments in human capital for unemployed are particularly important in the context of occupation switching and one of the main instruments to alleviate specific skill deficits. These programs are intended to adjust and extend knowledge, or qualify individu-

[53] An overview of the different instruments can be found in several evaluation studies one of which is e.g.Wunsch and Lechner (2008). There exist training measures that combine job-seeker assessment, and minor adjustment of skills in programs of up to 2 months or programs promoting the “acquisition of specific knowledge and skills” with a similar planned duration. Moreover, there are employment programs promoting "subsidized non-market jobs" and the general category of "further vocational training" with planned durations of 3 up to 24 months.

Table 5.1: Further vocational training and other instruments of ALMP in Germany 2000-2004

	2000	2001	2002	2003	2004
	Entries in 1000				
Further vocational training (FVT)	552	450	456	246	185
Training measures (TM)	477	565	877	1070	1188
Employment program (EP)	361	386	229	180	165
Wage subsidy program	206	193	225	202	182
Short-time work (average stock)	86	123	207	195	151
	Expenditures in million EUR				
Total expenditures on ALMP	22005	22317	22400	21197	19518
Further vocational training (FVT)	6808	6982	6701	5000	3616
Training measures	323	350	478	578	496
Employment program	5044	3847	3143	2273	1610
Wage subsidy program	1159	1191	1351	1529	1085

Source: BA (2000)

als for a different occupation such that the risk of staying in unemployment is substantially reduced.[54] This is particularly important for those unemployed who intend to move to a new occupation as their employment chances are potentially lower (relative to unemployed from that particular new occupation) due to lack of experience and human capital.

In Germany, programs that are supposed to improve human capital are summarized in the category further vocational training (FVT). As 5.1 shows, this type of policy measures is widely used and one of the most important instruments in ALMP in Germany. Relative to other programs FVT is the program with the most or second highest number of participants, even though the entries decline considerably towards the end of our observation period. Expenditures on FVT make up the largest share of total ALMP expenditures throughout, compared to other programs such as training measures (TM) or employment programs (EP). This is due to their longer durations of up to two years and the payment of special benefits (so-called maintenance allowance, MA) to unemployed during program participation.

Our analysis considers two specific programs within the group of FVT: (i) occupation-related and general training, and (ii) practice training in specific skills and practice firms. The first type provides participants with knowledge in specific skills necessary for a certain occupation (e.g. use of a particular software for an engineer, or learning special techniques as a carpenter) or they improve their general qualifications (e.g. computer skills, language skills). Typically these trainings are offered as classroom trainings and, thus, contain a large theoretical component. Instead, the second program type is expected to improve participants' skills within a more practical environment such as practice firms or even internships. Individuals educated as businessmen or merchants perform real tasks, however, in a simulated entrepreneurial situation and they learn how different areas of a firm interact. Carpenters, electricians or other more technical occupations are e.g. trained in practice

[54]These objectives are explicitly formulated also in the Social Code III §87, providing clear guidelines about the aim of the programs the employment agency and the caseworker, respectively, offer.

workshops at machines and workbenches with the help of instructors. Both program types have planned durations between 6 to 12 months, whereas the first type is most widely used from all FVT programs. Given very similar average durations of the two program types and their mutual objective of improving human capital we interpret them as one *treatment*.

Unemployed are eligible for these trainings once they have been working for a total of 12 months in the last 3 years before participation if: (i) the training is necessary to prevent (further) unemployment and, (ii) the caseworker approves the particular course after consultation with the unemployed.[55] In the first consultation the caseworker collects information on qualifications and skill deficits, the regional mobility of the unemployed, job search activities as well as health impairments, family background, and the target occupation. On the basis of this profiling, together with available training resources and regional labor market characteristics, the caseworker has to decide on program assignment. Usually, however, the decision about training is not necessarily made during the first meeting. The time of program participation does not reduce the unemployment benefit (UB) claim, i.e. every day of training prolongs the unemployment benefit period. In addition, instead of unemployment benefits participants receive maintenance allowance usually in the same amount as the previous unemployment benefits. Accordingly, unemployed might have an incentive to ask for a program in order to extend their unemployment benefit period.

Unemployed are eligible for these trainings once they had been working for a total of 12 months in the last 3 years before participation if: (i) the training is necessary to prevent (further) unemployment and, (ii) the caseworker approves the particular course after consultation with the unemployed.[56] In the first consultation the caseworker collects information on qualifications and skill deficits, the regional mobility of the unemployed, job search activities as well as health impairments, family background, and, importantly, the target occupation. On the basis of this profiling, together with available training resources and regional labor market characteristics, the caseworker has to decide on program assignment. Usually, however, the decision about training is not made during the first meeting. The time of program participation does not reduce the unemployment benefit (UB) claim, i.e. every day of training prolongs the unemployment benefit period. In addition, instead of unemployment benefits participants receive maintenance allowance usually in the same amount as the previous unemployment benefits. Accordingly, unemployed might have an incentive to ask for a program in order to extend their unemployment benefit period.

Note, however, that in 2003 there were some changes regarding the reduction of the UB claim and the assignment process. Program durations do not extend the claim period one by one anymore, but for two days of program participation the claim is now reduced by one day. The allocation of programs was altered such that unemployed now receive a voucher (so-called Bildungsgutschein) for a specific training course and they can choose

[55]These rules are explicitly stated in the Social Code III §77.

[56]These rules are explicitly stated in the Social Code III §77.

their training facility freely. Before, this choice was made by the caseworker. Even with this voucher, caseworker and unemployed have to formulate a specific educational goal and the caseworker has to approve the voucher on the basis of her assessment of the potential of successful program completion and subsequently increased employment chances. Schneider et al. (2006) show that these regulation changes alter the composition of the group of participants slightly.[57] The introduction of the voucher indirectly favors unemployed that perceive the free choice of training facilities as an advantage. Relatively better educated, younger and more mobile individuals now more often participate in programs, whereas less qualified, handicapped people or long-term unemployed are less likely to take part in training courses. However, given our highly informative data on individual characteristics (including educational attainment, age, worker mobility, or health impairments) and employment histories the described changes are unlikely to invalidate our identification strategy (as discussed later).

Eventually, caseworkers need to assess the program success and the subsequent employment chances carefully based on the individual's profile in order not to 'waste' financial resources. Since the target occupation is, however, only one characteristic in the profile and not of main importance in the program assignment process (program availability, financial constraints, skill deficits, or mobility considerations are more crucial) it seems unlikely that caseworkers systematically suggest target occupations to unemployed according to available program places. In other words, caseworkers suggest training courses based on the agreed target occupation and not vice versa. If caseworkers proposed occupations because of available program places our identification strategy would be invalid as the choice of target occupation and training program were jointly determined. We instead require exogeneity of the target occupation decision with respect to training. We believe this is plausible as caseworkers do not suggest target occupations (potentially unrelated to the qualifications of the unemployed) because there are places available in a program that is related to the proposed target occupation. This is also implausible as the first consultation rarely results in program assignment and program participation is based on qualification needs and not vice versa. Nevertheless, whether FVT programs are eventually more (or less) effective for unemployed workers who intend to move into a new occupation is a very relevant question from a policy point of view and should, thus, be considered carefully.

[57] According to Schneider et al. (2006) the changes in the composition of participants, though, are dominated by alterations in the structure of the allocation of programs types. In general, the entries in all FVT measures declined after the reform. The largest drop occurs in the two types we consider, whereas programs resulting in a recognized vocational degree are reduced relatively less.

4 Data, Sample Selection and the Definition of Occupation Switch

4.1 General Database

For our analysis we start off with a 2 percent random sample of German employees that are subject to social insurance contribution at least once in the years 1990-2008. The data is maintained by the Institute for Employment Research (IAB) and is collected from the employers' administrative process of registering (new) workers with the agency. Eventually, the data is based on several sources: social insurance notifications, program participation records, benefit payment files, and job seeker registers. Those are combined to one administrative database - the so-called Integrated Employment Biographies (IEB).[58] In total, our general sample contains about 1.4 million distinct individuals from which we select the observations relevant for our analysis.

This administrative data includes information on the target occupation of each unemployed which is one crucial variable for our investigation. Within their first meetings the caseworker and the unemployed agree on this target occupation. Contrasted with the last occupation we then compute the 'intention to switch occupation' for each individual (more details in section 4.3). The data further allows us to obtain exact durations (to the day) of e.g. employment and unemployment spells, program participation periods, or days of unemployment benefit claims. Since the duration of unemployment benefit claims as well as the amount of unemployment benefit payments are likely to affect individual incentives of program participation and job search activities we will use these variables in the analysis. We further benefit from the availability of various socio-demographic characteristics (such as age, schooling, higher education, marital status, mobility, health impairments, or last occupation) as well as the possibility to compute complete labor market histories (for at least 10 years prior to unemployment) for each individual. There is also information about past and future employer(s) for each individual. In addition, we augment the data with official data that capture the regional economic situation and the specific local labor market for each individual. One interesting variable here is an occupation specific measure of monthly labor market tightness defined as the ratio between the number of vacancies and unemployed at any given month.[59] Theory suggests a clear link between that measure and occupational mobility. Hence, we can approximate (and control for) occupational labor supply and demand which might influence the intention to switch occupation as well as the training participation decision.

[58] Further details on the data are available in Oberschachtsiek et al. (2009).

[59] We obtain the monthly number of vacancies on a two digit occupation level from the job openings registered with the employment agency.

4.2 Sample Definition and Program Participation

From the general data described above, we select all individuals that enter unemployment from unsubsidized employment or general non-employment (e.g. out of labor force) between April 2000 and December 2004.[60] The sample starts in April 2000 because of missing information and likely underreporting of job-seeker spells in the first months of 2000. It is restricted to the end of 2004, on the one hand, so that we obtain a 4 year outcome horizon which is preferable for detecting potential training effects of longer programs as argued by Lechner et al. (2011). On the other hand, following the well-known Hartz reforms in 2005, the administrative processes as well as the incentives for program participation changed dramatically with the introduction of the so-called "unemployment benefit II"[61] and the abolishment of social assistance payments.

The main additional restrictions we impose on the sample are related to past employment and past occupation switching. Moving to a new occupation is more strongly associated with losses of occupation specific human capital for individuals who are employed for a relatively long time in a certain occupation before entering unemployment. Individuals with this kind of labor market attachment have relatively lower employment chances if they change their occupation. In contrast unemployed with low previous labor market attachment and/or excessive previous occupation switching are not of interest in our study because they have accumulated less occupation specific human capital in the past. While individuals who frequently change jobs or occupations are clearly also in need for training, they are arguably relatively less prone to occupation specific human capital loss as a result of unemployment. In contrast, unemployed with high labor market attachment who have to change their occupation, maybe due to structural changes in the economy, are thus relatively more in need of FVT. Following these arguments, we impose the requirement that individuals in our sample were in unsubsidized employment for at least 9 months within one year before entering unemployment and had at most 2 different occupations (defined on a 2 digit level) in the last 3 years before entering unemployment. Furthermore we require them to have accumulated at least 12 months of work experience in the same occupation during the last 3 years before unemployment. This definition is quite restrictive and excludes about 65% of our initial sample reducing it from 351,000 observations to about 121,000.

Next, we only consider individuals aged 20 to 55 to avoid influences from schooling and retirement decisions. This affects about 15% of the remaining sample (5% are aged below

[60] We do not consider cases from Berlin due to its special situation as capital and formerly divided city which makes it rather incomparable to the whole of Germany. Furthermore note that multiple entries into unemployment are possible. However, given the conditions about previous employment and occupations, more than 90% of individuals in our final sample have only one unemployment spell.

[61] "Unemployment benefit II" combines unemployment assistance and social assistance payments (existent until the end of 2004) to one general welfare payment. This is available for all unemployed that are not eligible for unemployment benefits (anymore) fixed at a level that should ensure a minimum subsistence level.

20). Concerns about influences from vocational training or university education are, if any, only minor as we further exclude unemployed whose last employment was an apprenticeship and those who worked in marginal employment[62] (12.5% of the so far restricted sample). The latter is observed relatively more often for younger individuals and is also often used by students. Thus, the remaining individuals aged between 20 and 25 obtain their occupation specific human capital also from rather stable employment in a regular job. We also exclude some specific last occupations such as agricultural jobs or those classified as unskilled workers as they are not representative for the unemployed of interest in our study. Unemployed with agricultural jobs are mostly seasonal workers and, thus, unlikely to be considered for training programs.[63] The category 'unskilled workers' has simply no informational value as it does not describe any occupation. This reduces the sample by another 7.5% leaving us with about 83,000 observations before the eligibility condition, as described below, is imposed resulting in a final sample size of 33,414 observations.

In order to ensure program eligibility of all sample individuals we select only those unemployed who receive unemployment benefit or unemployment assistance just before program start.[64] This, however, requires knowing start dates also for non-participants which are not available by construction. We tackle this issue by simulating hypothetical start dates for non-participants following the approach in Lechner (1999). In particular, we assign start dates randomly to non-participants drawing from the distribution of actual program starts of participants regardless of their switching intention.[65] Given our empirical approach to identify the parameters of interest we need all relevant factors that jointly influence program participation, the intention to switch and labor market outcomes to be exogenous (besides being observable in the first place) with respect to program participation. This is obtained by measuring all variables at or relatively to unemployment start (except time until treatment).

Eventually program participation is defined on the basis of the start dates within each unemployment spell following the approach discussed in Lechner et al. (2011). Each unemployed individual who participates at least once in a vocational training within 12 months after entering unemployment is classified as participant whereas we consider only the first

[62] Marginal employment in Germany refers to jobs that are not subject to social insurance contributions and must not pay more than 400€ per month, or are temporarily limited to 2 months.

[63] Agricultural workers are typically subject to seasonal unemployment and improving their skills with training programs is not expected to prevent them from future unemployment once the next season is over.

[64] Actually, the main eligibility condition requires individuals to be employed for at least 12 months during the last 3 years before program start. This however, is the same condition as for receiving unemployment benefits or unemployment assistance. In addition to caseworker consultation, individuals are further eligible if 'only' participation in FVT can increase employment prospects and avoid (further) unemployment. Finally, the regulations require a vocational degree or at least three years of work experience for eligibility. However, since we select only individuals with valid information on their last occupation and require at least one year of employment in the past, the participants and non-participants are most likely to be eligible.

[65] In fact, program participants that intend to switch occupation and non-switchers start the program on average after 4.5 months in unemployment. There is no observable difference in the participants' distribution of program starts between switchers and non-switchers.

program participation.[66] Everybody who does not enter in vocational training or any other labor market program within this one-year-window is considered as non-participant. This includes also individuals who enter training, but only later than 12 months after entering unemployment. Finally we require training participation to start before December 2004 in order to obtain a 4-year outcome period which is preferable for the evaluation of rather long lasting human capital programs (about 6-12 months).[67] Following Fredriksson and Johansson (2003) and their discussion of the problems of this kind of participation definition we check the sensitivity of our results by extending the participation window to 18 months. A reduction of the window to 6 months is however infeasible as the number of program participants would become too small.

4.3 Definition of the Intention to Switch Occupation and Sample Descriptives

One crucial variable in our analysis relates to the intention to switch occupation for each individual. We derive this as a binary indicator from the pair of the last and the target occupation - variables available for each unemployed in our data. The target occupation represents the realizable placement desire of the unemployed and is formulated within the first consultation with the caseworker. In the first meeting the caseworker collects lots of individual information about qualifications, degrees, family background, and mobility of the unemployed as well as the last occupation. Subsequently, based on this information and the vacancy listings from the employment agency, caseworker and unemployed formulate a placement objective that also determines the target occupation. In later meetings the caseworker and the unemployed talk about specific training possibilities. This is the usual curriculum. Given this, the target occupation is fixed before any training participation is considered. Thus, the caseworker is unlikely to systematically suggest (other) target occupations according to available training resources.[68] This is an important point for the validity of our analysis and should be kept in mind.

The target occupation is originally available as a 3-digit occupation code in the data based on the Occupation Classification of the German Statistical Office (Destatis (1992)). However, we use a more aggregated measure based on 2-digits. While this still leaves enough

[66] It is theoretically possible, that individuals enter a second vocational training program within the 12-months-window. However, analyzing multiple program participations would require a dynamic evaluation approach. Instead, we use all states after the first program start as outcomes

[67] Based on this participation definition, all individuals participating in any program after 12 months are also classified as non-participants and thus are used as controls. Those cases account for only about 19% of all non-participants in our sample, which does not substantially weaken our treatment interpretation - "starting vocational training within 12 months of unemployment" vs "no program in 12 months". See Wunsch and Lechner (2008) for a further discussion and the interpretation of estimates based on this definition of treatment.

[68] These explanations are based on conclusions and descriptions from WZB and Infas (2005), Rübner and Sprengard (2011) and training material for caseworkers.

variation, we believe to reduce potential problems related to measurement error when using the finer classification. In addition, we observe the last occupation of each individual as well as the future occupation (i.e. only for those who find a job within our observation period). Accordingly, individuals whose target occupation differs from their last occupation (on 2-digit level) could be classified as "unemployed with an intention to switch occupation".[69] For our analysis, however, we need to assume a loss of human capital to be associated with occupational mobility. Hence, the simple difference in occupation codes is overly simplistic for our purposes as it does not allow any inference about the extent of human capital loss. Different occupation codes, even on a 2-digit level, do not always imply significantly different skill requirements in the new occupation and the two occupations can be rather close to each other in terms of their human capital needs.[70]

Going beyond simple inequality in occupational codes we define the intention to switch based on the skill distance measure proposed in Gathmann and Schönberg (2010). They use data from the German Qualification and Career Survey which includes information on skill requirements for each occupation on a 2-digit level to obtain specific values for the distances between different occupations. In short, their measure is based on the 19-dimensional Euclidian distance between occupation specific skill vectors. The normalized distance measure is defined between 0 and 1 with larger distances reflecting more different skill requirements in pairwise comparisons of occupations (see Robinson (2010) for a profound discussion on skill-distance measures). We simply use the measure they compute based on the data from the German Qualification and Career Survey[71] (2003 wave) and match these skill-based occupation distances to each last-occupation – target-occupation pair. We define switching on the distribution of distances of all unemployed whose target occupation is different from their last occupation before unemployment. By definition, the resulting distances for individuals with identical previous and target occupations are zero. Every distance larger than the median of distances of the sub-group of individuals, whose target occupation is different from the last occupation, i.e. skill distance > 0.06, is considered as "intention to switch". To see why, consider the resulting distribution of occupational distances for the unemployed in our data as depicted by Figure 5.A.1 in the Appendix. It shows that most target occupations "are not far away" from the original occupations in terms of their skill requirements. We do not want to consider these as occupation switches as we expect human capital loss for such moves to be negligible. This means that individuals who indicate a

[69]See Kambourov and Manovskii (2009) and Robinson (2010) for extensive discussions on possible definitions of occupation switching. Note, however, that most of the critiques are less relevant for the purpose of this paper, as we use administrative data which is less prone to measurement error.

[70]Some individuals, for instance, want to change from electrician to technical service personnel (occupation codes 31 and 62, respectively) or from mechanics to locksmith (28 to 27, respectively). It is not obvious, that those individuals change to an occupation that requires skill adjustments. Furthermore, the distance in the ordinal numbers of the occupation code does not provide any information on this issue.

[71]This survey is conducted jointly by the German Federal Institute for Vocational Education and Training (BIBB) and the Institute of Employment Research (IAB) and its main purpose is to track skill requirements of occupations.

target occupation that is relatively close to their original occupation (i.e. with a skill distance <0.06) are treated as stayers. Our definition of occupation switching is therefore more conservative than the alternative which is only based on inequality between the previous and the target occupation code of the unemployed. In a sensitivity analysis we reduce the skill distance threshold by 10% defining "intention to switch" if skill distance is larger than 0.054. Investigating the sensitivity of results also for a 10% increase of the chosen threshold would be ideal. The data, however, does not provide enough observations for this additional step.

In Table 5.2 we document the most frequent categories of intended occupation switches. This overview shows that there are basically two different occupation switches. Firstly, we observe switches that can be classified as skill upgrades such as the intended change from unskilled construction worker to bricklayer (n=157) or plasterer (n=59), or from office clerk to engineer (n=56) or technical service personnel (n=34), or from cleaning service worker to sales personnel (n=50). Secondly, there are several switches only between different occupation types, for example from sales personnel to accountant (n=86), or from cleaning service worker or cook to housekeeper (n=52 and n=33, respectively), or from office clerk to storekeeper (n=39). In our analysis we consider these two categories together because both groups of intended occupation switches require (considerable) skill adjustments which could be obtained through FVT. Note that our skill distance measure does not allow discriminating "occupation downgrades" (e.g. bricklayer to unskilled construction worker (n=53)). However, as the table indicates, such cases are rather rare and do not account for a large share of program participants. Finally, given the 68 occupation categories there are many possible changes which explain the low share of each specific switch. The share of participants in each intended switch (in Table 5.2) varies considerably from 25% (office clerk to engineer) to 3.4% (unskilled construction worker to plasterer).

Additional insight into the switching definition from a more aggregate perspective is provided in Table 5.3. It gives an overview of the shares within the groups of program participants and non-participants in our sample and the numbers and shares of intended and actual occupation switches.[72] Accordingly, about 67% of non-participants and 73% of participants are in unsubsidized employment after the end of unemployment. Note that this measure is not used as an outcome in our actual evaluation as it is not measured relative to program start. If anything this variable shows that 33% and 27%, respectively, never find unsubsidized employment within 4 years after program start. Instead, we use the first state after unemployment only for illustrative purposes and for obtaining the first occupation after unemployment for each individual. Conditional on being in unsubsidized employment we can see from the table that unemployed who intend to change their occupation are also more likely to actually switch relative to those who do not want to switch. 57% of

[72] Actual switches are obtained from the actual future occupation given employment based on the same occupation distance threshold as for intended switching.

Table 5.2: Most frequent intended occupation switches in the sample

Last occupation	Target occupation	cases	(%) of all switchers	share participants (%)
Unskilled Construction Worker	Bricklayer, Mason	157	2.43	3.82
Sales Personnel	Accountant, Bookkeeper	86	1.33	12.79
Accountant, Bookkeeper	Sales Personnel	79	1.22	15.19
Unskilled Construction Worker	Plasterer	59	0.91	3.39
Office Clerk	Engineer	56	0.87	25
Bricklayer, Mason	Unskilled Construction Worker	53	0.82	3.77
Waiter, Barkeeper, Innkeeper	Sales Personnel	52	0.81	7.69
Cleaning Service Worker	Janitor, Housekeeper	52	0.81	1.92
Cleaning Service Worker	Sales Personnel	50	0.77	6
Storekeeper, Warehouse Keeper	Sales Personnel	49	0.76	20.41
Assembler	Product/ Quality Inspector	45	0.7	6.67
Storekeeper, Warehouse Keeper	Office Clerk	42	0.65	14.29
Technical Service Personnel	Office Clerk	40	0.62	20
Office Clerk	Storekeeper, Warehouse Keeper	39	0.6	7.69
Janitor, Housekeeper	Cleaning Service Worker	36	0.56	11.11
Office Clerk	Technical Service Personnel	34	0.53	11.76
Nurse, Dietitian, Physical Therapist	Office Clerk	34	0.53	8.82
Cook	Janitor, Housekeeper	33	0.51	6.06
Sales Personnel	Storekeeper, Warehouse Keeper	32	0.5	15.63

Note: The total number of those with an intention to switch occupation in the sample according to our distance measure and the respective threshold is 6455.

Table 5.3: Intention to switch, actual switch, and program participation

		Non-participants			Participants		
		Intention to switch					
		No	Yes	Total	No	Yes	Total
Not employed	cases	7,765	2,108	9,873	598	137	735
	column %	31.49	35.4	32.25	26.01	27.35	26.25
Employed[a]	cases	16,895	3,846	20,741	1,701	364	2,065
	column %	68.51	64.6	67.75	73.99	72.65	73.75
Basis: Employed							
No actual switch	cases	12,512	1,643	14,155	1,179	102	1,281
	column %	74.06	42.72	68.25	69.31	28.02	62.03
Actual switch[b]	cases	4,383	2,203	6,586	522	262	784
	column %	25.94	57.28	31.75	30.69	71.98	37.97
Basis: Actual Switch							
Desired occupation[c]	cases		931			126	
	column %		42.26			48.09	

[a] The employment rate is obtained from the state after the end of unemployment and represents unsubsidized employment.
[b] Actual occupation switching is obtained from the first occupation after the end of unemployment. If the distance between the last occupation and the first occupation is larger than 0.06 the change is classified as switch.
[c] Individuals switch to their desired occupation if the skill distance between their intended occupation and their first occupation is less than 0.06. It is obtained as the share of those that intend to switch and actually do so.

non-participants (72% of participants) with a switching intention do switch whereas only 26% of non-participants (31% of participants) who do not want to change their occupation eventually switch. In addition we also observe relatively more participants with a switching intention to change occupation in the end compared to non-participants (72% versus 57%). Moreover, we compute the share of individuals ending up in their desired occupation of those who intend to and actually do switch. For non-participants these are about 42% and for program participants the share is slightly higher with 48%. Thus, we believe that there is a clear link between what we consider as the intention of occupation switch and realized occupation switches.

We further investigate the correlates of the a-priori intention to switch measure. Simply splitting the sample according to our binary variable Table 5.A.1 (Appendix) shows the mean values of selected covariates by sub-groups. The numbers suggest that individuals with switching intentions are on average more often males, married, less educated, unskilled workers with worse labor market histories. In addition, the average number of distinct occupations in the past is higher for individuals with switching intentions. This is line with the lower average value of the labor market tightness measure.

To check the robustness of these findings we run a simple probit regression. The predicted changes in the average probability of the intention to switch measure with respect to selected covariates (dY/dX) are provided by the coefficients in Table 5.4. Here, too, females are significantly less likely to indicate the desire to switch their occupation. Somewhat surprisingly, the same is true for workers younger than 30 – yet the profiles flatten for older age brackets. Also the presence of younger kids in the household shows a negative correlation. Bad health (as already indicated by the unconditional means) appears to drive individuals out of their original occupation. This seems reasonable since health impairments often limit the possibility to remain in the old occupation. Individuals with lower education (and eventually lower occupation specific human capital) are significantly more likely to have switching intentions - even when other factors are controlled for. The only notable exceptions here are those who hold a university degree.

The hypotheses of Lazear (2009) are further supported by the coefficients related to characteristics of the last employment: better earning individuals who have been in more stable and skilled employment are more likely to indicate that they are willing to stay in their old occupation. On the contrary, previous occupation switches are positively correlated with further occupation switching. Arguing along the lines of Lazear (2009) this is due to higher opportunity costs of these individuals once they decide to change their occupation. Moreover, larger numbers of vacancies in the old occupation (relative to job-seekers) are negatively related to the intention of occupation switching – as reflected by the coefficient on the labor market tightness variable. This suggests that workers are more likely to stay in their old occupation if their chances to find a new job in that occupation are better given the actual demands of employers. Again, that is what standard theory would predict. Finally, there seems to be a trade-off between geographical and occupational flexibility of workers: those workers who are willing to move to find a new job are also more likely to indicate that they want to stay in their old occupation (or at least an occupation that is closely related in terms of its skill demands).

The descriptive evidence provided here makes us confident that our measure indeed reflects individual's ex-ante (i.e. before any training) willingness of occupation switching and is, thus, useful for the purpose of the heterogeneity analysis that lies ahead.

5 Empirical Strategy

5.1 Identification

We are interested in isolating the effects that publicly sponsored vocational training programs have on future labor market outcomes (denoted as Y) of their participants, i.e. in the average treatment effect on the treated ($ATET$). In particular, our analysis aims at investigating training program effect heterogeneity with respect to occupation switching decisions. Instead of using training participation and the intention to switch occupation as a joint treatment, we analyze heterogeneity by distinguishing between the two strata of participants in the data – those who indicate an intention to switch their occupation at the time when they register as unemployed (s=1) and those who indicate the desire to find a new job in their old occupation (s=0). In that sense, we view the switching intention as an

Table 5.4: Probit estimation of the intention to switch occupation on relevant covariates

Dependent variable: Intention to switch occupation (based on distance measure)	**dY/dX**	**S.E.**
Individual characteristics		
Female	-0.0462***	-0.00536
Age<25	-0.0826***	-0.00992
Age 25-29	-0.0305***	-0.00817
Age 30-34	-0.0139*	-0.00727
Age 35-39	-0.00545	-0.00682
Age 40-44	Reference	
Age 45-50	0.00115	-0.00723
Age 50+	0.0102	-0.0074
School: no degree	-0.00576	-0.0077
School: Hauptschule	Reference	
School: upper secondary degree	-0.0125**	-0.0055
School: university entrance degree	-0.0274***	-0.00847
No vocational degree	0.0324***	-0.00568
Vocational training	Reference	
University degree	0.0332***	-0.0108
At least 1 child (yes=1)	0.00678	-0.0058
Age of youngest child < 3 years (yes=1)	-0.0205*	-0.0119
Age of youngest child 3 to 5 years (yes=1)	-0.00285	-0.00924
Age of youngest child > 5 years (yes=1)	Reference	
Last employment		
Log of wage in EUR (halfmonthly)	-0.0206***	-0.00481
Unskilled worker	0.0762***	-0.00711
Skilled worker	-0.0268***	-0.00731
Part-time worker	0.0143*	-0.00734
White-collar worker	Reference	
Duration last occupation (within last 5 years)	-0.00217***	-0.000129
Occupation specific labour market tightness	-0.122***	-0.0313
Labor market history in 4 years before entering unemployment		
Halfmonths employed	-0.0008	-0.000494
No previous unemployment (1=yes)	0.0148**	-0.00686
Halfmonths unemployed	-0.00111**	-0.000515
Number unemployment spells	-0.0200***	-0.00311
Halfmonths out of labour force	-0.00114**	-0.000506
Any previous training (1=yes)	-0.00759	-0.00761
Additional information		
1 occupation (2-digit) in last 5 years (yes=1)	Reference	
2 occupations (2-digit) in last 5 years (yes=1)	0.0398***	-0.0061
3 or more occupations (2-digit) in last 5 years (yes=1)	0.0528***	-0.0123
Job search not regionally limited	-0.0121***	-0.00427
Vacancy referrals: number per day	-0.0206	-0.015
Vacancy referrals: none	0.0107**	-0.00449
Health impairment affecting placement (yes=1)	0.0369*	-0.0192
Health impairment, disabled (yes=1)	0.0277***	-0.00887
Observations	33,402	
Log-likelihood	-14647.922	
McFadden Pseudo-R^2	0.1067	

Significance levels indicates by stars (*** p<0.01, ** p<0.05, * p<0.1).
Note: The coefficients in the column represent the predicted change in the average probability of the intention to switch occupation with respect to the particular covariate. Additional controls not shown in this table are last occupation, nationality, family status and regional characteristics (local unemployment rate, GDP per capita) as well as state and time fixed effects (year and quarter of entering unemployment).

individual attribute providing additional information for the caseworker's allocation process – relative to the information already used. Thus, we do not intend to obtain the impact of the intention to switch on employment but rather the differential impact of training for switcher and non-switcher.

As widely recognized in the microeconometric evaluation literature, the parameters of interest cannot be easily identified as the counterfactual situation of participants is unobservable. To illustrate the evaluation problem we follow the potential outcome school-of-thought established by Rubin (1974). Formally the $ATET$, i.e. the difference in average potential outcomes for the treated (D=1) denoted as Y^d with d=0 for the non-treatment state and d=1 for treatment, is defined as follows:

$$ATET_s = E(\Delta \Big| D = 1, S = s) = E(Y^1 \Big| D = 1, S = s) - E(Y^0 | D = 1, S = s) \qquad (1)$$

for the sub-group of switchers and non-switchers, respectively. The comparison of $ATET_1$ and $ATET_0$ allows us to make statements about potential effect heterogeneities. In order to assess program effectiveness we focus only on the effect on those unemployed who actually participate instead of investigating the impact for a person randomly drawn from the population which could be obtained by the average treatment effect (ATE). We also refrain from analyzing the effects on the non-treated, i.e. how FVT might affect those who are eligible but decide not to participate as there is no policy interest in extending the program to the entire population of unemployed. Since the last term in the equation above is not observable we approximate it using the outcomes of non-participants. Given, that we are not in an experimental setting with random training assignment we need further assumptions in order to use non-participants as a valid comparison group because $E(Y^0 \Big| D = 1, S = s) \neq E(Y^0 \Big| D = 0, S = s)$. A naive comparison of their mean outcomes would otherwise result in biased estimates of the $ATET$s as unemployed are selectively assigned to programs and selectively intend to switch occupation based on various individual characteristics. This is obvious given the caseworker's selection process and the apparent differences in covariates between participants and non-participants, and switchers and non-switchers, respectively (as shown in the descriptive statistics of our sample in Table 5.A.1. We account for the selection bias by following a selection on observables strategy. The underlying conditional independence assumption (Imbens (2004)) postulates that conditional on all covariates X that jointly influence treatment assignment and outcomes within each stratum $S = 0, 1$, the treatment status (D) can be seen as quasi random (exogenous) so that $(Y^1, Y^0) \perp D | X = x, S = s$, and, hence, $E(Y^0 \Big| D = 1, X = x, S = s) = E(Y^0 | D = 0, X = x, S = s)$. This implicitly assumes that all unobserved factors are - conditional on X - equally distributed between participants and non-participants within each stratum. In order to ensure that the estimated training effects between switcher and non-switcher do not vary because of differences in other observed characteristics (other than the switching intention) that influence their labor market outcomes we apply an additional adjustment step. In particular, we identify effect heterogeneity by defining one target group, namely participants who intend to switch their occupation, and adjust the covariate distributions of the other three subgroups to match the target group's one. The rationale for this additional adjustment step is not that we view the switching status as conditionally independent, i.e. as a treatment state. Instead we believe the variation in S to provide additional information to the caseworker that is not fully described by the other observed covariates.

By selecting the sample based on stable employment records (at least in the time before becoming unemployed) we make sure that all individuals are in fact eligible for the program. Participation in vocational training and switching intentions are determined in the meetings of the caseworker and the unemployed. While caseworkers have some discretion about who they assign, the approval is substantially based on their assessment of successful program completion, subsequent employment prospects, local labor market conditions as well as available program places and financial constraints of the local employment office. We capture the selectivity of this assignment process and the individual decisions using an extensive list of variables (all measured before program start). We include socio-demographic characteristics of the unemployed (e.g. gender, age, nationality, marital status, number of children, education, the degree of mobility, etc.), their labor market histories (past employment, unemployment and out-of labor force spells, past training participation, earnings, sector of employment, vocational degrees, characteristics of the last employer, etc.) and individual specific limitations (e.g. health and disability status) to control for individual employment prospects and program completion. Local labor market conditions are taken into account using regional characteristics such as the local unemployment rate, regional GDP or urbanity as well as occupation specific labor market tightness. From the perspective of the unemployed it is the same factors that, we argue, drive selection into participation and compliance with the rules of the program. Though, in order to account for the incentive to enhance unemployment benefit receipt with program participation we control for the claim and the amount of unemployment benefits measured at the beginning of the unemployment spell. With respect to unobserved characteristics of the unemployed, we are confident that systematic differences are sufficiently captured as these should otherwise have materialized in the past (Lechner and Wunsch (2009)).

Our identification strategy allows, through appropriate reweighting of the covariate distribution in the comparison groups, to match that of the target group and, thus, closely mimic an experimental set-up. When simply testing for $ATET_1 = ATET_0$ as defined above we implicitly test if

$$[E(Y^1|D=1,S=1) - E_{X|D=1,S=1}E(Y^0|D=1,X=x,S=1)]$$

$$- [E_{X|D=1,S=1}E(Y^1|D=1, X=x, S=0) - E_{X|D=1,S=1}E(Y^0|D=1, X=x, S=0)] = 0 \tag{2}$$

5.2 Estimation

Following our identification strategy, we estimate the reweighted conditional expectations for the outcomes of interest in each of the sub-populations. For this purpose we use the semi-parametric matching estimator as proposed in Lechner et al. (2011) which allows for effect heterogeneity and avoids unnecessary functional form assumptions.[73] We first estimate the probability of treatment (i.e. the propensity of being a training participant with switching intention) using parametric probit models that include all relevant covariates as described in the previous section.[74] We then conduct specification tests for omitted variables, but also rely on Lechner and Wunsch (2011) for the sets of included covariates who – based on IAB data - provide helpful guidance for this step. The estimated propensity scores are then used as a reweighting device. Instead of nearest neighbor matching the estimator used here is based on the idea of radius matching (Dehejia and Wahba (2002)). First the distance between each treated and the closest non-treated observation is computed based on the Mahalanobis metric using the covariates and the propensity scores. Here, the propensity scores are weighted 5 times more than the covariates. In a second step the radius around the treated observation is defined. For the particular value of the radius we rely on the one used in Lechner et al. (2011) as there is no specific algorithm available in the literature to obtain a certain value. Thus, the radius is defined as percentage of the maximum Mahalanobis distance – in our case the 90%-distance to the largest one-to-one match. Within this radius each non-treated observation is weighted proportional to the inverse of its Mahalanobis distance. Thus, the resulting weights correspond to the relative importance of non-treated observations in the overall comparison of the outcome variables. Based on the weights obtained from the propensity score estimation we ensure sufficient common support by dropping treated for whom no close match is available. In particular, we apply the procedure proposed by Dehejia and Wahba (2002) that removes all treated observations with propensity scores larger than the largest propensity score within the control group. In addition, we acknowledge the problem of areas with thin common support by using a trimming rule as suggested in Huber et al. (2013). It is shown that trimming can reduce the mean squared error substantially relative to no trimming. This

[73]The study by Huber et al. (2013) provides further details on the performance of this particular estimator in comparison to a broad range of other propensity score-based estimators. Moreover, Huber et al. (2012) provide additional results on the various tuning parameters that need to be set in this type of estimators and implement these in different software packages.

[74]Based on this probit estimation we are able to account for a large number of covariates circumventing the curse of dimensionality by summarizing the covariate information into a one-dimensional score. If the treatment status is conditionally independent (conditional on X) from the potential outcomes this holds also for any one-dimensional function of X (Rosenbaum and Rubin (1983)).

rule accounts for values of the propensity score of the treated that are rare among the controls. In the absence of trimming these controls would obtain too much weight in the estimation and, in turn, might dominate the results. To avoid this problem, trimming restricts the maximum weight these controls obtain, i.e. "all weights are set to zero if their share of the sum of all weights is larger than t%." We started with 6% trimming as used in Huber et al. (2013) as one particular value. This specification reduced the root mean squared error on average for all estimators they investigate – which includes the one we use here – by a fairly large amount compared to no trimming. An inspection of the common support (documented in Appendix B figures 5.B.1 to 5.B.3) shows that in our estimations there are some areas with relatively thin common support. We further reduced the level of trimming excluding observations with a weight larger than 1% (as share of the sum of all weights) to reduce the relatively high variance that we obtained. We kept this value, on the one hand, because a further reduction would have led to too few observations 'on support'. On the other hand, we choose 1% because the point estimates did not change much compared to a 6% trimming level but the variance was reduced. Eventually the estimated weights are then used in a final parametric regression model to achieve further bias reduction where we use linear and logit bias correction in particular. When using the weights we utilize the double robustness property, i.e. our results are consistent if either the participation model or the final regression model is correctly specified. Sampling uncertainty of the estimates is approximated based on 999 bootstrap replications of the effects. The confidence intervals are expected to be larger for the adjustment of the non-switcher participants to the switcher participants since these groups are the smallest. To eventually make statements about effect heterogeneity and program effects in general on a 95% confidence level we consider the 2.5th and 97.5th percentile of the bootstrap distributions.

6 Results

6.1 The Outcomes

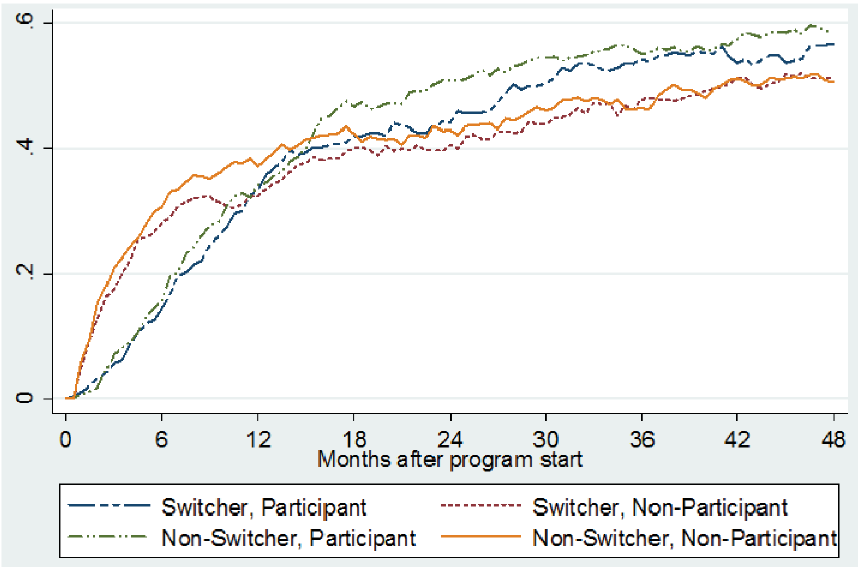
We investigate the effects of participation in vocational training between 2000 and 2004 and consider an outcome period of up to 4 years after training start (i.e. up to 2008). Because all outcomes are measured relative to the time of (hypothetical) program start we are able to make statements about initial lock-in effects that are usually observed while the programs are running (Van Ours (2004)) as well as medium-run effects. The following outcomes are investigated: First, employment rates at a half-monthly frequency and accumulated during different periods after training start. Successful reintegration into the labor market and stable employment are the formulated goals of FVT. Thus, we view this outcome as the most general, but nevertheless as very relevant to draw conclusions about program effects. Second, as the focus of this study is on occupational mobility, we investigate the rate of successful occupation. Second, as the focus of this study is on occupational mobility, we investigate the rate of successful occupation switching. As discussed in before successful training should qualify unemployed participants to meet the skills demanded by the labor market. In cases where an unemployed individual is unlikely to find a job in her old occupation (and is therefore considering a change) a positive effect on successful occupation switches is desirable from a policy perspective. Yet, this can only be evaluated for the selected group of individuals who are observed in employment at some point after training within the outcome period. Successful occupation switching is defined as a binary outcome measure being 1 if individuals are employed in an occupation different from their last occupation according to our distance measure and 0 otherwise, which includes non-employment as well as employment in an occupation that is equal or "similar" to the one before unemployment (i.e. the distance is <0.06). As for the crude employment rate, the outcomes on occupation switches are measured on half-monthly rates and accumulative. Third, we investigate indicators for accumulated earnings over time - they are interpreted as a crude measure for the productivity level of individuals.

The results of our covariate adjustments are presented in Table 5.B.1 which documents the estimated probit models for the three necessary adjustment steps described above where the dependent variables are training participation and intention to switch, respectively, and we obtain the propensity scores. The inspection of balancing tables (before any additional bias correction) based on the estimated propensity scores suggest a good performance of the reweighting of mean characteristics and, thus, a good match quality. Also the common support requirement seems satisfactory. According to the distribution of propensity scores, depicted in Figures 5.B.1, 5.B.2 and 5.B.3 the vast majority of treated observations are in the common support.

6.2 Effects of Training

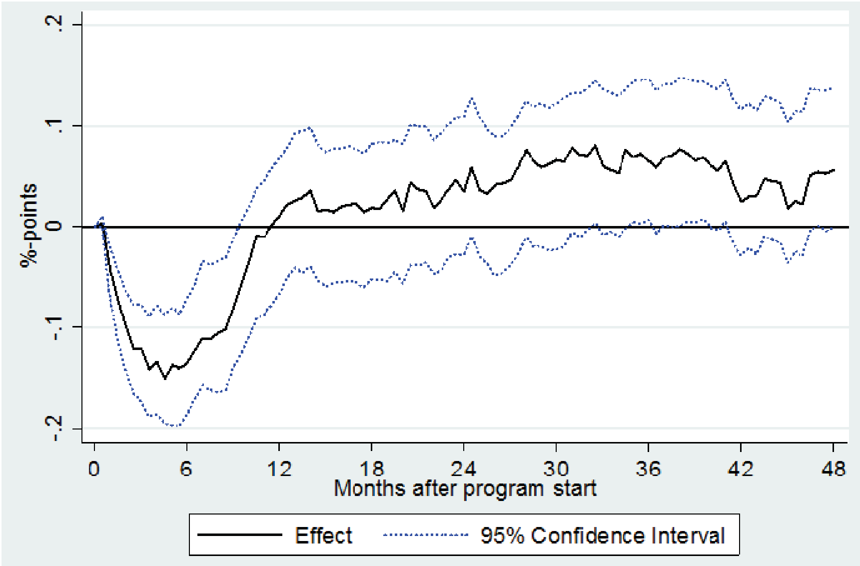
Figure 5.1 shows the mean (unsubsidized) employment rates of the four groups over time. For the two participant groups we see that their employment rates behave very similarly. This is especially true for the first 12 months after program start. During the initial period the non-participants have higher employment probabilities which we attribute to the lock-in effects of the program. By the end of the first year all groups have employment rates around 35 to 40%, while the rates of the participant groups grow faster afterwards - most likely because of the end of their training program (planned durations are between 6 to 12 months). At the end of the observation period the two trained groups show higher employment rates (>56%) than their respective comparison groups (<50%). This is confirmed by Figures 5.2 and 5.3 showing the training effects of participants over time. The negative impact on employment rates during the first year simply confirms the existence of significant lock-in effects. Moreover, both figures show positive employment effects of training for the participant groups approximately 2 years after program start. The point estimates suggest, that in the long-run participation in FVT leads to an approximately 8%-points higher employment rate. This effect remains relatively stable and becomes statistically significant towards the end of the observation period for non-switcher, whereas the effect for switcher declines after 3.5 years after program start and becomes insignificant. In Figure 5.4 we depict the difference in program effects for switcher and non-switcher to illustrate potential effect heterogeneities. Here, the impacts on employment rates over time do not show any heterogeneous program effects between these two groups as the difference is always very close to zero. Given the rather large confidence bounds we can however not conclude that unemployed with a switching intention really do not benefit more (or less) from FVT programs than non-switcher. This figure only allows to conclude that in case the effect is not zero it is also not larger or smaller than the confidence bounds, which is rather unsatisfactory though.

Figure 5.1: Employment probabilities



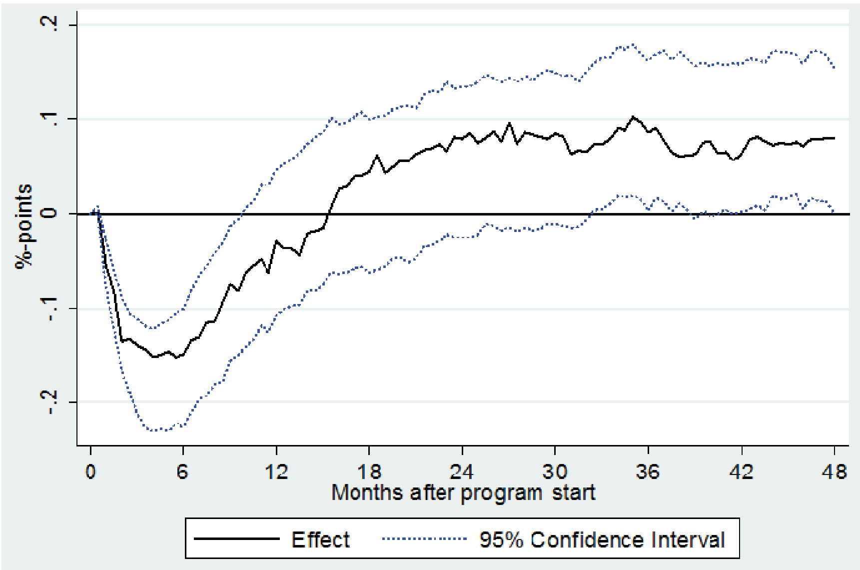
Note: Ordinate measures employment probability in %. Employment probabilities are obtained considering unsubsidized employment only.

Figure 5.2: Program effects for participants with switching intention: Employment probability in unsubsidized employment



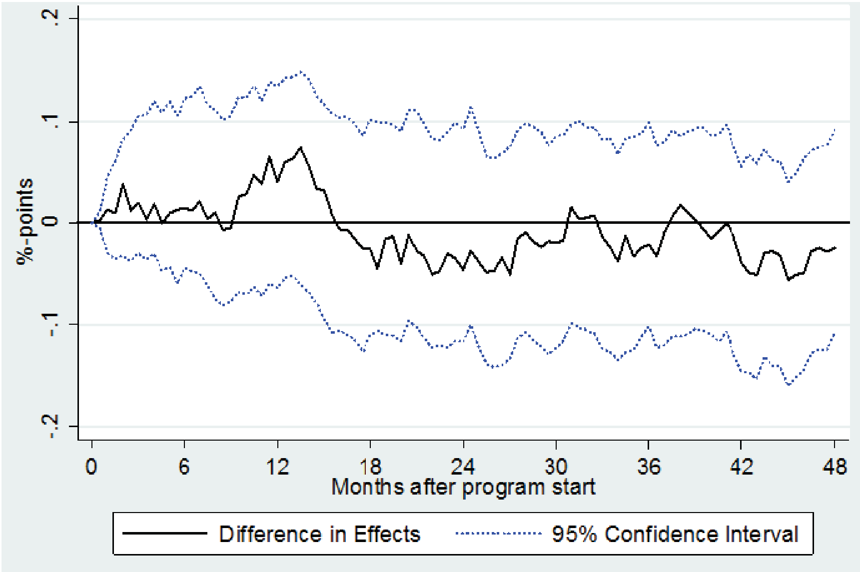
Note: The 95% confidence interval refers to the 2.5th and 97.5th percentiles of the distribution of effects obtained from 999 bootstrap replications.

Figure 5.3: Program effects for participants without switching intention: Employment probability in unsubsidized employment



Note: The 95% confidence interval refers to the 2.5th and 97.5th percentiles of the distribution of effects obtained from 999 bootstrap replications.

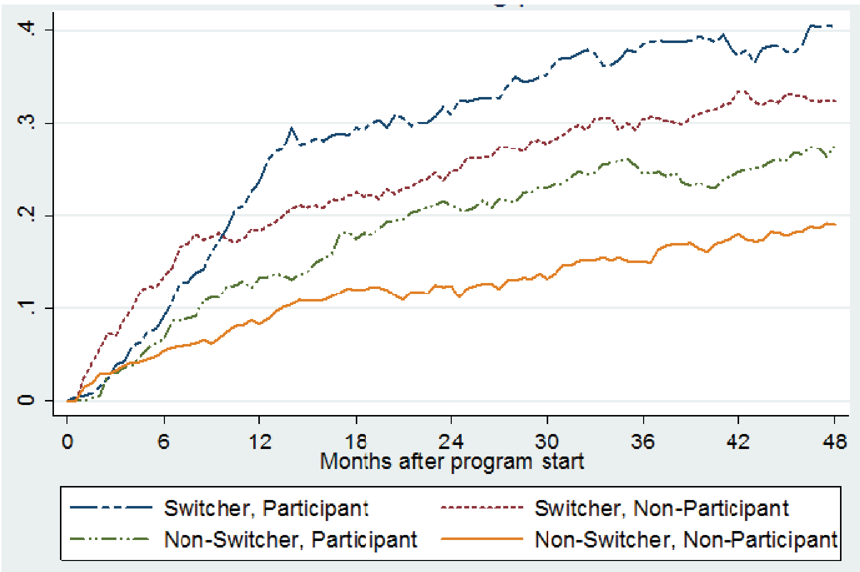
Figure 5.4: Difference in program effects between participants with and without switching intention



Note: The presented results are obtained by subtracting the effect of non-switcher from the one of switcher. Thus, a positive value of the difference indicates effect heterogeneity in favor of switcher. The 95% Confidence Interval refers to the 2.5th and 97.5th percentiles of the distribution of effects obtained from 999 bootstrap replications.

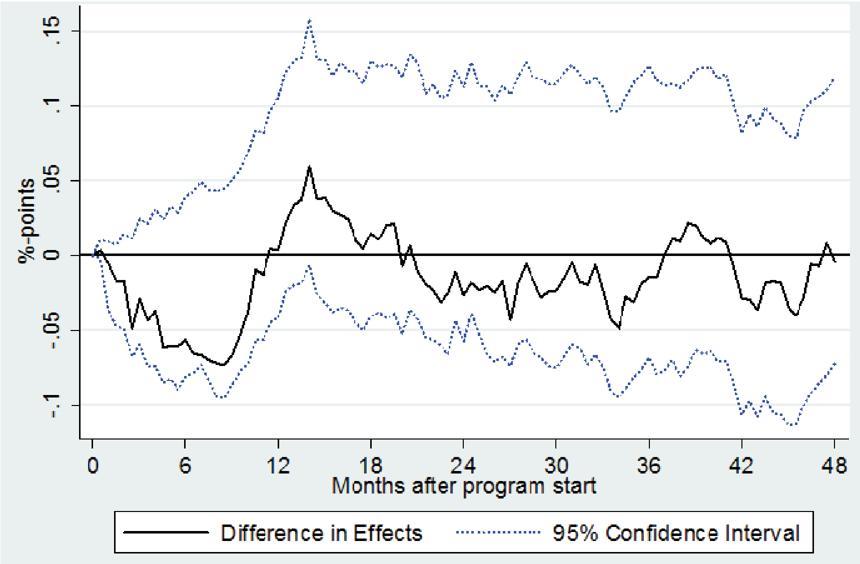
Figure 5.5 visualizes the development of employment probabilities with successful (ex-post) occupation change after training start. Remember that we defined this outcome variable to be 1 if individuals are employed in a new occupation according to our distance measure and 0 otherwise which includes non-employment and employment in an occupation that is 'close' to the old one. It appears that, on average, those who indicate an intention to switch at the beginning of their unemployment spell indeed have higher probabilities of being employed in an occupation that is different from their last occupation. This holds for program participants and non-participants. Again, the training participants in each stratum show higher probabilities of occupation change which is, again, defined as being employed in an occupation with a skill distance larger than 0.06 (away from the occupation of last employment). The training effects for the outcome of successful occupation switches are shown in Figures 5.B.3 and 5.B.4. Our analysis detects a positive effect of training participation on successful occupation switching afterwards. The effect is largest (about 9%- points) for both groups after approximately 3 years. For participants with switch intentions there is still a lock-in effect observable indicating that without program participation they would have found a job in a new occupation already in the first year. Later on the employment rate increases for participants and they experience significantly higher employment probabilities in a new occupation for the last 3 years of the observation period than if they had not participated in FVT. The development is somewhat different for non-switcher (Figure 5.B.4) where program participation does not result in lower employment rates in the initial time after training compared to non-participation. Already after 6 months the mean effect of the program becomes positive which implies that program participation (at least partly) induces higher employment rates in a new occupation compared to non-participation for individuals who initially do not intend to change their occupation. However, this observed effect is insignificant until approximately 3 years after program start. Eventually, there is again only little heterogeneity across the two groups as indicated by the difference in program effects (Figure 5.6). In the first 8 months after program start switcher actually seem to perform worse than non-switcher, whereas they catch up until the end of the first year. Afterwards the difference in effects is always close to zero and as before has a large confidence bound over the entire outcome horizon. Thus, our results of the effects of training on employment combined with successful occupation change do not provide conclusive evidence on effect heterogeneity either. Note also that these conclusions are limited as successful occupation switch is only observed for the selected group of individuals who are employed after training at least once within the outcome period.

Figure 5.5: Successful switching probabilities



Note: Ordinate measures employment probability in %. Individuals are employed in a different occupation if the difference between new and old occupation is larger than 0.06 according to our distance measures.

Figure 5.6: : Difference in program effects (successful switching) between participants with switching intention and those without



Note: The presented results are obtained by subtracting the effect of non-switcher from the one of switcher. Thus, a positive value of the difference indicates effect heterogeneity in favor of switcher. The 95% Confidence Interval refers to the 2.5th and 97.5th percentiles of the distribution of effects obtained from 999 bootstrap replications.

Turning to the effects of training on our accumulated outcomes, Table 5.5 shows the estimated average outcome levels as well as the program effects (and their differences) for the four sub-groups. The employment outcomes - measured in half-months individuals spend in unsubsidized employment - confirm the negative and significant lock-in effects for both strata from before. On average, participants with (no) switching intention are employed 4 (5) weeks less than their non-participant comparison group during the first year after program start. Towards the end of our observation period, in particular in the last year, we see that program participants (switcher and non-switcher) accumulate about 2 to 3 weeks (1.21 and 1.75 half-months, respectively) more in employment than they would have without participation. Also the cumulated effects for the last 2 years are positive. Program participation increases employment for switchers by about 5 weeks (only weakly significant) and for non-switcher by more than 7 weeks. In turn, participants with switching intention experience about 5 weeks (only weakly significant) less in unemployment in the last 2 years and non-switcher almost 8 weeks (3.79 half-months). Over the entire 4 years, however, participants cannot compensate for the initial lock-in effect and, thus, we do not find any significant effects of program participation on employment stability over the whole outcome period. Remember, however, that the estimated confidence bounds are rather large – in fact too large to conclude that the program has no impact. Regarding effect heterogeneity the last columns depict the differences in program effects for switchers and non-switcher. Again, there is no evidence that training has differential impacts for the two sub-groups.

The results also provide insights about the switching patterns of individuals in our sample. As expected, training appears to lower the attachment of unemployed workers to their previous occupations - by 5 and 4 weeks for switchers and non-switchers (-2.28 half-months and -1.91 half-months), respectively, during the last two years of the observation period. This indicates a positive effect on occupational mobility even though it is only significant for the subgroup of switchers. At the same time it shows no significant effect on the time employed in the intended occupation. Instead, the implied higher occupational mobility is a result of realized switches into occupations other than either the last occupation or the target occupation, i.e. into an alternative occupation as defined by our distance criterion. Switcher participants accumulate almost 6 weeks (2.72 half-months) and non-switcher participants about 13 weeks (6.18 half-months) more in a new occupation during the 3rd and 4th year after program start. Thus, training appears to stimulate occupational mobility, yet is less successful to place participants into their originally intended occupations.

Finally, we provide outcomes related to accumulated wage earnings. For the 4th year as well as for the last 2 years after program participation the results suggest some positive effects on total earnings - significant for both sub-groups for the last year of the observation period. However, as for the employment outcomes, these gains cannot compensate for the initial earnings losses during the first year (about 2000 € less for participants of both groups). Over the entire outcome period program participants do not accumulate more earnings than if they had not participated in FVT. Clearly, these results are closely related to the accumulated employment outcomes mainly reflecting the negative (positive) effects in the first (last) year after program start.

6.3 Sensitivity Analysis

Variation of the occupation distance measure

To demonstrate the sensitivity of our results with respect to the distance measure used to determine the intention to switch occupation we reduce the threshold by 10% to 0.054. Accordingly, all skill distances between last and target occupation larger than 0.054 are 'intended occupation switches'. Thus, we classify more individuals as switcher than previously. Note, however that now we define differences between last and target occupation as intended moves to a new occupation which we initially did not classify as those. For example, according to Table 5.C.1 (appendix) presenting again the most frequent intended occupation switches following the new distance threshold, changes from 'Office Clerk' to 'Account, Bookkeeper' and vice versa are now the main groups of intended occupation switches. Nevertheless, we still believe to capture mostly relevant switches, i.e. occupational moves that still require substantial skill adjustments. We conduct the same estimation steps as before to obtain the respective outcomes.

The results presented in Figures 5.C.1 and 5.C.1 in the appendix depict the evolution

Table 5.5: Program Effects of FVT on cumulated labor market outcomes

	With intention to switch			Percentile		Without intention to switch			Percentile		Difference	Percentile	
Outcome	Participants (1)	Non-participants (2)	Effect (3)	2.5th	97.5th	Participants (4)	Non-participants (5)	Effect (6)	2.5th	97.5th	(3)-(6)	2.5th	97.5th
Halfmonths employed													
1 year after	3.83	5.84	**-2.01**	**-3.02**	**-1.05**	4.17	6.61	**-2.44**	**-3.71**	**-1.54**	0.43	-0.78	2.16
4 years after	38.96	37.66	1.30	-3.33	5.87	41.60	39.55	2.05	-3.49	6.85	-0.75	-6.77	6.57
4th year after	13.16	11.95	**1.21**	**0.02**	**2.99**	13.75	12.00	**1.75**	**0.41**	**3.80**	-0.54	-2.80	1.47
3rd and 4th year after	25.25	22.57	2.68	-0.17	5.53	26.69	22.98	**3.71**	**0.68**	**7.12**	-1.03	-4.96	2.91
Unemployed 3rd and 4th year after	18.96	21.51	-2.55	-5.23	0.37	17.15	20.94	**-3.79**	**-7.00**	**-0.75**	1.24	-2.75	5.29
Halfmonths employed 3rd and 4th year													
In last occupation	4.04	6.33	**-2.28**	**-4.49**	**-0.43**	8.08	9.99	-1.91	-4.62	0.71	-0.37	-3.96	2.81
In intended occupation	5.78	3.77	2.01	-0.61	2.93	8.54	10.60	-2.06	-3.96	1.03	4.08	-0.44	5.79
In other occupation	19.22	16.50	**2.72**	**0.96**	**6.59**	21.05	14.88	**6.18**	**2.03**	**8.68**	-3.46	-6.06	2.52
Wage earnings													
1 year after	3231	5185	**-1954**	**-2959**	**-834**	3731	6040	**-2310**	**-3364**	**-1077**	355	-1106	1901
4 years after	32047	29904	2143	-3682	6877	36246	34433	1814	-2145	9064	329	-9280	5848
4th year after	10635	8855	**1780**	**78**	**3358**	11601	9973	**1628**	**739**	**4334**	152	-3087	1625
3rd and 4th year after	20607	17136	3471	-149	6252	22913	19526	**3387**	**1405**	**8189**	85	-5986	3119

Note: The reported 2.5th and 97.5th percentiles are based on 999 bootstrap replications of the effects. Numbers in bold indicate significance at a 95% level.

of the employment rates over time for each sub-group and the difference in training effects between switcher and non-switcher, respectively, are not different from the ones we obtained before. The same is true for the outcome of successful occupation switching (Figure 5.C.3). The cumulated employment outcomes (Table 5.C.2) neither indicate any different insights such that a variation of the distance measure in the proposed way does not result in new conclusions.

Variation in the training program participation definition – '18 months window'

According to our previous definition of program participation and the discussion of Fredriksson and Johansson (2003) we estimate our results again defining each unemployed individual who participates at least once in a vocational training within 18 months after entering unemployment as participant. This definition results in a new sample with in total 28872 observations. In Table 5.C.3 we present the mean values of selected characteristics of the individuals in this new sample. Except for different numbers of observations in each sub group – we now observe 569 program participants with a switching intention (501 before) - there are no substantial differences in the mean characteristics observable compared to the original sample we use for our primary results.

The findings, as before, suggest positive effects of training participation on employment rates (unsubsidized employment) 3 years after program start (see Figures 5.C.4 and 5.C.5) but again no significant difference between those effects (Figure 5.C.6) which would otherwise imply heterogeneity. In addition, the employment probabilities with successful occupation switching as well as the cumulated outcomes (Table 5.C.4) are very similar (especially with respect to statistical significance) to the results we obtain from our initial sample. Thus, our findings are not sensitive to the definition of program participation.

7 Conclusion

In this paper we investigate the effectiveness of vocational training programs with a special focus on the aspect of occupational mobility - something that the existing literature has largely neglected so far. More specifically, we discuss the interactions between vocational training and occupational mobility, and argue that these should be accounted for in any evaluation study. Based on evidence from an extensive body of previous research in labour economics it is safe to assume that job changes are often accompanied by losses of human capital and declining wages. This is especially true for unemployed individuals who - because of bleak reemployment prospects in their old occupation - switch between occupations with different skill requirements (Gathmann and Schönberg (2010)). If applied adequately, vocational training should cushion the negative effects of such occupation changes as it may help to (partly) overcome skill deficits and, thus, reduce subsequent productivity and wage losses. In addition, training towards a new occupation can provide a positive signal and raise the attractiveness of unemployed workers to potential employers. On the contrary, training might be less useful, if an unemployed worker aims to stay in his previous occupation where he faces favorable job prospects or when his skills are highly demanded by the labour market. Thus, the usefulness of training might well differ between individuals with and without intention of occupation switch.

While the theoretical arguments are clear, empirical evidence is not easy to establish because occupation switching is a choice variable that is usually only observed after training - and thus potentially endogenous. Based on large administrative data from Germany we identify two sub-groups – i.e. unemployed individuals who intend to switch their occupation and individuals who do not. In contrast to other studies we are in the lucky position to observe a reliable measure for switching intentions that is recorded before any decision about training participation is made, i.e. exogenous to training. Within each of the groups we identify and estimate the effect of training participation on subsequent labor market outcomes – i.e. chances of future unsubsidized employment, occupational mobility patterns, as well as accumulated wage earnings. We also compare the effectiveness across the two groups to learn more about potential effect heterogeneity. The availability of many individual characteristics enables us to use advanced econometric matching methods to reasonably capture the selectivity into training and occupation switching, respectively.

Given our medium-term outcome horizon of up to four years after program start our we find positive effects of training for both sub-groups - switchers (intention to switch) and stayers – with respect to employment probability, successful occupation switching and accumulated wage earnings. The medium term effects become most apparent when only the later periods are considered (i.e. 3-4 years after training). But, due to initial lock-in effects, we do not find significant effects of training on employment for the total outcome period (Van Ours (2004)). Contrasting with our theoretical predictions, however, the treatment effects for both sub-groups show similar patterns and we do not find significant differences between them. The interpretability of that result is, however, limited due to sampling uncertainty in the data. Most of the outcomes we consider provide no evidence for a differential impact of training except that training appears to be more effective in promoting successful (ex-post) occupation switches in the short run for individuals with (ex-ante) switching intentions.

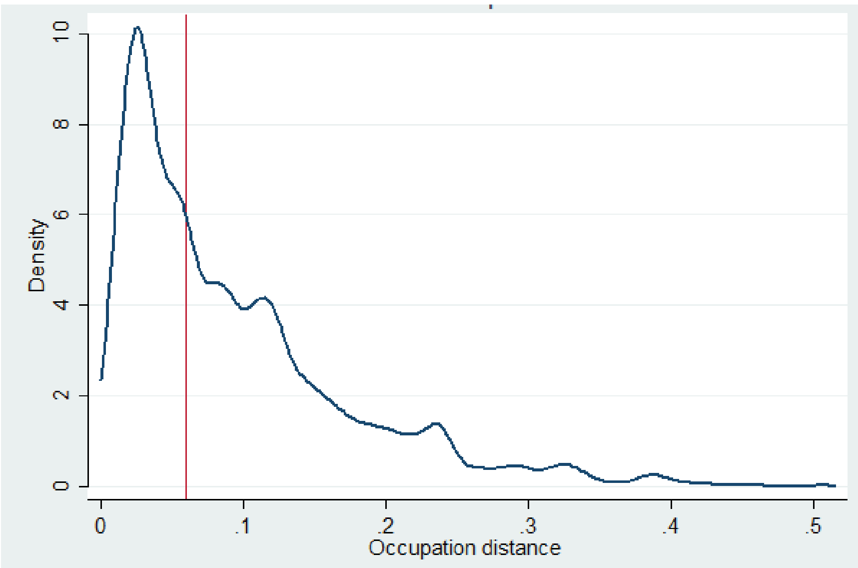
Our inability to demonstrate heterogeneous treatment effects should not suggest that the dimension of occupational mobility is to be deemed as unimportant and therefore ignored altogether. While we show that observable characteristics explain a decent part of the variation in switching intentions of unemployed workers in our sample, a large fraction remains unexplained. This unexplained part is likely to reflect unobserved individual traits, preferences, experiences and circumstances, among other things. While these dimensions are generally difficult to assess, also from the perspective of case workers, it is hard to argue that are irrelevant for labour market outcomes of individuals. Having reliable data about switching intentions might allow experienced case workers to draw relevant conclusions about such worker characteristics and, thus, help to optimize training allocation. Under the premise that occupational mobility is a crucial determinant of subsequent labour market pathways - and we know that it is (Kambourov and Manovskii (2009)) - information about switching intentions should achieve greater importance in the treatment of unemployed workers.

Kambourov et al. (2012) are the first authors to explicitly analyze the interaction between occupational mobility and labour market training. Although our paper does provide additional insights on the topic, it has also shortcomings that should be addressed in future research. First, a key issue we do not address is the direction of mobility, i.e. upward vs. downward mobility. Due to small sample sizes and the lack of a well defined ranking of occupations, we pooled all individuals with apparent switching intentions into one group. Yet, there are strong arguments of why these groups should be considered separately. Individuals who aim for upward mobility (regardless of the ranking scale) should be assigned to different types of training programs. (e.g. for them we would expect skill upgrades to be most important) and be very different in their characteristics when compared to downwardly mobile workers. A closer look at this dimension and a separation of the two subgroups should reveal further interesting insight. In addition, one should look more deeply into the content of training programs. Certain types of programs could be more efficient for workers with switching intentions, while other types of programs maybe be preferable for workers without switching intentions. Second, in line with existing literature, we classify occupation switches based on ISCO codes (further adding a minimum skill distance requirement to improve the robustness and minimize measurement error). This is, however, sub-optimal as the recording of ISCO codes is known to be plagued by random noise which results in an overestimate of occupation switching. Admittedly, we would have preferred a worker assessed measure of occupation switching instead. That variable could be collected at minimal cost (e.g. by just asking the unemployed worker directly, maybe even at different points during his unemployment spell) and would certainly improve the reliability of our analyses.

8 Appendix

A - Sample Descriptives

Figure 5.A.1: Distribution of occupation distances



Note: The vertical line indicates the median of this distribution and, thus, the threshold for our intention to switch occupation definition.

Table 5.A.1: Selected mean characteristics by sample subgroups

	Non-participants		Participants	
Intention to switch	No	Yes	No	Yes
Observations	24660	5954	2299	501
Baseline individual characteristics				
Age	38.32	39.20	37.58	38.32
School: no degree	0.08	0.13	0.03	0.06
School: Hauptschule	0.44	0.48	0.30	0.35
School: upper secondary degree	0.33	0.29	0.45	0.41
School: university entrance degree	0.14	0.10	0.22	0.18
No vocational training	0.22	0.36	0.12	0.21
Vocational training	0.71	0.59	0.78	0.71
University	0.06	0.05	0.10	0.09
Female	0.42	0.39	0.47	0.37
At least 1 child	0.32	0.36	0.35	0.41
Single	0.37	0.33	0.39	0.31
Couple not married	0.04	0.05	0.04	0.06
Single with child	0.05	0.05	0.05	0.07
Married	0.53	0.57	0.52	0.57
Timing of unemployment and program start				
Unempl start: in half-months (1=Jan 2000)	61.51	59.92	53.58	49.45
Unempl. Duration until program start: half-months	6.56	7.10	9.09	9.22
Last employment: non-firm characteristics				
Unskilled worker	0.21	0.41	0.12	0.34
Skilled worker	0.33	0.23	0.24	0.21
White-collar worker	0.31	0.20	0.51	0.31
Part-time worker	0.15	0.16	0.13	0.14
Last employment: firm characteristics				
Age of firm in years	12.83	12.50	12.24	12.19
Industry: Manufacturing	0.19	0.25	0.24	0.27
Industry: Construction	0.22	0.14	0.14	0.12
Industry: Retail	0.15	0.14	0.18	0.16
Industry: Service higher skilled	0.07	0.04	0.14	0.07
Industry: Health and social services	0.10	0.10	0.06	0.10
Industry: Service lower skilled, tourism	0.17	0.20	0.15	0.18
Industry: Primary sector, other services	0.10	0.13	0.08	0.11
Short-term labor market history (last 4 years before entering unemployment)				
Halfmonths employed	78.02	72.07	80.81	74.59
Halfmonths unemployed	8.54	11.21	6.71	9.24
Halfmonths out of labor force	7.99	10.86	7.07	10.28
Long-term labor market history (last 10 years before entering unemployment)				
Average employment duration	67.70	60.80	71.95	67.98
Number unemployment spells	1.45	1.48	1.13	1.17
Average unemployment duration	10.47	14.32	10.23	12.88
Unemployment benefits and claim				
Amount of unempl benefit per 2 weeks in EUR	334.50	317.50	320.34	307.84
UB claim <12 months	0.26	0.31	0.23	0.29
UB claim 12 months	0.44	0.41	0.54	0.51
UB claim >12 months	0.29	0.29	0.23	0.20
Additional information				
Halfmonths last occupation in last 5 years	85.83	70.00	87.28	71.56
1 occupation (2-digit) in last 5 years	0.66	0.47	0.63	0.48
2 occupations (2-digit) in last 5 years	0.27	0.40	0.29	0.38
3 or more occupations (2-digit) in last 5 years	0.07	0.13	0.07	0.14
Last employment: monthly wage in EUR	1806	1621	1904	1792
Cumulated earnings in last year	20051	18337	21528	20509
Labor market tightness	0.09	0.08	0.09	0.09

Note: All entries are mean values measured at the time of entering unemployment except for the variable 'unemployment duration until program start'

B - Estimation Results

Figure 5.B.1: Common support for program participation of switchers – distribution of estimated propensity scores

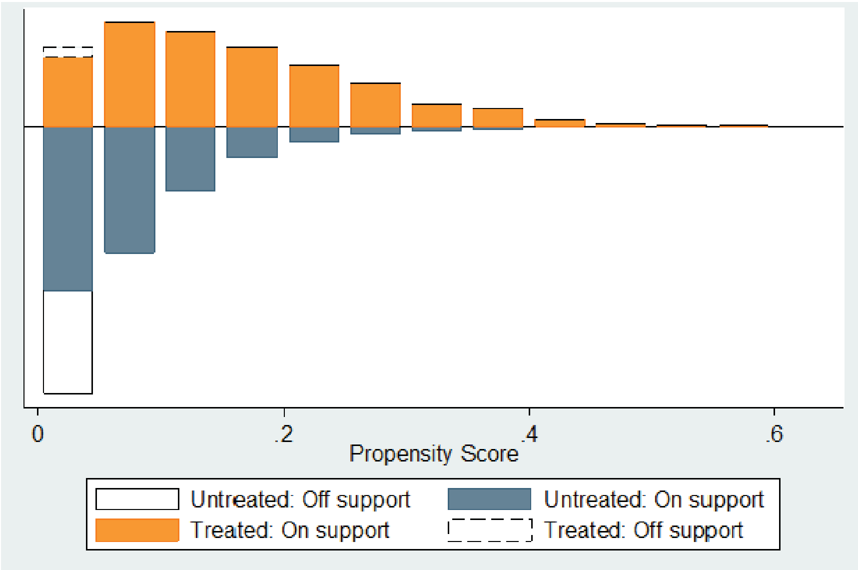


Figure 5.B.2: Common support for covariate adjustment of participants with switching intention to non-participants without switching intention – distribution of estimated propensity scores

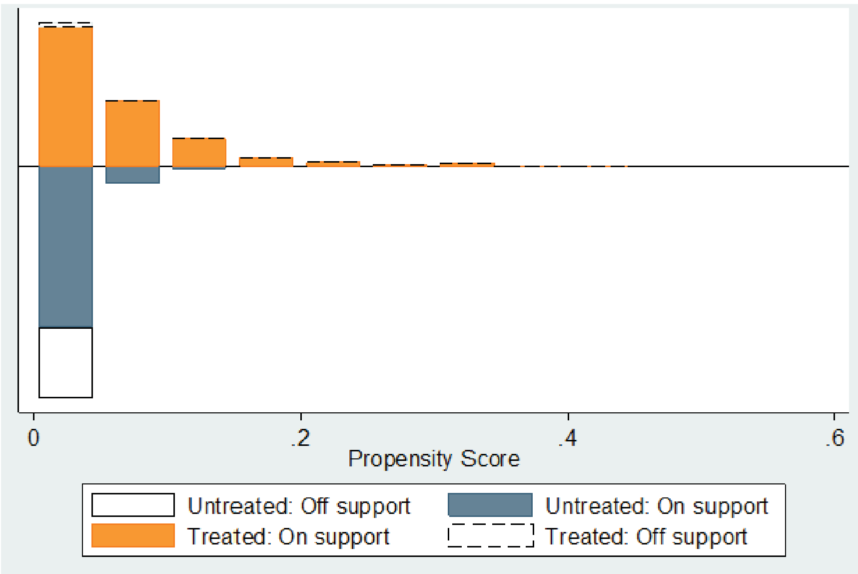


Figure 5.B.3: Common support for covariate adjustment of participants with switching intention to participants without switching intention – distribution of estimated propensity scores

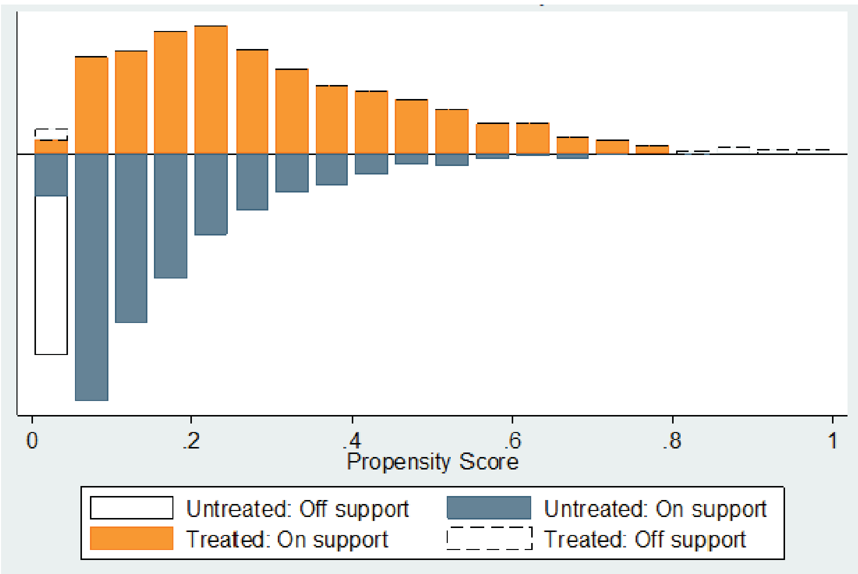
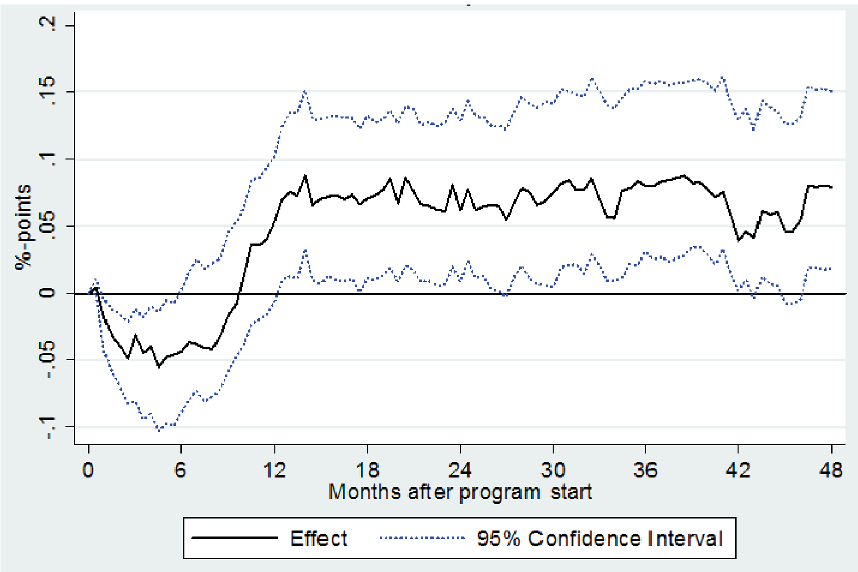
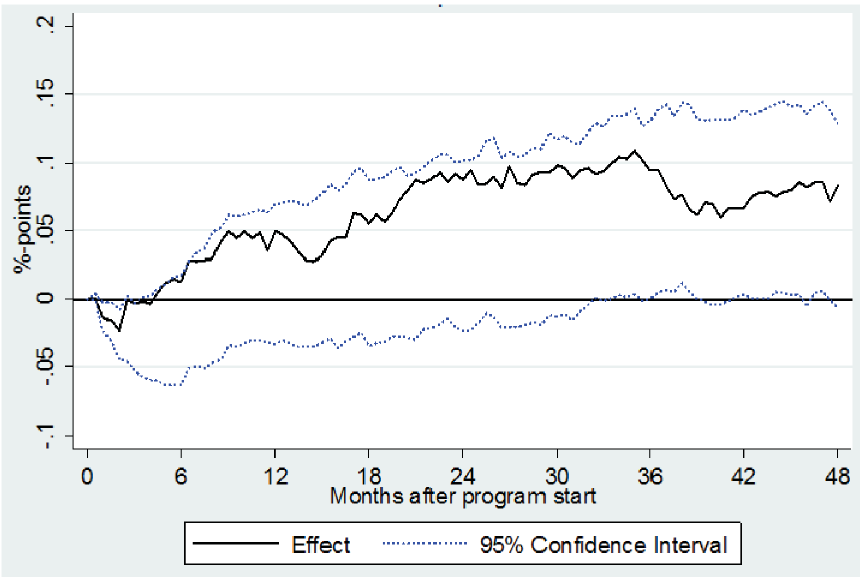


Figure 5.B.4: Program effects for participants with switching intention: Successful occupation switch



Note: Individuals are employed in a different occupation if the difference between new and old occupation is larger than 0.06 according to our distance measure. The 95% Confidence Interval refers to the 2.5th and 97.5th percentiles of the distribution of effects obtained from 999 bootstrap replications.

Figure 5.B.5: Program effects for participants without switching intention: Successful occupation switch



Note: Individuals are employed in a different occupation if the difference between new and old occupation is larger than 0.06 according to our distance measure. The 95% Confidence Interval refers to the 2.5th and 97.5th percentiles of the distribution of effects obtained from 999 bootstrap replications.

Table 5.B.1: Results of probit estimations of the three covariate adjustments to the target group 'participants with intention to switch'

	Dep. Var. Training participation		Dep. Var. Training participation		Dep. Var. Intention to switch	
	Switcher, Non-participant		Non-switcher, Non-participant		Non-switcher, Participant	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Individual characteristics						
Female	-0.155**	(0.0704)	-0.269***	(0.0550)	-0.141*	(0.0846)
Age < 25	0.0683	(0.135)	-0.123	(0.105)	-0.324**	(0.149)
Age 25-29	-0.107	(0.102)	-0.144*	(0.0813)	-0.156	(0.126)
Age 30-34	-0.0582	(0.0858)	-0.127*	(0.0682)	-0.150	(0.103)
Age 40-44	-0.0331	(0.0800)	-0.0132	(0.0632)	-0.0409	(0.0939)
Age 45-49	0.0494	(0.111)	0.0751	(0.0894)	-0.00933	(0.147)
Age > 50	-0.0572	(0.133)	-0.0570	(0.106)	0.161	(0.179)
school: no degree	-0.286**	(0.115)	-0.218**	(0.100)	0.157	(0.177)
school: Hauptschule	-0.195***	(0.0669)	-0.150***	(0.0534)	-0.00944	(0.0826)
school: university entrance degree	0.0392	(0.101)	0.0164	(0.0756)	-0.00978	(0.108)
no vocational training	-0.169**	(0.0728)	-0.115*	(0.0627)	-0.0185	(0.0991)
University	0.0391	(0.127)	0.112	(0.0964)	0.0600	(0.137)
Single	-0.121*	(0.0729)	-0.108*	(0.0586)	-0.187**	(0.0887)
Couple not married	0.124	(0.119)	0.0655	(0.0924)	0.228	(0.145)
Single with child	0.139	(0.118)	0.114	(0.0925)	0.113	(0.141)
German	0.152	(0.0984)	0.112	(0.0837)	0.00304	(0.142)
At least 1 child	-0.265	(0.424)	0.179	(0.341)	0.407	(0.628)
1 child	0.0125	(0.154)	-0.00362	(0.129)	-0.0578	(0.210)
2 children	0.0707	(0.155)	0.0481	(0.130)	-0.00612	(0.211)
Job search not regionally limited	-0.0373	(0.0538)	-0.0806*	(0.0424)	-0.147**	(0.0648)
Age of youngest child < 3 years	0.350	(0.413)	-0.162	(0.331)	-0.494	(0.619)
Age of youngest child 3 to 5 years	0.265	(0.407)	-0.153	(0.324)	-0.403	(0.600)
Age of youngest child > 5 years	0.270	(0.413)	-0.170	(0.330)	-0.492	(0.610)
Health impairment	-0.179	(0.115)	-0.0513	(0.0953)	-0.0518	(0.147)
Disabled affecting placement	0.198	(0.143)	0.150	(0.119)	0.300	(0.193)
Job looked for: technical	-0.0206	(0.0804)				
Job looked for: construction	-0.318***	(0.0985)				
Job looked for: production	0.0911	(0.101)				
Job looked for: service low-skilled	-0.284***	(0.0818)				
Job looked for: other	-0.386***	(0.0918)				
Claim and unemployment benefit (UB) payments						
Unemployment benefit days at start	-0.000535	(0.000386)	-0.000382	(0.000307)	-0.000888*	(0.000482)
UB claim 0-11 months	-0.0951	(0.0854)	-0.123*	(0.0682)	-0.147	(0.111)
UB claim >12 months	-0.262**	(0.118)	-0.199**	(0.0925)	0.0798	(0.159)
Unemployment benefits	-0.00177***	(0.000423)	-0.000836***	(0.000322)	0.000657	(0.000453)
UB: 201-300 Euro (halfmonthly)	0.106	(0.0847)	0.0810	(0.0680)	0.104	(0.101)
UB: 301-400 Euro (halfmonthly)	0.308***	(0.111)	0.164*	(0.0894)	0.164	(0.129)
UB: > 401 Euro (halfmonthly)	0.371**	(0.159)	0.246*	(0.128)	-0.0624	(0.178)
Short and medium term labor market historya						
Halfmonths unemployed (1 year)	-0.104**	(0.0434)	-0.0785**	(0.0339)	-0.0901	(0.0572)
Halfmonths in program (2 years)	0.00497	(0.0101)				

Halfmonths employed (4 years)	0.00438	(0.00816)	0.0227**	(0.00965)	0.0329*	(0.0182)
Halfmonths unemployed (4 years)	-0.00178	(0.00847)	0.0188*	(0.00990)	0.0259	(0.0182)
No unemployment (4 years)	-0.234	(0.178)			-0.328	(0.219)
Halfmonths program (4 years)			-0.0104	(0.00699)		
At least 1 program (4 years)	-0.0688	(0.130)				
Halfmonths out of labor force (4 years)	0.00997	(0.00756)	0.0269***	(0.0101)	0.0328*	(0.0188)
Number unemployment spells (1 year)			0.331**	(0.141)		
Number unemployment spells (4 years)			-0.0306	(0.0670)		
Mean program duration (2 years)					0.0155	(0.0134)
Mean out of labor force duration (3 years)			-0.0135**	(0.00602)	-0.0144	(0.00963)
Mean out of labor force duration (4 years)			0.0128***	(0.00469)	0.0145*	(0.00756)
Long-term labor market historya						
Halfmonths employed (6 years)	0.00760**	(0.00329)	-0.0152***	(0.00534)	-0.0256***	(0.00928)
Halfmonths employed (10 years)	-0.00494***	(0.00187)	-0.000862	(0.000935)		
Halfmonths unemployed (6 years)	0.00920**	(0.00468)	-0.0129**	(0.00650)	-0.0247**	(0.00989)
Halfmonths unemployed (10 years)	-0.0103***	(0.00253)	-0.00395**	(0.00188)		
Halmonths program (6 years)			0.0127**	(0.00587)		
Halfmonths program (10 years)	0.0100**	(0.00510)				
Halfmonths out of labor force (6 years)			-0.0226***	(0.00491)	-0.0289***	(0.00941)
Halfmonths out of labor force (10 years)	-0.00377**	(0.00158)				
Number unemployment spells (6 years)			-0.123*	(0.0629)		
Number unemployment spells (10 years)			0.0345	(0.0338)		
Average employment duration (6 years)	-0.00274*	(0.00159)			0.00187	(0.00116)
Average employment duration (10 years)	0.00138	(0.00113)				
Average unemployment duration (6 years)	0.00197	(0.00269)	0.000423	(0.00241)	0.00554	(0.00374)
Average program duration (10 years)	-0.00822	(0.00692)	-0.00506	(0.00436)		
Halfmonths since last unemployment			-0.000690	(0.00106)		
Halfmonths since last program			0.000883	(0.000856)		
Halfmonths since last out of labor force	0.000504	(0.000403)	0.000543	(0.000418)		
Last unemployment: 12-48 months	-0.295*	(0.167)	0.0966	(0.102)	-0.357*	(0.215)
Last unemployment: > 48 months	-0.0571	(0.0901)	0.102	(0.177)	-0.0471	(0.0903)
No last program	-0.264**	(0.121)	-0.169	(0.116)		
No last out of labor force			0.145*	(0.0799)		
Employment history						
Employment duration in last firm (3 years)	-0.000189	(0.00216)	0.000539	(0.00159)	0.00162	(0.00236)
Employment duration in last occupation (3 years)	-0.000307	(0.00289)	-0.00742***	(0.00237)	-0.0116***	(0.00359)
Employed in 2 firms (5 years)	0.126	(0.0884)	-0.113*	(0.0649)	-0.188*	(0.102)
Employed in 3 or more firms (5years)	0.0256	(0.116)	-0.313***	(0.0874)	-0.547***	(0.135)
2 occupations (5 years)	-0.0861	(0.0863)	0.354***	(0.0630)	0.403***	(0.0981)
3 or more occupations (5 years)	0.0263	(0.120)	0.647***	(0.0930)	0.902***	(0.145)
Last employment						
Wage in EUR (halfmonthly)	0.000420***	(0.000109)	6.62e-05	(4.53e-05)	0.000114	(0.000132)
Unskilled worker	-0.0590	(0.0954)	0.235***	(0.0739)	0.797***	(0.113)
Skilled worker	-0.185*	(0.0982)	-0.178**	(0.0762)	0.118	(0.111)
Part-time worker	0.0580	(0.108)	0.0544	(0.0783)	0.383***	(0.119)
Industry: other	-0.194**	(0.0991)	-0.122	(0.0754)	0.154	(0.121)
Industry: Other Services (tourism, culture, public)	-0.0447	(0.0922)	-0.0917	(0.0697)	0.0367	(0.106)
Industry: Construction	-0.0912	(0.106)	-0.337***	(0.0752)	-0.197*	(0.112)
Industry: Retail	-0.0139	(0.0946)	-0.0666	(0.0718)	-0.00368	(0.107)
Industry: Services high-skilled	0.101	(0.131)	-0.0516	(0.0959)	-0.162	(0.133)

Industry: Education, health, social services	-0.0598	(0.114)	-0.0420	(0.0870)	0.402***	(0.138)
Occupation: technical	0.122	(0.0955)				
Occupation: construction	0.0600	(0.108)				
Occupation: production	0.0374	(0.112)				
Occupation: services low-skilled	0.129*	(0.0775)				
Occupaiton: other	-0.0159	(0.105)				
		Last employer				
Total number of employees	-0.000127**	(5.99e-05)	-0.000126**	(5.29e-05)	-0.000105	(7.30e-05)
Number of employees missing (firm closed)	-0.0916	(0.193)	0.0707	(0.151)	-0.216	(0.229)
Number of employees: 20-49	0.0190	(0.0810)	0.0890	(0.0637)	0.0569	(0.0982)
Number of employees: 50-199	-0.0196	(0.0775)	0.0875	(0.0625)	-0.0250	(0.0941)
Number of employees: >199	0.0807	(0.0973)	0.162**	(0.0802)	0.138	(0.118)
Age of firm in half-months	0.000119	(0.000142)	-6.59e-05	(0.000112)	-0.000208	(0.000172)
Share of unskilled workers	0.0467	(0.170)	0.132	(0.138)	0.124	(0.212)
Share of skilled workers	0.173	(0.170)	0.235*	(0.134)	0.241	(0.199)
Share of part-time workers	-0.231	(0.207)	0.0116	(0.160)	0.156	(0.250)
Average age in firm	-0.00140	(0.00300)	-0.000482	(0.00220)	-0.00444	(0.00354)
Average wage full-time employees in EUR	-0.000221	(0.00239)	-8.72e-05	(0.00177)	0.00117	(0.00275)
		Entry timing into unemployment and program				
Beginning of unemployment	-0.00431	(0.00500)	0.000530	(0.00399)	0.00100	(0.00641)
Beginning of unemployment: 2000	-0.000413	(0.232)	0.249	(0.187)	0.176	(0.292)
Beginning of unemployment: 2001	0.113	(0.137)	0.176	(0.110)	0.0990	(0.168)
Beginning of unemployment: 2003	-0.0396	(0.148)	-0.202*	(0.118)	-0.222	(0.186)
Beginning of unemployment: 2004	0.0650	(0.249)	-0.209	(0.200)	-0.106	(0.321)
Beginning of unemployment: Dec-Feb	-0.103	(0.0742)	-0.0703	(0.0594)	-0.0523	(0.0909)
Beginning of unemployment: Jun-Aug	0.0277	(0.0772)	-0.000977	(0.0613)	0.0302	(0.0937)
Beginning of unemployment: Sep-Nov	-0.123	(0.0949)	-0.141*	(0.0764)	-0.0950	(0.117)
Time to program: 0-1 half-months	-0.401***	(0.116)	-0.296***	(0.0938)	-0.181	(0.153)
Time to program: 2 half-months	-0.113	(0.0874)	-0.0721	(0.0690)	0.0839	(0.114)
Time to program: 7-12 half-months	0.140**	(0.0671)	0.166***	(0.0538)	-0.0724	(0.0843)
Time to program: at least 12 half-months	0.342***	(0.0699)	0.351***	(0.0561)	-0.0698	(0.0851)
		Region dummies				
Baden-Wurttemberg	0.113	(0.119)	0.0398	(0.0955)	0.162	(0.150)
Bavaria	0.0492	(0.117)	0.00571	(0.0931)	0.0231	(0.143)
Lower-Saxony, Bremen	0.339***	(0.106)	0.255***	(0.0833)	0.331**	(0.132)
Schleswig-Holstein, Hamburg	0.358***	(0.134)	0.318***	(0.105)	0.353**	(0.161)
Hessen	-0.0537	(0.133)	0.00629	(0.108)	-0.0787	(0.167)
Rhineland-Palatinate, Saarland	0.245*	(0.127)	0.255**	(0.100)	0.234	(0.153)
Brandenburg	0.178	(0.185)	0.170	(0.148)	0.257	(0.218)
Mecklenburg-West Pomerania	0.281	(0.216)	0.226	(0.174)	0.510**	(0.254)
Saxony-Anhalt	0.469**	(0.195)	0.475***	(0.156)	0.780***	(0.239)
Saxony	0.0457	(0.175)	0.0853	(0.141)	0.420*	(0.215)
Thuringia	0.532***	(0.170)	0.457***	(0.135)	0.784***	(0.209)
		Regional characteristics				
GDP per capita in 10.000 EUR	-0.0172	(0.0466)	-0.0132	(0.0371)	0.00305	(0.0583)
Average time to next big city by public transport	0.00113*	(0.000637)	0.000928*	(0.000512)	0.000420	(0.000805)
Share of foreigners	0.473	(1.138)	0.485	(0.902)	0.875	(1.379)
Unemployment rate	-1.182	(1.215)	-0.820	(0.949)	-1.056	(1.452)
Rural area	-0.259*	(0.157)	-0.0880	(0.126)	-0.260	(0.196)
Number of small cities	0.00758	(0.0125)	0.00264	(0.00980)	0.0168	(0.0153)

Number of big cities	-0.0110	(0.0659)	-0.0323	(0.0535)	-0.0261	(0.0836)
Net migration	-0.0427	(0.240)	-0.0170	(0.187)	0.302	(0.290)
Labor market tightness	0.426	(0.372)	-0.0860	(0.298)	-0.319	(0.440)
Labor market tightness missing	0.0689	(0.134)	0.0686	(0.106)	0.000341	(0.155)
Constant	-0.249	(0.787)	-0.959	(0.651)	0.0869	(1.181)
Observations	6,455		25,161		2,800	
Pseudo-R2	0.138		0.141		0.155	

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

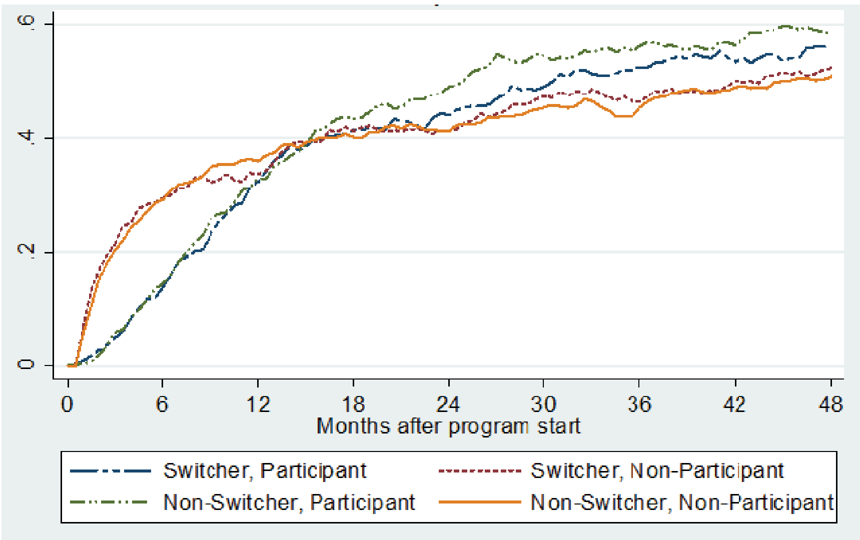
Note: a) The years in brackets refer to the time period before entering unemployment during which the respective variable is measured.

C Sensitivity Analysis

The following figures and tables present the results from our sensitivity analysis with respect to a different threshold of the skill distance measure. If last occupation and target occupation are further apart than 0.054 (according to the skill distance measure used in this study) they are classified as intended occupation switches.

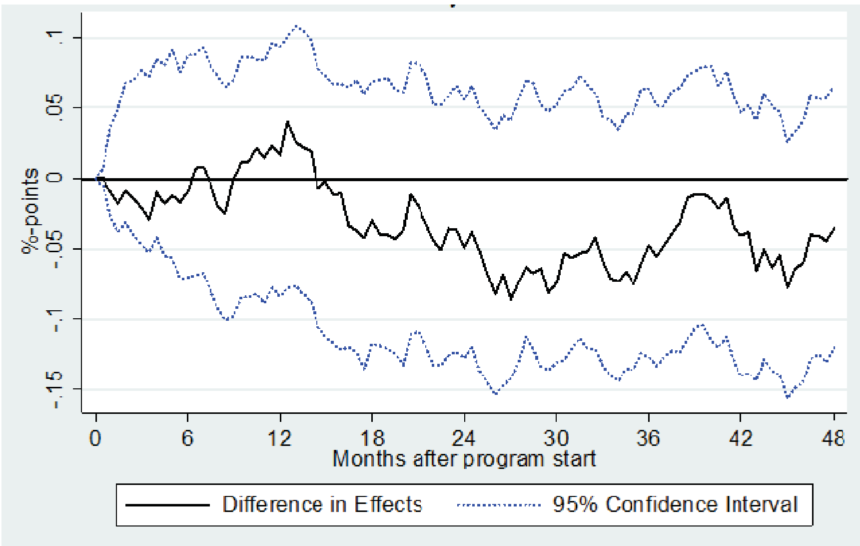
Variation of the occupation distance measure

Figure 5.C.1: Employment probabilities



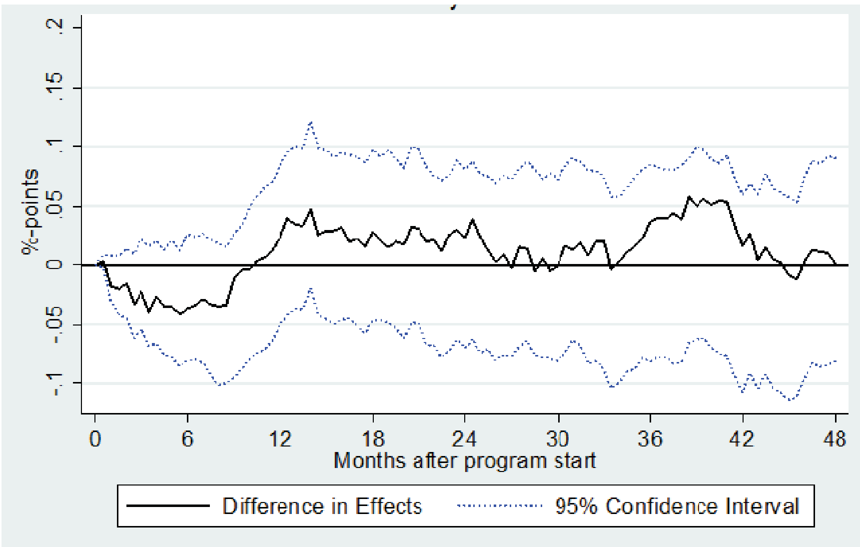
Note: Ordinate: Employment probability in %. Employment probabilities are obtained considering unsubsidized employment only. Skill distances between last and target occupation above 0.054 are defined as switch intention.

Figure 5.C.2: Difference in program effects between participants with and without switching intention



Note: The presented results are obtained by subtracting the effect of non-switcher from the one of switcher. Thus, a positive value of the difference indicates effect heterogeneity in favor of switcher. The 95% Confidence Interval refers to the 2.5th and 97.5th percentiles of the distribution of effects obtained from 999 bootstrap replications.

Figure 5.C.3: Difference in program effects (successful switching) between participants with switching intention and those without



Note: The presented results are obtained by subtracting the effect of non-switcher from the one of switcher. Thus, a positive value of the difference indicates effect heterogeneity in favor of switcher. The 95% Confidence Interval refers to the 2.5th and 97.5th percentiles of the distribution of effects obtained from 999 bootstrap replications.

Table 5.C.1: Most frequent intended occupation switches in the sample

Last occupation	Target occupation	cases	(%) of all switchers	share participants (%)
Office Clerk	Accountant, Bookkeeper	202	2.81	20.79
Unskilled Construction Worker	Bricklayer, Mason	157	2.18	3.82
Accountant, Bookkeeper	Office Clerk	120	1.67	16.67
Sales Personnel	Accountant, Bookkeeper	86	1.20	12.79
Truck Driver, Conductor	Storekeeper, Warehouse Keeper	84	1.17	8.33
Accountant, Bookkeeper	Sales Personnel	79	1.10	15.19
Storekeeper, Warehouse Keeper	Truck Driver, Conductor	75	1.04	9.33
Technical Service Personnel	Engineer	60	0.83	40.00
Unskilled Construction Worker	Plasterer	59	0.82	3.39
Office Clerk	Engineer	56	0.78	25
Bricklayer, Mason	Unskilled Construction Worker	53	0.74	3.77
Waiter, Barkeeper, Innkeeper	Sales Personnel	52	0.72	7.69
Cleaning Service Worker	Janitor, Housekeeper	52	0.72	1.92
Cleaning Service Worker	Sales Personnel	50	0.7	6
Storekeeper, Warehouse Keeper	Sales Personnel	49	0.68	20.41
Assembler	Product/ Quality Inspector	45	0.63	6.67
Storekeeper, Warehouse Keeper	Office Clerk	42	0.58	14.29
Technical Service Personnel	Office Clerk	40	0.56	20
Engineer	Technical Service Personnel	39	0.54	25.64
Office Clerk	Storekeeper, Warehouse Keeper	39	0.54	7.69
Janitor, Housekeeper	Cleaning Service Worker	36	0.5	11.11

Note: The total number of those with an intention to switch occupation in the sample according to the adjusted distance measure (0.054) is 7189 (before: 6455).

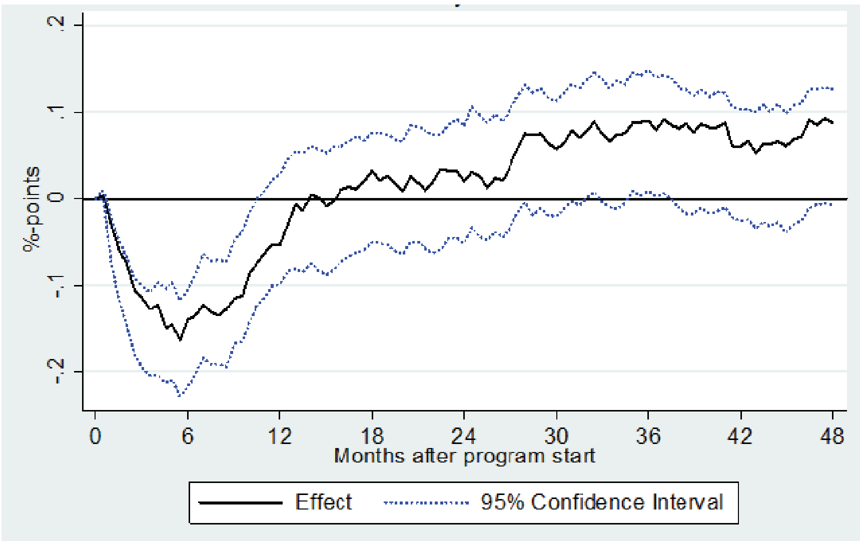
Table 5.C.2: Program Effects of FVT on cumulated labor market outcomes (sensitivity: distance measure)

	With intention to switch			Percentile		Without intention to switch			Percentile		Difference	Percentile	
Outcome	Participants (1)	Non-participants (2)	Effect (3)	2.5th	97.5th	Participants (4)	Non-participants (5)	Effect (6)	2.5th	97.5th	(3)-(6)	2.5th	97.5th
Halfmonths employed													
1 year after	3.64	6.32	**-2.68**	**-3.33**	**-1.55**	3.79	6.37	**-2.58**	**-3.52**	**-1.57**	-0.1	-1.17	1.45
4 years after	38.31	39.01	-0.7	-4.42	4.05	40.87	38.48	2.39	-2.72	6.58	-3.09	-8.2	4.13
4th year after	13.06	11.95	1.11	-0.17	2.46	13.82	11.75	**2.07**	**0.52**	**3.52**	-0.97	-2.76	1.12
3rd and 4th year after	24.91	23.01	1.89	-0.52	4.64	26.8	22.42	**4.38**	**1.09**	**6.62**	-2.49	-5.4	1.99
Unemployed 3rd and 4th year after	19.43	21.34	-1.92	-4.45	0.64	17.38	21.62	**-4.24**	**-6.57**	**-1.01**	2.32	-1.72	5.64
Halfmonths employed 3rd and 4th year													
In last occupation	4.3	7.23	**-2.93**	**-4.25**	**-0.8**	10.65	11.72	-1.08	-4	0.98	-1.85	-4.14	2.03
In intended occupation	6.18	5.47	0.71	-0.71	2.83	10.96	10.69	0.27	-3.34	1.14	0.43	-0.41	5.06
In other occupation	18.09	13.98	**4.12**	**0.55**	**5.55**	18.17	13.6	**4.57**	**2.44**	**7.81**	-0.45	-5.77	1.7
Wage earnings													
1 year after	3146	5977	-2832	-3535	-1426	3552	5678	**-2126**	**-3234**	**-1131**	-706	-1773	1219
4 years after	32462	33536	-1073	-5138	4896	36344	32889	3455	-1497	9899	-4528	-11603	3433
4th year after	10938	9813	1125	-234	2895	11869	9746	**2123**	**823**	**4338**	-998	-3564	1115
3rd and 4th year after	20990	19076	1915	-782	5252	23313	18814	**4499**	**1728**	**8380**	-2584	-7166	1826

Note: The reported 2.5th and 97.5th percentiles are based on 999 bootstrap replications of the effects. Numbers in bold indicate significance at a 95% level.

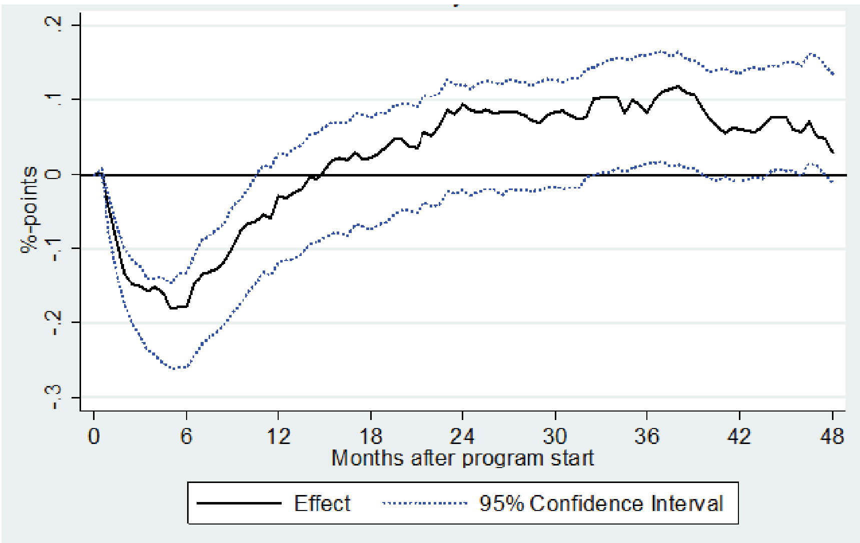
Variation in the training program participation definition – ‘18 months window’

Figure 5.C.4: Program effects for participants with switching intention: Employment probability in unsubsidized employment



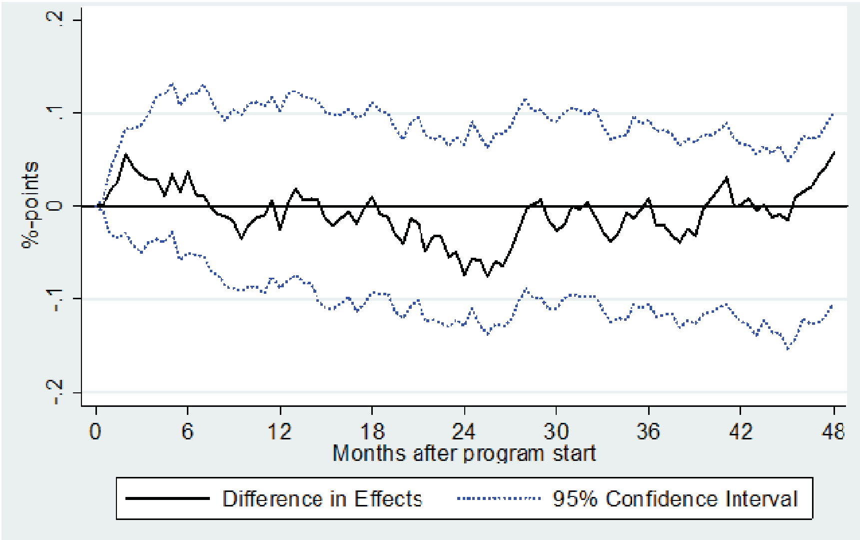
Note: The 95% Confidence Interval refers to the 2.5th and 97.5th percentiles of the distribution of effects obtained from 999 bootstrap replications.

Figure 5.C.5: Program effects for participants without switching intention: Employment probability in unsubsidized employment



Note: The 95% Confidence Interval refers to the 2.5th and 97.5th percentiles of the distribution of effects obtained from 999 bootstrap replications.

Figure 5.C.6: Difference in program effects between participants with switching intention and those without



Note: The presented results are obtained by subtracting the effect of non-switcher from the one of switcher. Thus, a positive value of the difference indicates effect heterogeneity in favor of switcher. The 95% Confidence Interval refers to the 2.5th and 97.5th percentiles of the distribution of effects obtained from 999 bootstrap replications.

Table 5.C.3: Selected mean characteristics by sample subgroups

	Non-participants		Participants	
Intention to switch	No	Yes	No	Yes
Observations	20937	4860	2506	569
Baseline individual characteristics				
Age	38.22	39.30	37.75	38.39
School: no degree	0.08	0.13	0.03	0.06
School: Hauptschule	0.45	0.49	0.30	0.35
School: upper secondary degree	0.33	0.28	0.45	0.41
School: university entrance degree	0.14	0.10	0.22	0.18
No vocational training	0.23	0.37	0.12	0.21
Vocational training	0.71	0.58	0.78	0.70
University	0.06	0.05	0.10	0.09
Female	0.42	0.40	0.47	0.37
At least 1 child	0.32	0.36	0.35	0.40
Single	0.37	0.33	0.38	0.30
Couple not married	0.05	0.05	0.04	0.05
Single with child	0.04	0.05	0.06	0.07
Married	0.54	0.57	0.52	0.58
Timing of unemployment and program start				
Unempl start: in half-months (1=Jan 2000)	61.18	60.27	52.40	47.79
Unempl. Duration until program start: half-months	6.89	7.49	10.64	11.42
Last employment: non-firm characteristics				
Unskilled worker	0.21	0.42	0.12	0.33
Skilled worker	0.34	0.22	0.24	0.21
White-collar worker	0.30	0.20	0.51	0.31
Part-time worker	0.16	0.16	0.13	0.15
Last employment: firm characteristics				
Age of firm in years	12.93	12.58	12.22	12.02
Industry: Manufacturing	0.19	0.25	0.23	0.27
Industry: Construction	0.22	0.15	0.14	0.12
Industry: Retail	0.15	0.13	0.18	0.17
Industry: Service higher skilled	0.06	0.04	0.14	0.07
Industry: Health and social services	0.10	0.10	0.07	0.11
Industry: Service lower skilled, tourism	0.17	0.21	0.15	0.17
Industry: Primary sector, other services	0.10	0.13	0.08	0.11
Short-term labor market history (last 4 years before entering unemployment)				
Halfmonths employed	77.88	71.61	80.59	74.70
Halfmonths unemployed	8.54	11.50	6.89	9.20
Halfmonths out of labor force	8.15	11.00	7.12	10.18
Long-term labor market history (last 10 years before entering unemployment)				
Average employment duration	67.39	60.58	72.27	67.50
Number unemployment spells	1.46	1.48	1.13	1.17
Average unemployment duration	10.30	14.04	10.38	12.99
Unemployment benefits and claim				
Amount of unempl benefit per 2 weeks in EUR	332.10	315.45	322.55	307.42
UB claim <12 months	0.27	0.32	0.23	0.28
UB claim 12 months	0.44	0.40	0.54	0.51
UB claim >12 months	0.29	0.28	0.24	0.21
Additional information				
Halfmonths last occupation in last 5 years	85.72	69.85	87.05	70.79
1 occupation (2-digit) in last 5 years	0.67	0.47	0.64	0.47
2 occupations (2-digit) in last 5 years	0.26	0.39	0.30	0.39
3 or more occupations (2-digit) in last 5 years	0.07	0.14	0.07	0.14
Last employment: monthly wage in EUR	1800	1610	1914	1773
Cumulated earnings in last year	19924	18201	21630	20314
Labor market tightness	0.09	0.08	0.09	0.08

Note: All entries are mean values measured at the time of entering unemployment except for the variable 'unemployment duration until program start'

Table 5.C.4: Program Effects of FVT on cumulated labor market outcomes (18 months participation window)

	With intention to switch			Percentile		Without intention to switch			Percentile		Difference	Percentile	
Outcome	Participants (1)	Non-participants (2)	Effect (3)	2.5th	97.5th	Participants (4)	Non-participants (5)	Effect (6)	2.5th	97.5th	(3)-(6)	2.5th	97.5th
Halfmonths employed													
1 year after	3.68	6.13	-2.45	-3.6	-1.64	3.79	6.47	**-2.68**	**-4.18**	**-2.02**	0.23	-0.87	2.02
4 years after	38.25	37.09	1.17	-4.59	4.49	39.55	37.66	1.88	-4.15	5.03	-0.72	-7.03	5.87
4th year after	12.98	11.14	1.84	-0.19	2.69	13.25	11.46	**1.8**	**0.23**	**3.43**	0.04	-2.66	1.43
3rd and 4th year after	24.94	21.6	3.33	-0.12	5.21	25.72	21.87	**3.85**	**0.44**	**6.3**	-0.51	-4.67	3.19
Unemployed 3rd and 4th year after	18.96	22.16	-3.2	-5.24	0.22	18.12	21.7	**-3.59**	**-6.32**	**-0.49**	0.39	-2.98	4.84
Halfmonths employed 3rd and 4th year													
In last occupation	4.13	5.7	**-1.57**	**-4.35**	**-0.6**	9.44	10.55	-1.11	-4.4	0.35	-0.45	-3.52	2.6
In intended occupation	5.52	4.33	1.2	-0.4	3.03	9.24	10.49	-1.25	-4.13	0.53	**2.45**	**0.25**	**6.07**
In other occupation	19.39	15.89	**3.49**	**0.89**	**6.15**	18.71	14.43	**4.29**	**2.39**	**8.12**	-0.8	-5.77	2.42
Wage earnings													
1 year after	3024	4903	-1879	-3307	-1320	3209	5543	**-2334**	**-3693**	**-1587**	455	-1151	1844
4 years after	30677	27548	3129	-3962	5625	33820	30421	3399	-3499	6988	-269	-8266	6079
4th year after	10190	7942	**2248**	**53**	**3026**	11047	8835	**2212**	**248**	**3665**	36	-2806	1712
3rd and 4th year after	19849	15639	**4210**	**115**	**5936**	21870	17076	**4794**	**523**	**7059**	-584	-5235	3290

Note: The reported 2.5th and 97.5th percentiles are based on 999 bootstrap replications of the effects. Numbers in bold indicate significance at a 95% level.

Part VI

References

ABDULAI, A. (2010): "Socio-economic characteristics and obesity in underdeveloped economies: does income really matter?" *Applied Economics*, 42, 157–169.

AKEE, R., E. SIMEONOVA, W. COPELAND, A. ANGOLD, AND J. COSTELLO (2010): "Does More Money Make You Fat? The Effects of Quasi-Experimental Income Transfers on Adolescent and Young Adult Obesity," *IZA Discussion Paper No. 5135.*

ASFAW, A. (2007): "Do Government Food Price Policies Affect the Prevalence of Obesity? Empirical Evidence from Egypt," *World Development*, 35, 687–701.

BA (2000): "Arbeitsstatistik - Jahreszahlen 2000, 2001, 2002, 2003, 2004, 2005," *Amtliche Nachrichten der Bundesanstalt für Arbeit*, 49-54.

BALL, K. AND D. CRAWFORD (2005): "Socioeconomic status and weight change in adults: a review," *Social Science & Medicine*, 60, 1987–2010.

BAUM, C. (2009): "The effects of cigarette costs on BMI and obesity," *Health economics*, 18, 3–19.

BAUM II, C. AND C. RUHM (2009): "Age, socioeconomic status and obesity growth," *Journal of Health Economics*, 28, 635–648.

BECKER, G. S. (1962): "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, 70, 9–49.

BHALLA, A. S., S. YAO, AND Z. ZHANG (2003): "Causes of Inequalities in China, 1952 to 1999," *Journal of International Development*, 15, 939–955.

BLEICH, S., D. CUTLER, C. MURRAY, AND A. ADAMS (2008): "Why is the developed world obese?" *Annual Review of Public Health*, 29, 273–295.

BRAMALL, C. (2009): *Chinese Economic Development*, Taylor & Francis.

CABALLERO, B. (2007): "The global epidemic of obesity: An overview," *Epidemiologic Reviews*, 29, 1–5.

CARD, D., J. KLUVE, AND A. WEBER (2010): "Active Labour Market Policy Evaluations: A Meta-Analysis," *The Economic Journal*, 120, 452–477.

CASE, A. AND A. MENENDEZ (2009): "Sex differences in obesity rates in poor countries: Evidence from South Africa," *Economics & Human Biology*, 7, 271–282.

Cawley, J. (2004): "The impact of obesity on wages," *Journal of Human Resources*, 39, 451–474.

Cawley, J., J. Moran, and K. Simon (2010): "The impact of income on the weight of elderly Americans," *Health Economics*, 19, 979–993.

Chou, S., M. Grossman, and H. Saffer (2004): "An economic analysis of adult obesity: results from the Behavioral Risk Factor Surveillance System," *Journal of Health Economics*, 23, 565–587.

Chou, S., I. Rashad, and M. Grossman (2008): "Fast-Food Restaurant Advertising on Television and Its Influence on Childhood Obesity," *Journal of Law and Economics*, 51, 599–618.

Currie, J., S. DellaVigna, E. Moretti, and V. Pathania (2010): "The effect of fast food restaurants on obesity and weight gain," *American Economic Journal: Economic Policy*, 2, 32–63.

Cutler, D., E. Glaeser, and J. Shapiro (2003): "Why have Americans become more obese?" *Journal of Economic Perspectives*, 93–118.

Deaton, A. (2003): "Health, Inequality, and Economic Development," *Journal of Economic Literature*, 41, pp. 113–158.

Dehejia, R. H. and S. Wahba (2002): "Propensity Score Matching Methods for Nonexperimental Causal Studies," *The Review of Economics and Statistics*, 84, 151–161.

Destatis (1992): "Die Klassifizierung der Berufe des Statistischen Bundesamtes in der Fassung für den Mikrozensus - Ausgabe 1992," *Statistisches Bundesamt Deutschland*.

DeVellis, R. (2003): *Scale Development: Theory and Applications*, Sage Publications, Inc.

Doak, C. and B. Popkin (2008): "The rapid emergence of obesity in developing countries," *Nutrition and Health in Developing Countries*, 617–638.

Du, S., T. Mroz, F. Zhai, and B. Popkin (2004): "Rapid income growth adversely affects diet quality in China–particularly for the poor!" *Social Science & Medicine*, 59, 1505–1515.

Eide, E. and M. Showalter (2001): "The effect of grade retention on educational and labor market outcomes," *Economics of Education Review*, 20, 563–576.

Etil, F. (2007): "Social norms, ideal body weight and food attitudes," *Health Economics*, 16, 945–966.

Fernald, L. (2007): "Socio-economic status and body mass index in low-income Mexican adults," *Social Science & Medicine*, 64, 2030–2042.

Filmer, D. and L. Pritchett (2001): "Estimating Wealth Effects Without Expenditure Data-Or Tears: An Application To Educational Enrollments In States Of India," *Demography*, 38, 115–132.

Fitzenberger, B., A. Osikominu, and R. Völter (2008): "Get Training or Wait? Long-Run Employment Effects of Training Programs for the Unemployed in West Germany," *Annals of Economics and Statistics*, 321–355.

Fitzenberger, B. and S. Speckesser (2006): "Employment effects of the provision of specific professional skills and techniques in Germany," *Empirical Economics*, 32, 529–573.

Fredriksson, P. and P. Johansson (2003): "Program evaluation and random program starts," *IFAU Discussion Paper*.

Galal, O. M. (2002): "The nutrition transition in Egypt: Obesity, undernutrition and the food consumption context," *Public Health Nutrition*, 5, 141–148.

Gathmann, C. and U. Schönberg (2010): "How General Is Human Capital? A Task Based Approach," *Journal of Labor Economics*, 28, 1–49.

Gortmaker, S., B. Swinburn, D. Levy, R. Carter, P. Mabry, D. Finegood, T. Huang, T. Marsh, and M. Moodie (2011): "Changing the future of obesity: science, policy, and action," *The Lancet*, 378, 838–847.

Greene, J. and M. Winters (2007): "Revisiting grade retention: An evaluation of Florida's test-based promotion policy," *Education Finance and Policy*, 2, 319–340.

Grimm, M. (2011): "Does inequality in health impede economic growth?" *Oxford Economic Papers*, 63, 448–474.

Gruber, J. and M. Frakes (2006): "Does falling smoking lead to rising obesity?" *Journal of Health Economics*, 25, 183–197.

Hao, R. and Z. Wei (2010): "Fundamental causes of inland–coastal income inequality in post-reform China," *The annals of regional science*, 45, 181–206.

Heckman, J. J., R. J. LaLonde, and J. A. Smith (1999): "Chapter 31 The economics and econometrics of active labor market programs," in *Handbook of Labor Economics*, ed. by O. C. Ashenfelter and D. Card, Elsevier, vol. 3, Part A of *Handbook of Labor Economics*, 1865–2097.

Heilig, G. (2006): "Many Chinas? The Economic Diversity of China's Provinces," *Population and Development Review*, 32, 147–161.

Holz, C. (2008): "China's economic growth 1978-2025: What we know today about China's economic growth tomorrow," *World Development*, 36, 1665–1691.

Hotz, V. J., G. W. Imbens, and J. A. Klerman (2006): "Evaluating the Differential Effects of Alternative Welfare-to-Work Training Components: A Reanalysis of the California GAIN Program," *Journal of Labor Economics*, 24, 521–566.

Huber, M., M. Lechner, and A. Steinmayr (2012): "Radius matching on the propensity score with bias adjustment: finite sample behaviour, tuning parameters and software implementation," *Mimeo*.

Huber, M., M. Lechner, and C. Wunsch (2013): "The performance of estimators based on the propensity score," *Journal of Econometrics*, 175, 1 – 21.

Imbens, G. and T. Lemieux (2008): "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, 142, 615–635.

Imbens, G. W. (2004): "Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review," *The Review of Economics and Statistics*, 86, 4–29.

Jacob, B. and L. Lefgren (2004): "Remedial education and student achievement: A regression-discontinuity analysis," *Review of economics and statistics*, 86, 226–244.

Jacob, B. A. and L. Lefgren (2009): "The Effect of Grade Retention on High School Completion," *American Economic Journal: Applied Economics*, 1, 33–58.

Jacobson, L. S., R. J. LaLonde, and D. G. Sullivan (1993): "Earnings Losses of Displaced Workers," *The American Economic Review*, 83, 685–709.

Jensen, R. T. and N. H. Miller (2008): "Giffen behavior and subsistence consumption," *The American economic review*, 98, 1553.

Jimerson, S. (2001): "Meta-analysis of grade retention research: Implications for practice in the 21st century," *School Psychology Review*, 30, 420–437.

Jisheng, Y. (2010): "The fatal politics of the PRC's great leap famine: The preface to tombstone," *Journal of Contemporary China*, 19, 755–776.

Jolliffe, D. (2010): "Overweight and poor? On the relationship between income and the body mass index," *IZA Discussion Paper No. 5366*.

Jones-Smith, J. and B. Popkin (2010): "Understanding community context and adult health changes in China: Development of an urbanicity scale," *Social Science & Medicine*, 71, 1436–1446.

Kakwani, N., A. Wagstaff, and E. Van Doorslaer (1997): "Socioeconomic inequalities in health: Measurement, computation, and statistical inference," *Journal of Econometrics*, 77, 87–103.

Kambourov, G. and I. Manovskii (2009): "Occupational specificity of human capital," *International Economic Review*, 50, 63–115.

Kambourov, G., I. Manovskii, and M. Plesca (2012): "Occupational Mobility and the Returns to Training," *Working Paper*.

Kan, K. and W. Tsai (2004): "Obesity and risk knowledge," *Journal of Health Economics*, 23, 907–934.

Kanbur, R. and X. Zhang (2005): "Fifty years of regional inequality in China: a journey through central planning, reform, and openness," *Review of Development Economics*, 9, 87–106.

Kletzer, L. G. (1998): "Job Displacement," *Journal of Economic Perspectives*, 12, 115–136.

Kluve, J. (2010): "The effectiveness of European active labor market programs," *Labour Economics*, 17, 904–918.

Kluve, J. and C. M. Schmidt (2002): "Can training and employment subsidies combat European unemployment?" *Economic Policy*, 17, 409–448.

——— (2009): "The growth of obesity and technological change," *Economics & Human Biology*, 7, 283–293.

Lakdawalla, D., T. Philipson, and J. Bhattacharya (2005): "Welfare-enhancing technological change and the growth of obesity," *American Economic Review*, 95, 253–257.

Lazear, E. P. (2009): "Firm-Specific Human Capital : A Skill-Weights Approach Edward P," *Journal of Political Economy*, 117, 914–941.

Lechner, M. (1999): "Earnings and Employment Effects of Continuous Off-the-Job Training in East Germany after Unification," *Journal of Business & Economic Statistics*, 17, 74–90.

Lechner, M., R. Miquel, and C. Wunsch (2011): "Long-Run Effects of Public Sector Sponsored Training in West Germany," *Journal of the European Economic Association*, 9, 742–784.

Lechner, M. and C. Wunsch (2009): "Are Training Programs More Effective When Unemployment Is High ?" *Journal of Labor Economics*, 27, 653–692.

——— (2011): "Sensitivity of matching-based program evaluations to the availability of control variables," *Discussion Paper*.

Lee, D. S. and T. Lemieux (2010): "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 48, 281–355.

Levine, J. (2008): "Obesity in China: Causes and solutions," *Chinese Medical Journal*, 121, 1043–1050.

Liu, H. (2008): "The China health and nutrition survey: an important database for poverty and inequality research," *Journal of Economic Inequality*, 6, 373–376.

Lorence, J. and A. Dworkin (2006): "Elementary grade retention in Texas and reading achievement among racial groups: 1994–2002," *Review of Policy Research*, 23, 999–1033.

Lu, Y. and D. Goldman (2010): "The Effects of Relative Food Prices on Obesity – Evidence from China: 1991-2006," *NBER Working Paper No. 15720.*

Mammen, K. and C. Paxson (2000): "Women's Work and Economic Development," *The Journal of Economic Perspectives*, 14, pp. 141–164.

Manacorda, M. (2012): "The Cost of Grade Retention," *Review of Economics and Statistics*, 94, 596–606.

Manski, C. F. (1989): "Schooling as experimentation: a reappraisal of the postsecondary dropout phenomenon," *Economics of Education Review*, 8, 305–312.

Martorell, P. and I. McFarlin Jr (2011): "Help or hindrance? The effects of college remediation on academic and labor market outcomes," *The Review of Economics and Statistics*, 93, 436–454.

McCrary, J. (2008): "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 142, 698–714.

McLaren, L. (2007): "Socioeconomic status and obesity," *Epidemiologic reviews*, 29, 29–48.

Monteiro, C., W. Conde, B. Lu, and B. Popkin (2004a): "Obesity and inequities in health in the developing world," *International Journal of Obesity*, 28, 1181–1186.

Monteiro, C., E. Moura, W. Conde, and B. Popkin (2004b): "Socioeconomic status and obesity in adult populations of developing countries: A review," *Bulletin of the World Health Organization*, 82, 940–946.

Neal, D. (1995): "Industry-Specific Human Capital : Evidence from Displaced Workers," *Journal of Labor Economics*, 13, 653–677.

Ng, S., E. Norton, D. Guilkey, and B. Popkin (2012): "Estimation of a dynamic model of weight," *Empirical Economics*, 1–31, 10.1007/s00181-012-0547-7.

Ng, S., E. Norton, and B. Popkin (2009): "Why have physical activity levels declined among Chinese adults? Findings from the 1991-2006 China health and nutrition surveys," *Social Science & Medicine*, 68, 1305–1314.

Oberschachtsiek, D., P. Scioch, C. Seysen, and J. Heining (2009): "Stichprobe der Integrierten Erwerbsbiografien IEBS * Handbuch für die IEBS in der Fassung 2008," *FDZ Datenreport*, 2009.

OECD (2004): "OECD Employment Outlook 2004," Tech. rep., OECD, Paris.

——— (2008): *Higher Education to 2030, Volume 1, Demography*, OECD Publishing.

Ou, D. (2010): "To leave or not to leave? A regression discontinuity analysis of the impact of failing the high school exit exam," *Economics of Education Review*, 29, 171–186, special Issue in Honor of Henry M. Levin.

Parent, D. (2000): "Industry-Specific Capital and the Wage Profile: Evidence from the National Longitudinal Survey of Youth and the Panel Study of Income Dynamics," *Journal of Labor Economics*, 18, 306–23.

Pereira, M. (2006): "The possible role of sugar-sweetened beverages in obesity etiology: A review of the evidence," *International Journal of Obesity*, 30, 28–36.

Philipson, T. and R. Posner (2003): "The long-run growth in obesity as a function of technological change," *Perspectives in Biology and Medicine*, 46, 87.

——— (2008): "Is the obesity epidemic a public health problem? A decade of research on the economics of obesity," *NBER Working Paper No. 14010.*

Podder, N. and B. Tran-Nam (1994): "A new approach to estimating Engel elasticities from concentration curves," *Oxford Economic Papers*, 262–276.

Popkin, B. (1994): "The nutrition transition in low-income countries: An emerging crisis," *Nutrition Reviews*, 52, 285–298.

——— (1998): "The nutrition transition and its health implications in lower-income countries," *Public Health Nutrition*, 1, 5–21.

——— (1999): "Urbanization, lifestyle changes and the nutrition transition," *World Development*, 27, 1905–1916.

——— (2008): “Will China’s nutrition transition overwhelm its health care system and slow economic growth?” *Health Affairs*, 27, 1064–1076.

Popkin, B., S. Kim, E. Rusev, S. Du, and C. Zizza (2006): “Measuring the full economic costs of diet, physical activity and obesity-related chronic diseases,” *Obesity Reviews*, 7, 271–293.

Pyatt, G. (2003): “Development and the distribution of living standards: a critique of the evolving data base,” *Review of Income and Wealth*, 49, 333–358.

Robinson, C. (2010): “Occupational Mobility , Occupation Distance and Basic Skills: Evidence from Job-Based Skill Measures,” *Working Paper*.

Roderick, M. and J. Nagaoka (2005): “Retention under Chicagos high-stakes testing program: Helpful, harmful, or harmless?” *Educational Evaluation and Policy Analysis*, 27, 309.

Rosenbaum, P. R. and D. B. Rubin (1983): “The central role of the propensity score in observational studies for causal effects,” *Biometrika*, 70, 41–55.

Rosin, O. (2008): “The economic causes of obesity: A survey,” *Journal of Economic Surveys*, 22, 617–647.

Rubin, D. B. (1974): “Estimating causal effects of treatments in randomized and nonrandomized studies.” *Journal of Educational Psychology*, 66, 688–701.

Rübner, M. and B. Sprengard (2011): “Beratungskonzeption der Bundesagentur für Arbeit - Grundlagen,” *Bundesagentur für Arbeit*.

Rutstein, S. (2008): “The DHS wealth index: approaches for rural and urban areas,” *Washington, DC: Macro International Inc.*

Sahn, D. (2009): “Weights on the rise: where and for whom?” *Journal of Economic Inequality*, 7, 351–370.

Sahn, D. and S. Younger (2009): “Measuring intra-household inequality: Explorations using the body mass index,” *Health Economics*, 18, 13–36.

Sarlio-Lahteenkorva, S., K. Silventoinen, and E. Lahelma (2004): “Relative weight and income at different levels of socioeconomic status,” *American Journal of Public Health*, 94, 468.

Schneider, H., K. Brenke, L. Kaiser, J. Steinwede, B. Jesske, and A. Uhlendorff (2006): “Evaluation der Maßnahmen zur Umsetzung der Vorschläge der Hartz-Kommission - Modul 1b,” *IZA Research Report*.

Schroeter, C., J. Lusk, and W. Tyner (2008): “Determining the impact of food price and income changes on body weight,” *Journal of Health Economics*, 27, 45–68.

Shimokawa, S. (2008): “The labour market impact of body weight in China: A semiparametric analysis,” *Applied Economics*, 40, 949–968.

Sianesi, B. (2004): “An Evaluation of the Swedisch System of Active Labor Market Programs in the 1990s,” *The Review of Economics and Statistics*, 86, 133–155.

Sobal, J. and A. J. Stunkard (1989): “Socioeconomic status and obesity: A review of the literature.” *Psychological Bulletin*, 105, 260–275.

Tafreschi, D. (2011): “The Income Body Weight Gradients in the Developing Economy of China,” *Economics Working Paper Series*.

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

Van de Poel, E., O. O’Donnell, and E. Van Doorslaer (2009): “Urbanization and the spread of diseases of affluence in China,” *Economics & Human Biology*, 7, 200–216.

——— (2012): “Is there a health penalty of China’s rapid urbanization?” *Health Economics*, 21, 367–385.

Van Ours, J. C. (2004): “The locking-in effect of subsidized jobs,” *Journal of Comparative Economics*, 32, 37–55.

——— (2011a): “The concentration index of a binary outcome revisited,” *Health Economics*, 20, 1155–1160.

Wagstaff, A., P. Paci, and E. Van Doorslaer (1991): “On the measurement of inequalities in health,” *Social science & medicine*, 33, 545–557.

Wagstaff, A., E. Van Doorslaer, and N. Watanabe (2003): “On decomposing the causes of health sector inequalities with an application to malnutrition inequalities in Vietnam,” *Journal of Econometrics*, 112, 207–223.

Wagstaff, A. and N. Watanabe (2003): “What difference does the choice of SES make in health inequality measurement?” *Health Economics*, 12, 885–890.

Wu, Y. (2006): “Overweight and obesity in China,” *British Medical Journal*, 333, 362–363.

Wunsch, C. and M. Lechner (2008): “What Did All the Money Do ? On the General Ineffectiveness of Recent West German Labour Market Programmes,” *Kyklos*, 61, 134–174.

WZB and Infas (2005): "Evaluation der Maßnahmen zur Umsetzung der Vorschläge der Hartz-Kommission - Modul 1a: Neuausrichtung der Vermittlungsprozesse," *Bericht Bundesministerium für Wirtschaft und Arbeit*.

Yang, D. (1999): "Urban-biased policies and rising income inequality in China," *American Economic Review*, 89, 306–310.

Yang, D. T. (2002): "What has caused regional inequality in China?" *China Economic Review*, 13, 331–334.

Zhu, Y.-g. and K. C. Jones (2010): "Urbanisation and health in China," *The Lancet*, 376, 232–233.

Part VII

Curriculum Vitae

Education

2011/12 - 2012/11	Visiting Researcher, Erasmus University Rotterdam, Host: Prof. Eddy van Doorslaer
2007/09 - present	Ph.D. in Economics and Finance, University of St.Gallen, Thesis supervisor: Prof. Dr. Michael Lechner
2004/10 - 2006/06	M.A. International Economics, University of Göttingen, Thesis supervisor: Prof. Michael Grimm
2005/02 - 2005/10	Academic Exchange, University of Groningen
2000/10 - 2004/07	B.A. Economics, University of Göttingen, Thesis supervisor: Prof. Michael Grimm

Professional Experience

2007/08 - 2011/08	*Research assistant*, Chair Prof. Dr. Michael Lechner, University of St.Gallen
2009/10 - 2011/06	*Teaching assistant* in Introductory Economics 1 & 2, University of St.Gallen
2009/10 - 2010/01	*Teaching assistant* Econometric Methods, University of St. Gallen
2006/10 - 2007/04	*Consultant*, German Company for Technical Cooperation (GTZ), Frankfurt
2005/11 - 2006/04	*Intern*, International Labour Organization (ILO), Geneva
2004/07 - 2004/10	*Intern*,German-Iranian Chamber of Industry and Commerce, Tehran

Grants

2007/08 - 2011/05	SNF Grant, Causal Analysis in Microeconometrics
2011/12- 2012/11	SNF Grant, Fellowships for prospective researchers
2013/05- 2015/04	Marie Curie Fellowship (FP7)