Empirical Essays on Different Aspects of Labor Economics

Von der Wirtschaftswissenschaftlichen Fakultät der
Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des akademischen Grades

Doktorin der Wirtschaftswissenschaften
- Doctor rerum politicarum -

genehmigte Dissertation
von Diplom-Mathematikerin Katja Alena Sonderhof
geboren am 19.11.1978 in Darmstadt

2010

Erstgutachter:	Prof. Dr. Patrick Puhani
Zweitgutachter:	Prof. Dr. Olaf Hübler
Tag der Disputation:	07.05.2010

Diese Arbeit widme ich meinen Eltern,
meiner Schwester und Andreas

Danksagung

Für diese Doktorarbeit schulde ich vielen Menschen ein großes Dankeschön. Zuallererst möchte ich mich bei meinem Doktorvater, Patrick Puhani, für seine kontinuierliche Hilfe, seine Unterstützung und Förderung, die konstruktive Kritik und die sehr gute Betreuung bedanken. Vor allem durch gemeinsame Projekte konnte ich viel von ihm lernen und so meine eigenen Arbeiten verbessern. Ich bin Patrick Puhani und meinem Zweitgutachter, Olaf Hübler, sehr dankbar für die gute Ausbildung und die produktive Zusammenarbeit.

Des Weiteren möchte ich mich bei meinen Eltern bedanken, ohne die ein Studium und eine Doktorarbeit niemals möglich gewesen wären. Sie haben mich nicht nur finanziell, sondern vor allem durch wertvolle Ratschläge und aufbauende Worte in den letzten Jahren sehr unterstützt. Das gleiche gilt für meine Schwester Anita, die immer ein offenes Ohr für Probleme und einen guten Vorschlag zur Ablenkung bereit hatte.

Darüber hinaus danke ich meinen Kollegen der TU Darmstadt und der Leibniz Universität Hannover für die gute Zusammenarbeit und die netten Kontakte. Vor allem Timo und Philip möchte ich für die vielen netten Kaffeepausen, aber auch für gute fachliche Ratschläge danken. Ein Dank auch an Laura, Andrea, Dennis, Anja, Falko, Thomas, Christian, Tatjana, Tobias, Stefan, Anh, Marina, Oliver, Karola, Friederike, Meike und Corinna, die mich auf der einen Seite durch fachliche Diskussionen, aber auch durch private Unternehmungen immer wieder zum durchhalten motiviert haben. Gleichermaßen gilt mein Dank meinen Darmstädter Freundinnen Ana, Antje, Bine, Eva, Julia, Julia und Sinje. Ebenso sei allen denen ein Dankeschön ausgesprochen, die nicht namentlich erwähnt wurden, aber zum Gelingen dieser Arbeit beigetragen haben.

Schließlich möchte ich mich bei Andreas für die tägliche Unterstützung und seine unendliche Geduld bedanken. Er hat mir in den letzten vier Jahren immer zur Seite gestanden, mir weitergeholfen, mich aufgebaut und mit voller Kraft unterstützt.

Abstract

This dissertation analyses four different topics in labor economics. In the main introduction, a short summary of all results is given. This abstract gives a very short overview of all topics that are covered in this dissertation: Chapter 1 focuses on estimating the effect of an extension of maternity leave from 18 to 36 months on young women's participation in job-related training. It is shown that maternity leave extension negatively affects job-related training for young women, especially when focusing on employer-arranged training. In Chapter 2 the effect of a reduction of sick pay on absence and on health-related outcomes is evaluated. Results show that a reduction of sick pay reduced absence from work significantly, while there is no effect on subjective health indicators. In Chapter 3 the relationship between overweight and wages is estimated. Results indicate discrimination against overweight women, since they receive significantly lower wages than women of healthy weight. Chapter 4 studies maternal labor supply and how it is related to childhood overweight, especially when taking birth order and age differences between siblings into account. Findings indicate that childhood overweight is positively related to the amount of hours worked by the mother. Moreover, this relationship is more pronounced for only children, lastborns and children with large age differences to their siblings.

Keywords: policy evaluation, discrimination, overweight

Kurzzusammenfassung

Diese Dissertation beschäftigt sich mit verschiedenen Aspekten der Arbeitsökonomik. In der Haupteinleitung werden die einzelnen Kapitel und die Ergebnisse kurz zusammengefasst. In dieser Kurzzusammenfassung wird ein Überblick über die einzelnen Kapitel und die darin enthaltenen Themen gegeben. In Kapitel 1 wird eine Reform zur Verlängerung des Erziehungsurlaubes im Hinblick auf mögliche negative Konsequenzen für junge Frauen am Arbeitsmarkt evaluiert. Es zeigt sich, dass durch die Verlängerung der Erziehungsurlaubszeiten junge Frauen signifikant weniger Weiterbildung bekommen, vor allem wenn es um Weiterbildung auf Betreiben des Arbeitgebers hin geht. Kapitel 2 evaluiert eine Reform zur Kürzung der Lohnfortzahlung im Krankheitsfall. Es zeigt sich, dass die gekürzte Lohnfortzahlung im Krankheitsfall die Fehlzeiten signifikant reduziert hat, während sie dabei keinen Einfluss auf das subjektive Gesundheitsempfinden gab. In Kapitel 3 geht es um den Zusammenhang zwischen Übergewicht und Löhnen. Die Ergebnisse zeigen einen signifikanten Zusammenhang zwischen Gewicht und Löhnen bei Frauen, wobei es diesen Zusammenhang bei Männern nicht zu geben scheint. Kapitel 4 beschäftigt sich mit dem Arbeitsangebot von Müttern und wie es mit Übergewicht bei den Kindern zusammenhängt. Es zeigt sich, dass Arbeitsstunden der Mutter positiv korreliert sind mit der Wahrscheinlichkeit, dass das Kind übergewichtig ist. Dieser Zusammenhang wird stärker bei Einzelkindern, Letztgeborenen und Kinder die eine große Altersdifferenz zu ihren Geschwistern haben.

Schlagwörter: Politikevaluation, Diskriminierung, Übergewicht

Contents

Chapter 3:
The Relationship between Overweight and Wages for Women in Germany: 107

Chapter 4:
Maternal Labor Supply and Childhood Overweight 146

List of Tables

List of Figures

Main Introduction

This dissertation consists of four chapters that are separate empirical research projects supported by the German Research Foundation (DFG) as is part of the research initiative ‘Flexibility in Heterogeneous Labour Markets’. The four chapters cover various aspects of empirical labor economics. While Chapter 1 and 2 are on policy evaluation of labor market reforms (both co-authored with Patrick Puhani) the remaining chapters study the economics of obesity, focusing on discrimination (Chapter 3) and maternal labor supply (Chapter 4). All research project focus on Germany and use German datasets; nevertheless, results are set in contrast with evidence from other countries in order to get an international perspective. In this introductory part, a short summary of all four chapters is given.

In Chapter 1, three representative individual-level datasets for West Germany are used to estimate the effect of an extension of maternity leave from 18 to 36 months on young women’s participation in job-related training. Since only young women of childbearing age are affected by the reform, difference-in-differences identification strategies are used to identify a causal effect of the reform on young women’s training participation. Results indicate that maternity leave extension negatively affects job-related training for young women - even if they do not have children - especially when the focus is on employer-arranged training. There is tentative evidence that young women partly compensated for this reduction in employer-arranged training by increasing training on their own initiative.

Chapter 2 evaluates the effects of a reduction in sick pay from 100 to 80% of the wage. Unlike previous literature, apart from absence from work, this study also considers effects on doctor/hospital visits and on subjective health indicators. Moreover, both switch-on and switch-off effects are estimated, because the reform was repealed two years later. Results show that the reform reduced the annual number of absent days by two days. Quantile regression reveals higher point estimates (both in absolute and relative terms) at higher quantiles, meaning that the reform predominantly reduced long durations of absence. In terms of health, the reform reduced the average number of days spent in hospital by almost half a day, but there is no robust evidence for negative effects on health outcomes or perceived liquidity constraints.

Chapter 3 estimates the relationship between overweight and wages with a large German dataset and finds that lower wages for obese women are likely to be due to discrimination. Obese women earn 2.4 percent lower wages than women having a BMI in the recommended range, while women who are in the top 10 percent of the body mass index get 4.3 percent lower wages than thinner women. The focus of this chapter is on whether these differences in wages are due to reduced productivity of overweight women or due to discrimination against them. These two hypotheses are tested using four different subgroup designs: I test whether gender-composition of coworkers plays a role and if contact to customers or coworkers matters when it comes to wages of overweight women. Moreover, I divide the sample into employed and self-employed women and into young and older women to test which group faces lower wages when overweight. Results of these subgroup estimations clearly support the discrimination hypothesis.

In Chapter 4, the correlation between maternal employment and overweight children is analyzed. Using German Micro Census data, it is clearly shown that there exists a strong relationship between the mother’s working activities and childhood overweight. Children of mothers in a fulltime employment have an up to 3.2 percentage points higher probability to be overweight. Moreover, it is found that birth order and age differences between siblings are significantly related to the probability of being overweight. Only children and lastborns have a higher probability to be overweight than firstborns or middleborns if the mother is working. Furthermore, the relationship between maternal employment and overweight children is stronger for children with large age difference to their siblings.

Chapter 1:

The Effects of Maternity Leave Extension on Training for Young Women*

Joint work by
Patrick A. Puhani and Katja A. Sonderhof

* This chapter is based on an earlier discussion paper version: IZA Discussion Paper (2008), No. 3820.

Acknowledgements: This research was supported by German Research Foundation (DFG) within the project 'Labour Market Effects of Social Policy' which is part of the research initiative 'Flexibility in Heterogeneous Labour Markets'. We are grateful to Christian Dustmann, Lena Edlund, Dominique Meurs, Uta Schönberg, Alfonso Sousa-Poza and seminar participants at the Universities of Darmstadt, Hannover and Hohenheim, ERMES, Université Paris II, the 2008 EEA, EALE and German Economic Association meetings, and participants at the research initiative's IAB and ZEW meetings in Nuremberg and Mannheim, respectively, for helpful comments. All remaining errors are our own.

1.1 Introduction to Chapter 1

Most industrialized countries have some form of maternal leave policy that effectively grants employment protection to women around childbirth. Arguments in favor of this employment protection refer to the well-being of both young children and their mothers. From a labor perspective, employment protection through maternity leave might increase the attachment of mothers to their employer or the labor force in general. However, it may also have the opposite effect in that maternity leave combined with maternity benefits can be seen as a subsidy to leave the labor market temporarily with potential long-term consequences.

Whereas previous studies on maternity leave with employment and wages as outcome variables have frequently discussed the role of human capital accumulation and depreciation, we know of no study relating human capital investments like training directly to maternity leave.[1] Therefore, in this paper, we estimate the effect of prolonged maternity leave on the human capital investments of women of working and childbearing age. To this end, we exploit the natural experiment of a 1992 extension in the employment protection (maternity leave) period in Germany from 18 to 36 months, which propelled Germany to the top position in the ranking of legislatively mandated maternity leave durations among industrialized countries.[2] To assess the effect of this reform on the human capital investments of young women workers, we draw on three individual-level datasets, all of which ask information on job-related training for women and men of different age groups.

It is well-established empirically that women are generally less attached to the labor force than men and that they receive less job-related training. For example, Barron, Black and

[1] Present discounted value of earnings, of which wage profiles and employment histories are major ingredients, might be the most appropriate outcome variable for the financial impact of maternity leave. However, measurement of the impacts on overall lifecycle wage and employment profiles is complicated by the frequent lack of long panel data. Conversely, impacts on wages at a certain point in the lifecycle may fail to take account of effects like steepened wage profiles. For example, when women have to bear a higher share of the costs of firm-specific training because of extended maternity leave, their early-career wages may fall, although Hashimoto's (1981) model would predict that they will also reap a higher share of the returns later in their careers. Thus, without data on lifecycle wage profiles, estimates with wages as the outcome might be difficult to interpret.

[2] See http://www.childpolicyintl.org/issuebrief/issuebrief5table1.pdf

Loewenstein (1993) show that U.S. workers with weaker attachment to the labor market are allocated to jobs offering less training, while women are employed in positions associated with shorter durations of job-related training. Similarly, Royalty (1996) finds a significant relationship in the U.S. between the predicted probability of job turnover and the probability of receiving training. Thus, the fact that women change their job positions more frequently accounts for about one fourth of the gender gap in training. For Britain, Green (1991) analyzes the differences in job-related training between young women and young men and between older women and older men. In comparison to young men, young women have less than half the chances of receiving training, although no differences are found between older women and older men.

Although these studies do not explicitly relate maternity leave to the incidence of training for women, they implicitly raise the question of whether prolonged maternity leave might affect job-related training for young women. The effect on training might be negative because a very long maternity leave reduces a young woman's labor market attachment, at least for the duration of the leave. As a consequence, employers should be less likely to invest in young women's human capital and place them in career paths with less training. Theoretically, the opposite effect might also prevail: if employers are forced to reemploy a woman even after a long leave, they might make the best of the situation and make up for lost human capital through intensified training. In the end, it is an empirical question which effect predominates.

Previous research has analyzed the relationships between both maternity leave and labor force participation and maternity leave and wages. For instance, Waldfogel (1999) finds no negative effects for the Family and Medical Leave Act's (FMLA) introduction of a 12-week maternity leave on the wages or employment of young women. Hashimoto et al. (2004) also find the effects to be negligible. Indeed, Waldfogel (1998) suggests that maternity leave may even increase young women's employment and wages, a finding corroborated by

Zveglich and van der Meulen Rodgers's (2003) investigation of a similar reform in Taiwan that introduced an 8-week maternity leave. Nevertheless, these findings contrast with those of Lai and Masters (2005) for Taiwan, as well as with Gruber's (1994) finding of a negative effect on wages of variations in maternity benefits across the U.S. They also contrast with the results of European studies that use reforms or other control group designs with longer maternity leave periods (up to three years). Among these, Ondrich et al. (2003) and Lalive and Zweimüller (2005), based on data from Germany and Austria, respectively, find that extended maternity leave results in short-run reductions in labor supply, while Schönberg and Ludsteck (2007) estimate negative long-run effects on wages in Germany. Likewise, in an analysis of policy variation in a panel of European countries, Ruhm (1998) reports increased employment due to parental leave (de facto maternity leave) but lower wages.

The remainder of this paper is organized as follows. Section 1.2 provides an overview of maternity leave regulations in Germany, especially with respect to the 1992 reform investigated here. Section 1.3 describes the datasets and the research design, after which Section 1.4 presents the difference-in-differences estimates of the effects of maternity leave extension on job-related training for women of childbearing age. Overall, these estimates show that the extension reduces training for young women, even for those who do not have children. A separate look at different types of training shows that it is particularly employer-arranged training that has been reduced by the extension of maternity leave. Point estimates suggest that young women are in return trying to compensate the reduction in employer-arranged training by increasing training on their own initiative. Section 1.5 concludes the paper.

1.2 Maternity leave in Germany

The duration of maternity leave as guaranteed by law in Germany exceeds that of other industrialized countries. For example, only since 1993 have U.S. federal regulations given women the right to take a 12-week maternity leave from work, even though many firms previously had their own maternity leave schemes. In contrast, as early as 1952 Germany enacted the first law protecting mothers (Mutterschutzgesetz) with a mandated 12-week maternity leave, which was extended in 1965 to 14 weeks (i.e., six weeks before and eight weeks after the predicted birth date). In 1979, this maternity leave duration was extended to an optional additional four months (decided on by the mother), and since 1986 the government has repeatedly increased the maximum duration of maternity leave (see Table 1.1), with the largest increase being the 1992 extension of the maximum duration from 18 to 36 months.[3]

One intention of policy makers when increasing the maximum maternity leave duration was to protect women from unemployment following the birth of a child. Another was to improve the welfare of children. Since public childcare facilities for children younger than three years of age are not generally available in Germany (having only recently gained broader political support in the western part of the country), all women are supposed to be given the opportunity to care for their children for up to three years.

By law, women also have the right to return to a job with their previous employer following maternity leave, not necessarily the same job but one comparable to that held before the leave. Nevertheless, not all women take this opportunity to return to the labor force. For example, Ondrich et al. (2003) and Weber (2004) find that a longer duration of maternity leave has a negative impact on the probability of women returning to the labor market, a finding also reported by Lalive and Zweimüller (2005) for Austria. For the U.S., Klerman and

[3] Since 1986, fathers have also been allowed to take part of the leave, but, according to the Federal Ministry of Families, Seniors, Women and Youth, only 1.5 percent of fathers make use of this opportunity.

Leibowitz (1990) show that because of better childcare facilities and less maternity leave protection, mothers return to the labor market sooner than in the past. Similarly, Waldfogel and Berger (2004) report that more than 80 percent of U.S. women working before childbirth return to work within 12 months after childbirth, while 55 percent return within 12 weeks after childbirth. In Germany, however, only around 55percent of all women working before a first birth return to the labor market within 24 months (Gustafsson et al. (1996)).

Figure 1.1 shows calculations of the average maternity leave durations for women working before childbirth based on biographical information from the German Socio-Economic Panel (GSOEP). In the first graph, we calculate the average period out of the labor force due to childbirth by adding the duration of formal maternity leave to the number of months after the leave until a mother was reemployed. In the second, we plot the average duration of maternity leave taken by mothers who return to work directly when the official maternity leave ends. The difference between the two lines is driven by the fact that in Germany many mothers stay at home with their children for many years, even after their maternity leave entitlement has run out. It should also be noted that we have very few observations (between 10 and 70 per year), so the numbers shown here have high sampling variance.

For both graphs, we have censored all durations at 36 months (the maximum maternity leave duration in Germany since 1992) because we are only interested in how far maternity leave extension drives career breaks up to that limit. As it turns out, maternity leave extension is associated with an increase in average career break durations due to childbirth. Keeping in mind the sampling variance, career break durations increased from around 20 months in the late 1980s to around 25 months in the early 1990s. If we only consider mothers who return to work directly following the official maternity leave (which may be for shorter periods than the legal limit), we observe a sharper increase in career breaks due to maternity leave, from around 5 to 10 months in the 1980s to between 15 and 20 months (and over 25 months in one estimate) in the 1990s. Moreover, the pattern in the curve of Figure 1.2, which outlines the increase in the share of time spent in official maternity leave by *all* young women aged 20 to 35 (excluding post-leave career breaks), is similar to that showing the length of official maternity leave. This figure also plots the development of fertility, which has declined somewhat but not dramatically over the last two decades, meaning that the extension of maternity leave has seemingly had no overwhelming effect on birth rates.

Thus, Figure 1.1 suggests that, ceteris paribus, mothers' labor force attachment decreases through the direct effect of maternity leave extension, especially for those women who return to the labor force within the first three years after childbirth. As it is difficult for employers to predict who will become a mother and when, all else being equal, the extension of the leave period has probably decreased the expected job attachment of all female employees at childbearing age, even though, as discussed later in Section 1.3, other factors besides maternity leave expansion might be impacting the labor supply of young women.

The literature also indicates that job-related training is likely to at least partly entail investment in firm-specific human capital. Theoretical results in Becker (1962) and Hashimoto (1981) raise the hypothesis that the reduction in young women's job attachment due to prolonged maternity leave will decrease firms' willingness to invest in job-related training for women of childbearing age (or at least reduce their willingness to bear the costs). Likewise, young women's willingness to invest in job-related training may also decrease due to a reduction in expected returns to that investment. Alternatively, young women may want to compensate the reduced willingness of employers to invest in their human capital, by undertaking more training on their own initiative. It is, however, an empirical question which effect dominates. We evaluate the impact of extended maternity leave on the incidence of job-related training for young women in the following.

1.3 Data and methodology

1.3.1 The treatment group and data sets

From the employer's perspective, extension of the maternity leave period constitutes an increase in employment protection for women of childbearing age. That is, if increased protection rights for young women are not reflected in implicit or explicit contracts that compensate employers for young women's extended maternity leave, women of childbearing age can expect diminished employment opportunities, such as less job-related training (cf. Lazear (1990)). However, unlike Schönberg and Ludsteck (2007), who consider extended maternity leave a treatment for mothers only and use mothers subject to shorter maternity leave as controls to measure labor force participation and wages as outcomes, we are interested in extended maternity leave rights as a treatment that affects *all* women of childbearing age with job-related training as an outcome of that treatment. Therefore, in our research design, the treatment group consists of all women of childbearing age, defined as those between 20 and 35 years of age. We exclude women between 36 and 39 because we cannot tell whether or not an employer perceives these women as being of childbearing age.[4]

In the subsequent analysis, we draw on three individual-level datasets that represent the West German workforce. East Germany was excluded because at the time of the reform, it was experiencing a major transition whose related factors are difficult to filter out from the effect of the maternity leave extension. In addition, the prereform points of observation are mostly from the 1980s when East Germany was under communist rule and thus excluded from the data. The three datasets used are the Report System [on] Further Education (Berichtssystem Weiterbildung, BSW)[5], the German Socio-Economic Panel (GSOEP),[6] and

[4] According to administrative birth records for Germany, 8.3 percent of all new mothers in 1990 were 36 years of age or older. This share is rising over time. For example, in the year 2000, it was already 11.5 percent. However, the share of all new mothers aged 40 or older is much lower at 1.8 and 2.5 percent in the years 1990 and 2000, respectively.

[5] More information on these data is available from the Central Archive for Empirical Social Research, University of Cologne web site: http://info1.za.gesis.org/DBKSearch12/SDesc.asp

the Qualification and Careers Survey (Qualifikation und Berufsverlauf, IAB-BIBB)[7]. A brief description of the datasets can be found in Appendix 1.2.

We restrict the sample to persons who are currently employed and hence attached to the labor market because by definition, persons not working cannot receive job-related training. Hence, we ignore the potential effect of extended maternity leave on training that works *directly* through (temporarily) reduced labor supply in order to focus on the effect for young women attached to the labor market (and thus potentially interested in job-related training). Nevertheless, because the three datasets we use measure the incidence of past job-related training for the last 1, 3 and 5 years (BSW, GSOEP, and IAB-BIBB, respectively), in two of the datasets we cannot avoid capturing some of the potential direct effect through the reduced labor supply that results from maternity leave.

Despite differences in the size of the event window referred to by the various surveys, all three datasets exhibit a large degree of communality in training incidence, with training participation in the BSW and GSOEP varying between a quarter and a third (see Table 1.2). In the IAB-BIBB data, participation is somewhat higher (between a third and almost one half) because this survey asks for training during the previous 5 years (compared to 1 year in the BSW and 3 years in the GSOEP).

As Table 1.2 shows, all datasets report an increase in training participation over time, a finding that holds true for all age-gender groups. Moreover, consistent with the growing emphasis on lifelong learning, training participation increased more among older (aged 40–55) than younger workers (aged 20–35). Note, however, that in 1994, after the extension of maternity leave, young women had the lowest incidence of employer-arranged training but the

---

[6] The GSOEP is probably the most frequently used individual-level data set for Germany. For more information, see http://www.diw.de/english/soep/29012.html

[7] The Qualification and Careers Survey (IAB-BIBB), which specializes in job descriptions, was also used by Spitz-Oener (2006). More information is available at http://www.gesis.org/Datenservice/Themen/38Beruf.htm

highest training incidence of training on the employee's initiative.[8] Neither of these facts held in 1988, before the extension of maternity leave. Formal testing of these before-after comparisons is carried out in Section 1.4.

1.3.2 Potential control groups

In tracking the development of job-related training of young women before and after the increase in the maternity leave period, we consider three demographic groups as reference points to construct control group designs: young men of similar age to the treatment group (20–35), women aged 40–55 years, and young men together with women and men aged 40–55. Similar treatment-control group designs are used in Gruber (1994), Ruhm (1998), Waldfogel (1999) and Lai and Masters (2005). We exclude persons older than 55 years from all analyses because this group's outcomes may be affected by other factors like early retirement, which may evolve differentially between men and women. In addition, training is less important to the older worker because the closer the retirement age, the lower the returns to investment.

Before comparing changes in training participation before and after the maternity leave extension for different age-gender groups, we check whether the extension of the maternity leave period did indeed lower young women's labor market attachment in relation to the potential control groups. This assessment is important because theory suggests that labor market attachment may be a key determinant of employers' willingness to support job-related training (Hashimoto, 1981). Likewise, observation of young women's labor force participation is important because general trends in female labor force participation may overlap with the effects of maternity leave on labor force participation and thus also influence job-related training. Hence, we must show an association between the German government's

[8] The question in the BSW asks whether job-related training was a) arranged by the company, b) arranged on the recommendation of a supervisor, or c) on your own initiative. We subsume answers a) and b) under 'employer-arranged training'.

extension of maternity leave duration and a decrease in young women's labor force participation relative to the control group. As Figure 1.1 has already shown, for young mothers, actual maternity leave periods have increased.

Figure 1.3 to Figure 1.5 profile the development of the full time equivalent (FTE) labor force participation rates of our treatment group (young women, irrespective of whether they are mothers) in relation to various controls. Because we restrict our sample to employees, self-employed are excluded; however, the graphs are robust to the inclusion of self-employed workers. We expect no abrupt change in labor market participation owing to maternity leave extension because hesitation to exploit the extended leave to its full extent is quite plausible in the face of uncertainty about how the employer will deal with this new situation. This view is borne out by the gradual increase in the average maternity leave period exhibited in Figure 1.1 and Figure 1.2.

Figure 1.3 shows the full time equivalent (FTE) labor force participation of young women relative to the older women controls in two datasets: the GSOEP and the Micro Census[9]. Even though the GSOEP's smaller sample size results in more erratic results than the Micro Census data, both datasets suggest that young women's labor force participation has decreased over the last two decades relative to that of older women. It should also be noted that the more reliable evidence from the Micro Census data suggests a much deeper decline in young versus older women's labor force participation in the late 1980s and early 1990s; that is, exactly during the period when maternity leave duration was massively extended (from 6 to 10 months in 1986, 12 months in 1988, 15 months in 1989, 18 months in 1990, and 36 months in 1992). This decline in relative participation is sizeable, at about 5 percentage points between the 1980s and 1993 according to the Micro Census. This steep downward trend

[9] The Micro Census (MZ) is a one-percent sample of the population (the scientific community receives only a 70 percent sample of that one percent) and asks similar questions to a census. For political reasons, there has been no census in Germany since 1987, so the Micro Census acts as a substitute. For more information, see http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/EN/press/abisz/Mikrozensus__e,templateId=renderPrint.psml

flattens in the mid-1990s, although it remains negative despite no further reforms to maternity leave.

Figure 1.4 presents a comparison between young women and young men. Although the labor force participation of the former is lower than that of the latter, young women have seemingly been catching up over time. Nevertheless, the Micro Census data clearly suggest that the long-run trend in catching up with young men stalled after 1992 (when the maternity leave period was doubled from 18 to 36 months) until about 2000. Hence, the short time series presented here is consistent with a permanent reduction in the labor force attachment of young women relative to their male peers. Taking into account that this reduction overlaps with an upward trend that dominates the data, we expect no decrease in young women's job-related training relative to young men. On the contrary, an increase is to be expected. This increase is actually observed in the data. However, because young men are not an adequate control group due to the trends observed here, we do not present the results with young men as the control group in this paper.[10]

The third alternative for the control group design compares young women to older women and relates this difference to young versus older men. Consequently, Figure 1.5 depicts the difference in the differences of FTE labor force participation rates between young and older women and young and older men. This development is similar to that for the older women control group: young women's labor force attachment declines relative to older women, and the gap between young and older women declines in relation to the gap between young and older men. This pattern holds true during the period of maternity leave expansion and in the years after 1992 until the (positive) difference between these two gaps remains constant or even increases again from the late 1990s onwards. Therefore, we expect a

[10] In results not shown here, it turns out, however, that despite of the catch-up in labor supply of young women in relation to young men, young women have lost in terms of *employer-arranged* training in relation to young men after the extension of maternity leave. When considering training *in general*, however, they have caught up. Yet, consistent with the relative labor supply trends shown here, this catch-up in terms of training in general was slowed down in the period when maternity leave was extended (compared to a placebo period).

decrease in the relative incidence of job-related training for young women with this control group design.

Based on the control group designs just presented, we estimate two sets of regression equations. The first is an estimate of the difference in training incidence between young and older women before and after the 1992 reform:

$$training_{it} = \alpha + \beta_1 X_{it} + \beta_2 after_{it} + \beta_3 young_{it} + \tau_1 (young_{it} \times after_{tt}) + \varepsilon_{it} \qquad (1)$$

where *training* is an indicator variable that is equal to 1 if training has occurred. *Young* and *after* are dummy variables indicating whether a women is young (20 to 35 years) and whether an observation refers to a post-1992 time point. The vector X denotes other control variables. In this difference-in-differences setup, the effect of interest is τ_1, which we expect to be negative because of the relative labor supply developments shown in Figure 1.3. If we have panel data (as in the GSOEP) instead of repeated cross sections (as in the other two datasets), we adjust standard errors for clustering (Bertrand, Duflo and Mullainathan, 2004).

If older women and young and older men are used as controls, we estimate a difference-in-difference-in-differences model using the following equation:

$$\begin{aligned} training_{it} = {} & \alpha + \beta_1 X_{it} + \beta_2 after_{it} + \beta_3 female_{it} + \beta_4 young_{it} \\ & + \beta_5 (female_{it} \times young_{it}) + \beta_6 (female_{it} \times after_{it}) + \beta_7 (young_{it} \times after_{it}) \quad (2) \\ & + \tau_2 (female_{it} \times young_{it} \times after_{it}) + \varepsilon_{it} \end{aligned}$$

with τ_2 as the coefficient of interest, which, as argued in Figure 1.5, is expected to be negative. The regression results are presented below.

1.4 Results

1.4.1 Before-after estimates by age and gender

We formalize the comparison of changes in training incidence in providing before-after estimates for the four age-gender groups: young women as the treatment group and older women and young and older men as the controls (see Table 1.3). We estimated results for four types of specifications. First, as would be appropriate if the before-after comparison was not confounded by any compositional effects or if any compositional effects were the outcome of extending the maternity leave period, we used no control variables (e.g., if young women invested less in education, education would be endogenous and thus should not be controlled for). We then successively increased the set of control variables in specifications 2 through 4, first by including dummy variables for education (i.e., high school diploma/A-level/Abitur and a tertiary polytechnic or university degree) and controlling for age and age squared to account for possible changes in the age distributions within age groups. In specification 3, we also added job characteristics using dummy variables for full-time, white-collar, and civil servant employment. Finally, in specification 4 we incorporated dummy variables for civil status (i.e., for being married and having children). Thus, those variables most likely to be endogenous were included last in the four specifications. In other words, we believed that family status and children might be affected by extended maternity leave, whose original intention was to facilitate women's work-life-family balance in order to increase fertility. If so, the civil status variables should not be included among the controls. Similar arguments might apply to the occupational and educational variables, but probably to a lesser extent. It turned out that controlling for these sets of variables had only a minor impact on the estimates. Hence, in the subsequent tables we only report estimates based on the specification with the full set of control variables.

As Table 1.3 shows, according to the BSW data, the smallest increase in training incidence between the 1988 and 1994 surveys (referring to training in 1987 and 1993, respectively) is among young women. That is, the point estimate in specification 4 exhibits an increase in training participation of 5.7 percentage points, significant only at the 10 percent level, compared to a 6.1 percentage point estimate for young men, significant at the 5 percent level. The point estimates for older women and men are even larger and highly significant, at 8.8 and 10.1 percentage points, respectively.

Although the difference between the increases in training for young women and men seem rather small (5.7 versus 6.1 percentage points), this contrast becomes much more pronounced when we distinguish between different types of training (only possible consistently over time in the BSW data). Young women's probability to have taken part in employer-arranged training only increased by 2.5 percentage points (statistically insignificant), the number for young men, however, is 7.0 percentage points (significant at the 5 percent level). By contrast, young women seem to have partially compensated for this divergence by investing more in training on their own (rather than their employer's) initiative: the increases in the training incidences for this type of training are 4.5 percentage points for young women (significant at the 10 percent level), but only 1 percent for young men (insignificant). Hence, overall young women have not only experienced somewhat lower increases in training than young men, but also a change in the type of training they receive in relation to young men: the before-after estimates suggest that employers were less interested in the training of young women in relation to young men after the extension of maternity leave. What is striking is that the BSW data report similar (and significant) increases in employer-arranged training for three age-gender groups: young men (7.0 percentage points), older women (7.3 percentage points) and older men (6.2 percentage points), but not for young women (insignificant 2.5 percentage points). As a result, young women compensated this development by an increase in their own initiative to obtain training.

When we compare the BSW results with the other two datasets, we can only consider training incidence in general, but not by the type of training (employer-arranged or not). As can be seen in Table 1.3, similar to the BSW data, the GSOEP and the IAB-BIBB data show an increase in training of older relative to younger workers (irrespective of gender). Hence, the BSW data seem to measure the same thing as the other two data sets. However, they give better information on the type of training.

Note that for training arranged by the employer, the 'age effect' in the increase in training is not observed any more (that is to say, although older workers have higher *general* training increases than younger workers, they do not exhibit higher *employer-arranged* training increases than young workers). Because training arranged by the employer is more relevant than training in general, we will put special emphasis on the BSW data in the following, but use GSOEP and IAB-BIBB data for robustness checks.

1.4.2 Difference-in differences estimates

As Table 1.4 demonstrates, job-related training is much more common among white-collar than among blue-collar workers (e.g., 33 versus 13 percent in the 1988 BSW survey). Among white-collar workers, training participation is higher in larger than in very small firms (28 versus 17 percent in the 1988 BSW survey), perhaps because the latter find it more difficult to substitute for workers who are currently in training. Maternity leave reform should be more likely to have an impact on a group of workers with a high training incidence. We thus also report estimates where we restrict the sample to white-collar workers in firms with at least 20 employees to see whether the estimates for this subsample are more pronounced than those for all workers. Unfortunately, the information on firm size varies between datasets so that in the IAB-BIBB data, the firm-size limit must be set to 10 instead of 20 employees. In the BSW and GSOEP data, however, we are able to limit the sample to white-collar workers in firms with at least 20 employees.

Table 1.5 presents the difference-in-differences estimates for the three datasets and two control group designs.[11] As argued in connection with relative labor supply developments (see Section 1.3), we expect young women to lose in terms of training incidence relative to older women because of the (accelerated) decrease in their labor supply after maternity leave was extended. The point estimates in Table 1.5 generally confirm this hypothesis, and the findings are statistically significant for two of the three datasets (GSOEP and IAB-BIBB).The point estimates are -4.9, -13.5, and -9.6 percentage points for the BSW, GSOEP, and IAB-BIBB datasets, respectively.[12]

Additionally, because average training participation differs between datasets, we also provide estimates of the change in training participation for young women relative to the prereform level. The resulting estimates imply a relative decline in training participation by 19, 44, and 29 percent in the BSW, GSOEP, and IAB-BIBB datasets, respectively. When restricted to white-collar workers in firms with at least 20 employees, the effects are even larger at -6.4, -21.8, and -13.1 percentage points in the BSW, GSOEP, and IAB-BIBB datasets, respectively. Especially large and significant are the estimates in those datasets that refer to a longer event window, such as training in the previous 3 and 5 years (the GSOEP and IAB-BIBB, respectively). As pointed out previously, the longer the event window, the larger the estimates will be in absolute value in that they include the direct effect of prolonged maternity leave on job-related training through temporary reduction of the labor supply due to maternity leave. Moreover, although the BSW, which only refers to the previous year, also suggests a large effect (a 19 percent reduction in job-related training for young women, and also 19 percent reduction when the sample is restricted to white-collar workers in larger

[11] Again, because control variables do not make a noteworthy difference to the estimates, we only report the specifications for the full set of control variables.
[12] As was the case with the before-after estimates, there is hardly any variation in the estimates across specifications with different control variables.

firms), the coefficient estimate is not significant. Therefore, we interpret these estimates as only tentative evidence that extended maternity leave reduces the incidence of job-related training for young women *in general* (below, we will see that there is ample evidence that *employer-arranged* training has been reduced).

In a second set of estimates using older women and young and older men as controls, we use a difference-in-difference-in-differences strategy to compare the changes in training incidence of young versus older women in relation to the changes of young versus older men. Based on the relative labor supply behavior reported earlier (see Figure 1.5), we expect negative estimates for τ_2, a hypothesis confirmed by all point estimates (see Table 1.5, last column). For the BSW, GSOEP, and IAB-BIBB datasets, respectively, the point estimates suggest a -1.8, -5.5, and -2.0 percentage point change in young women's training participation. The estimates are not statistically significant. When restricted to white-collar workers in firms with at least 20 employees, the corresponding estimates are -8.7, -14.6 (significant at the 10 percent level), and -4.9 in the BSW, GSOEP, and IAB-BIBB datasets, respectively.

Note that the just presented investigation of the impact of maternity leave reform on the incidence of any job-related training makes no distinction between types of training, which, unlike schooling, is poorly classified in most surveys. Nevertheless, unlike the other two datasets, the BSW data has information on whether training was arranged directly by the employer or taken on the employee's own initiative (information lacking in the other datasets).[13] We therefore apply the same estimates as above but distinguish between different types of training.

[13] The GSOEP provides information on these training aspects, but the questions are inconsistent across the years.

1.4.3 Employer-arranged training vs. training on the worker's initiative

Only the BSW data provide information on the role of the employer in job-related training. In the following, we calculate separate estimates for job-related training arranged directly by the employer and training on the employee's own initiative. The incidence of the two types of training for the four age-gender groups in the BSW data is reported in Table 1.2. Whereas in the first year of observation (1988), 26 percent of all workers in the sample received some type of job-related training, only 14 percent received training arranged by the employer.

Table 1.6 shows estimation results for these two types of training using the same control group designs as before. Again, we report two blocks of estimates, one for the full sample and one for white-collar workers in firms with at least 20 employees. The fact that only a few estimates are statistically significant may be due to the sample size. However, it should be noted that all the point estimates for employer-arranged training are negative. For the subsample of white-collar workers in larger firms, by contrast, all point estimates for training on the employee's initiative are positive (albeit not statistically significant), but smaller than the negative ones for employer-arranged training. Hence, young women seem at best to have partially compensated for the reduced interest in training by their employers.

For employer-arranged training, the point estimates indicate a reduction in young women's training participation of 4.6 or 7.0 percentage points depending on the control group. For young white-collar women, these estimates are larger and statistically significant, at 9.6 and 15.7 percentage points. In relation to the training incidence before maternity leave extension, these point estimates are huge, implying a reduction in employer-arranged training of between 35 and 54 percent for all young women and between 53 and 87 percent for young white-collar women in firms with at least 20 employees. Moreover, estimated increases in training on the employee's initiative lie between 20 and 43 percent. We must not overemphasize these large numbers because of the large standard errors attached to the estimates. Nevertheless, point estimates are in a similar range irrespective of the control group chosen (for training on the employee's initiative this is only true if the sample is restricted to white-collar workers in larger firms).

1.4.4 Effects for young women without children

Maternity leave extension should affect women of childbearing age even if they currently have no children because they are at risk of leaving the employer for up to three years with a right to return. This risk is enough to make them part of the treatment group. Therefore, to check whether the results so far also apply to women without children, we repeat the estimates provided in Table 1.5 and Table 1.6 for young women who do not have children as treatment group (the control groups remain unchanged). These results, presented in Table 1.7 and Table 1.8, show very similar point estimates to those obtained for the full samples, which include young women with children. However, as shown in Table 1.8, in the BSW estimates for different types of training, none of the estimates remain statistically significant once young women with children are excluded (cf. Table 1.6). However, the point estimates still remain consistently negative and economically significant both for general training and for employer-arranged training, but positive in three out of four cases for training on the employee's initiative. The statistical insignificance may simply be the result of reduced sample size and correspondingly large standard errors. In sum, there is some evidence that maternity leave extension has reduced job-related training even for young women without children. Again it seems that this result is mainly driven by a reduction in employer-arranged training, which has only been partly compensated by training on young female employees' own initiative.

1.4.5 Placebo estimates

The estimates so far seem to suggest that young women receive less employer-arranged job-related training because of the extension of maternity leave from 18 to 36 months. Methodologically, we have relied on the difference-in-differences assumption that – in the absence of maternity leave extension – the training gap between treatment and control group would have remained constant. Because this identifying assumption is not testable and because there may be differential trends in training participation between treatment and control groups even in the absence of training, we carry out placebo estimates. This is to say, we estimate the same difference-in-differences models for a period in which no change in maternity leave took place. Because extensions have been frequent since 1979 and more data are available for recent years, we choose a postreform period for such estimates. However, owing to data availability constraints, we can only use 1997 and 2003 data from the BSW and 2000 and 2004 data from the GSOEP and must exclude the IAB-BIBB, whose last two waves occurred in 1991 and 1998.

The placebo estimation results for the three categories of training in the BSW data are provided in Table 1.9 for both the full sample and the subsample of white-collar workers in firms with at least 20 employees. These placebo estimates correspond to the results for the extension of maternity leave given in Table 1.6. Whereas in Table 1.6 the estimates for employer-arranged training are all negative and somewhat similar in magnitude, with half of them being statistically significant, none of the placebo estimates in Table 1.9 are simultaneously negative and statistically significant. In general, the placebo test is not so convincing when all workers – including blue-collar – are considered (upper panel of Table 1.9). There are, however, clear contrasts in the point estimates of the reform and placebo periods both for training in general and for employer-arranged training when we restrict the sample to white-collar workers in larger firms, i.e. the group of workers for whom job-related training seems most relevant. This comparison provides further support for the hypothesis that young women's participation in training has been held back by the maternity leave extension. Concerning training on the employee's initiative the point estimates in the placebo time period are similar to the ones for the period when maternity leave was extended. Hence, the evidence that young women partly compensated the reduction in employer-arranged training by their own initiative is only tentative.

The placebo estimates based on the GSOEP also support the assumption that our previous results on reduced training for young women due to extended maternity leave were not spurious. Whereas our GSOEP-based estimates using older women and older women together with young and older men as controls were significantly negative (see Table 1.5), the corresponding placebo estimates, given in Table 1.10, are all positive and insignificantly different from zero. Moreover, as the table shows, these placebo results hold for both the full sample and white-collar workers in firms with at least 20 employees.

1.5 Conclusions of Chapter 1

Even though policies that support the family-work balance are contentious on both sides of the Atlantic, maternity leave that guarantees a post-leave right to return to work is an important component of family policies. Whereas some countries like the United States opt for very short maternity leave periods (i.e., 12 weeks), Germany lies at the other extreme, having extended maternity leave with a right to return to work with the same employer from 18 to 36 months, which ranks in the highest maternity leave durations in industrialized countries. In this paper, we use difference-in-differences estimates to investigate the effect of this extension on the human capital investment of young women workers.

Specifically, drawing on three individual-level datasets that represent West German workers, we measure participation in job-related training as a proxy for human capital investment, taking care to consider long-term trends in labor force participation when interpreting our difference-in-differences estimates using alternative control groups. Similar to the previous literature, we choose older women and older women together with young and older men as control groups. We particularly focus on one dataset which distinguishes between employer-arranged training and training on the employee's initiative.

We find significant evidence that maternity leave extension reduced employer-arranged training for young women. There is also tentative evidence that young women partly compensated for this reduction in employer-arranged training by undertaking more training on their own initiative.

Taken together with extant findings on extended maternity leave in European countries, our results point to negative economic consequences of protective measures like maternity leave of up to three years (as in Germany) for all young working women, even those without children. These negative effects must be weighed against the potential job security benefits for those who become mothers and the potential benefits for their children. However, as Dustmann and Schönberg's (2008) regression discontinuity estimates illustrate, this latter may be close to zero.

References for Chapter 1

Barron, J.M., D.A. Black and M.A. Loewenstein (1993) Gender Differences in Training, Capital, and Wages, Journal of Human Resources, Vol. 28(2), 343-364.

Becker, G. (1962) Investment in Human Capital: A Theoretical Analysis, Journal of Political Economy, Vol. 70(5), 9-49.

Bertrand, M., E. Duflo and S. Mullainathan (2004) How Much Should We Trust Difference-in-Differences Estimates?, Quarterly Journal of Economics, Vol. 119(1), 249-275.

Dustmann, Ch. and U. Schönberg (2008) The Effect of Expansions in Leave Coverage on Children's Long-Term Outcomes, IZA Discussion Paper No. 3605, Bonn.

Green, F. (1991) Sex Discrimination in Job-Related Training, British Journal of Industrial Relations, Vol. 29(2), 295-304.

Gustafsson, S.S., C.M.M.P. Wetzels, J.D. Vlasblom and S. Dex (1996) Women's Labor Force Transitions in Connection with Children: A Panel Data Comparison Between Germany, Sweden and Great Britain, Journal Population Economics, Vol. 9(3), 223-246.

Gruber, J. (1994) The Incidence of Mandated Maternity Benefits, American Economic Review, Vol. 84(3), 622-641.

Hashimoto, M. (1981) Firm-Specific Human Capital as a Shared Investment, American Economic Review, Vol. 71(3), 475-482.

Hashimoto, M., R. Percy, T. Schoellner and B.A. Weinberg (2004) The Long and Short of It: Maternity Leave Coverage and Women's Labor Market Outcomes, IZA Discussion Paper No. 1207, Bonn.

Klerman, J.A. and A. Leibowitz (1990) Child Care and Women's Return to Work After Childbirth, American Economic Review, Vol. 80(2), 284-288.

Kreyenfeld, M. (2001) Employment and Fertility – East Germany in the 1990s, Rostock University, doctoral thesis.

Lai, Y.-Ch. and S. Masters (2005) The Effects of Mandatory Maternity and Pregnancy Benefits on Women's Wages and Employment in Taiwan, 1984-1996, Industrial and Labor Relations Review, Vol. 58(2), 274-281.

Lalive, R. and J. Zweimüller (2005) Does Parental Leave Affect Fertility and Return-to-Work? Evidence from a "True Natural Experiment", IZA Discussion Paper No. 1613, Bonn.

Lazear, E. (1990) Job Security Provisions and Employment, Quarterly Journal of Economics, Vol. 105(3), 699-726.

Ondrich, J., C.K. Spiess, Q. Yang and G.G. Wagner (2003) The Liberalization of Maternity Leave Policy and the Return to Work after Childbirth in Germany, Review of Economics of the Household, Vol. 1(1-2), 77-110.

Royalty, A.B. (1996) The Effects of Job Turnover on the Training of Men and Women, Industrial and Labor Relations Review, Vol. 49(3), 506-521.

Ruhm, Ch.J. (1998) The Economic Consequences of Parental Leave Mandates: Lessons from Europe, Quarterly Journal of Economics, Vol. 113(1), 285-317.

Schönberg, U. and J. Ludsteck (2007) Maternity Leave Legislation, Female Labor Supply, and the Family Wage Gap, IZA Discussion Paper No. 2699, Bonn.

Spitz-Oener, A. (2006) Technical Change, Job Tasks and Rising Educational Demands: Looking Outside the Wage Structure, Journal of Labor Economics, Vol. 24(2), 235-270.

Waldfogel, J. (1998) The Family Gap for Young Women in the United States and in Britain: Can Maternity Leave make a Difference?, Journal of Labor Economics, Vol. 16(3), 505-545.

Waldfogel, J. (1999) The Impact of the Family and Medical Leave Act, Journal of Policy Analysis and Management, Vol. 18(2), 281-302.

Waldfogel, J. and L.M. Berger (2004) Maternity Leave and the Employment of New Mothers in the United States, Journal of Population Economics, Vol. 17(2), 331-349.

Weber, A.M. (2004) Wann kehren junge Mütter auf den Arbeitsmarkt zurück? Eine Verweildaueranalyse für Deutschland [When Do Young Mothers Return to the Labor Market, A Duration Analysis for Germany, in German], ZEW Discussion Paper No. 04-08, Mannheim.

Wohn, K. (2007) Effizienz von Weiterbildungsmessung [Efficiency of Measuring Training Participation, in German]. Rat für Sozial- und Wirtschaftsdaten [German Council for Social and Economic Data], Working Paper No. 19.

Zveglich, Jr., J.E. and Y. van der Meulen Rodgers (2003) The Impact of Protective Measures for Female Workers, Journal of Labor Economics, Vol. 21(3), 533-555.

Tables and Figures of Chapter 1

Table 1-1: Increase of maximum maternity leave duration

Year	Duration	Maternity leave
1979	4 months	Introduction of a 4-months maternity leave, which can be taken in addition to the 14 weeks retention period. Maternity benefits up to 750 deutschmarks (about €375) per month) paid by the government
1986	10 months	Maternity leave can be taken by mother or father. Both are allowed to work for up to 19 hours per week. Maternity leave can be exchanged once between mother and father. Less than 2 percent of men take this opportunity. Parental benefits of 600 deutschmarks (about €300) per month paid for 10 months by the government
1988	12 months	Duration of maternity leave is extended to 12 months. Parental benefits of 600 deutschmarks (about €300) per month paid for 12 months by the government
1989	15 months	Duration of maternity leave is extended to 15 months. Parental benefits of 600 deutschmarks (about €300) per month paid for 15 months by the government
1990	18 months	Duration of maternity leave is extended to 18 months. Parental benefits of 600 deutschmarks (about €300) per month paid for 18 months by the government
1992	36 months	Duration of maternity leave is extended to 36 months. Demand for maternity leave can be exchanged three times between mother and father. Parental benefits of 600 deutschmarks (about €300) per month paid for 24 months by the government

Source: Kreyenfeld (2001).

Table 1-2: Descriptive statistics: training participation

a) All datasets	BSW		GSOEP		IAB-BIBB	
	1988	1994	1989	2000	1991	1998
All	0.26	0.34	0.30	0.36	0.35	0.42
Young women	0.26	0.30	0.31	0.35	0.33	0.37
Older women	0.18	0.31	0.16	0.34	0.26	0.39
Young men	0.32	0.36	0.38	0.38	0.37	0.39
Older men	0.25	0.36	0.32	0.37	0.39	0.51
n	3,112	2,147	2,764	5,639	16,682	17,564

b) Detailed information only in BSW	BSW Training arranged by employer		BSW Training on one's own initiative	
	1988	1994	1988	1994
All	0.14	0.19	0.12	0.14
Young women	0.13	0.15	0.13	0.16
Older women	0.09	0.16	0.10	0.14
Young men	0.16	0.22	0.15	0.14
Older men	0.16	0.22	0.08	0.14
n	3,112	2,147	3,112	2,147

Source: Report System Further Education (BSW); German Socio-Economic Panel (GSOEP); Qualification and Careers (IAB-BIBB); own calculations.

Table 1-3: Before-after estimates

a) All datasets

BSW - Training in general	
Young women n=1,188	0.057* (0.031)
Older women n=1,016	0.088*** (0.029)
Young men n=1,405	0.061** (0.029)
Older men n=1,456	0.101*** (0.027)

GSOEP - Training in general	
Young women n=1,716	0.002 (0.030)
Older women n=1,849	0.128*** (0.024)
Young men n=2,175	-0.004 (0.028)
Older men n=2,539	0.039* (0.023)

IAB-BIBB - Training in general	
Young women n=7,513	0.012 (0.012)
Older women n=6,823	0.104*** (0.012)
Young men n=9,560	0.006 (0.011)
Older men n=10,072	0.085*** (0.010)

b) Detailed information only in BSW data

BSW - Training arranged by employer	
Young women n=1,188	0.025 (0.024)
Older women n=1,016	0.073*** (0.022)
Young men n=1,405	0.070*** (0.025)
Older men n=1,456	0.062*** (0.024)

BSW - Training on one's own initiative	
Young women n=1,188	0.045* (0.027)
Older women n=1,016	0.047*** (0.023)
Young men n=1,405	0.010 (0.024)
Older men n=1,456	0.066*** (0.020)

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. The estimates are based on regressions with the following set of control variables: age, age squared, dummy variables for high school (A-level, *Abitur*) and tertiary degrees, dummy variables for full-time employment, white-collar job, civil-service employment, for being married and for having children. Because control variables only have a minor impact on these estimates, we only report the results with the full set of controls.
Source: Report System Further Education (BSW); German Socio-Economic Panel (GSOEP); Qualification and Careers (IAB-BIBB); own calculations.

Table 1-4: Training participation for subgroups

	Blue-collar worker	White-collar worker	White-collar worker in firms with at least 20 employees	White-collar worker in firms with less than 20 employees
BSW	0.13	0.33	0.28	0.17
GSOEP	0.13	0.41	0.34	0.16
IAB-BIBB	0.20	0.44	0.37	0.25

Note: Figures refer to the survey years before the reform: 1988 (BSW), 1989 (GSOEP) and 1991 (IAB-BIBB).
Source: Report System Further Education (BSW); German Socio-Economic Panel (GSOEP); Qualification and Careers (IAB-BIBB); own calculations.

Table 1-5: Difference-in-differences estimates

BSW	DiD Control group: Older women	DiDiD Control group: Older women and all men
Full sample	-0.049	-0.018
	(0.042)	(0.057)
Relative deviation	-0.19	-0.07
n	2,204	5,065
White-collar workers in firms with at least 20 employees	-0.064	-0.087
	(0.058)	(0.082)
Relative deviation	-0.19	-0.26
n	1,378	2,873
GSOEP		
Full sample	-0.135***	-0.055
	(0.038)	(0.052)
Relative deviation	-0.44	-0.18
n	3,508	8,146
White-collar workers in firms with at least 20 employees	-0.218***	-0.146*
	(0.056)	(0.080)
Relative deviation	-0.51	-0.34
n	1,991	4,362
IAB-BIBB		
Full sample	-0.096***	-0.020
	(0.017)	(0.022)
Relative deviation	-0.29	-0.06
n	14,336	33,968
White-collar workers in firms with at least 10 employees	-0.131***	-0.049
	(0.024)	(0.032)
Relative deviation	-0.31	-0.11
n	8,448	18,216

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. The estimates are based on regressions with the following set of control variables: age, age squared, dummy variables for high school (A-level, *Abitur*) and tertiary degrees, dummy variables for full-time employment, white-collar job, civil-service employment, for being married and for having children. Because control variables only have a minor impact on these estimates, we only report the results with the full set of controls.
Source: Report System Further Education (BSW); German Socio-Economic Panel (GSOEP); Qualification and Careers (IAB-BIBB); own calculations.

Table 1-6: Difference-in-differences estimates:
Results for different types of training – BSW

Full sample	DiD Control group: Older women	DiDiD Control group: Older women and all men
Job-related training (general)	-0.049	-0.018
	(0.042)	(0.057)
Relative deviation	-0.19	-0.07
Job-related training (arranged by employer)	-0.046	-0.070
	(0.033)	(0.047)
Relative deviation	-0.35	-0.54
Job-related training (on one's own initiative)	-0.003	0.049
	(0.035)	(0.043)
Relative deviation	-0.02	0.38
n	2,203	5,065
White-collar workers in firms with at least 20 employees		
Job-related training (general)	-0.064	-0.087
	(0.058)	(0.082)
Relative deviation	-0.19	-0.26
Job-related training (arranged by employer)	-0.096**	-0.157**
	(0.047)	(0.077)
Relative deviation	-0.53	-0.87
Job-related training (on one's own initiative)	0.030	0.065
	(0.049)	(0.065)
Relative deviation	0.20	0.43
n	1,378	2,873

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. The estimates are based on regressions with the following set of control variables: age, age squared, dummy variables for high school (A-level, *Abitur*) and tertiary degrees, dummy variables for full-time employment, white-collar job, civil-service employment, for being married and for having children.
Source: Report System Further Education (BSW); own calculations.

Table 1-7: Difference-in-differences estimates for young women without children

	DiD Control group: Older women	DiDiD Control group: Older women and all men
BSW		
Full sample	-0.069	-0.024
	(0.049)	(0.063)
Relative deviation	-0.23	-0.08
n	1,712	4,568
White-collar workers in firms with at least 20 employees	-0.049	-0.060
	(0.067)	(0.088)
Relative deviation	-0.14	-0.17
n	1,106	2,598
GSOEP		
Full sample	-0.170***	-0.088
	(0.044)	(0.057)
Relative deviation	-0.45	-0.23
n	2,972	7,610
White-collar workers in firms with at least 20 employees	-0.246***	-0.173**
	(0.060)	(0.083)
Relative deviation	-0.52	-0.37
n	1,735	4,106
IAB-BIBB		
Full sample	-0.103***	-0.014
	(0.019)	(0.024)
Relative deviation	-0.28	-0.04
n	11,784	31,416
White-collar workers in firms with at least 10 employees	-0.133***	-0.044
	(0.026)	(0.035)
Relative deviation	-0.30	-0.10
n	7,101	16,869

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. The estimates are based on regressions with the following set of control variables: age, age squared, dummy variables for high school (A-level, *Abitur*) and tertiary degrees, dummy variables for full-time employment, white-collar job, civil-service employment, for being married and for having children. Because control variables only have a minor impact on these estimates, we only report the results with the full set of controls.

Source: Report System Further Education (BSW); German Socio-Economic Panel (GSOEP); Qualification and Careers (IAB-BIBB); own calculations.

Table 1-8: Difference-in-differences estimates for young women without children – Results for different types of training – BSW

	DiD	DiDiD
	Control group:	Control group:
Full sample	Older women	Older women and all men
Job-related training	-0.069	-0.024
(general)	(0.049)	(0.063)
Relative deviation	-0.23	-0.08
Job-related training	-0.033	-0.056
(arranged by employer)	(0.040)	(0.052)
Relative deviation	-0.25	-0.43
Job-related training	-0.035	0.029
(on one's own initiative)	(0.039)	(0.048)
Relative deviation	-0.22	0.22
n	1,712	4,568
White-collar workers in firms with at least 20 employees		
Job-related training	-0.049	-0.060
(general)	(0.067)	(0.088)
Relative deviation	-0.14	-0.17
Job-related training	-0.062	-0.123
(arranged by employer)	(0.056)	(0.077)
Relative deviation	-0.34	-0.68
Job-related training	0.011	0.057
(on one's own initiative)	(0.053)	(0.069)
Relative deviation	0.06	0.32
n	1,106	2,601

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. The estimates are based on regressions with the following set of control variables: age, age squared, dummy variables for high school (A-level, *Abitur*) and tertiary degrees, dummy variables for full-time employment, white-collar job, civil-service employment, for being married and for having children.
Source: Report System Further Education.

Table 1-9: Placebo tests – BSW

	DiD Control group: Older women	**DiDiD** Control group: Older women and all men
Full sample		
Job-related training	-0.007	-0.096
(general)	(0.051)	(0.071)
Job-related training	0.024	-0.042
(arranged by employer)	(0.045)	(0.061)
Job-related training	-0.035	-0.055
(on one's own initiative)	(0.040)	(0.057)
n	1,817	3,736
White-collar workers in firms with at least 20 employees		
Job-related training	0.088	0.001
(general)	(0.070)	(0.099)
Job-related training	0.069	-0.014
(arranged by employer)	(0.065)	(0.092)
Job-related training	0.016	0.017
(on one's own initiative)	(0.055)	(0.084)
n	1,075	2,128

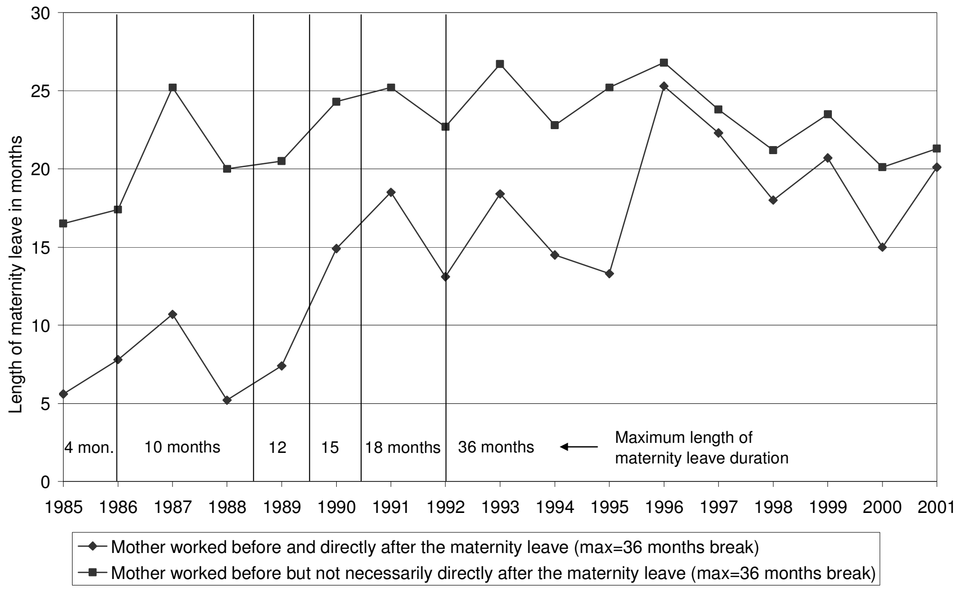
Note: *, ** and *** denote significance at the 10 %, 5 % and 1 % level, respectively. The estimates are based on regressions with the following set of control variables: age, age squared, dummy variables for high school (A-level, *Abitur*) and tertiary degrees, dummy variables for full-time employment, white-collar job, civil-service employment, for being married and for having children.
Source: Report System Further Education (BSW); own calculations.

Table 1-10: Placebo tests – GSOEP

	DiD Control group: Older women	**DiDiD** Control group: Older women and all men
Full sample		
Job-related training	0.029	0.023
(general)	(0.034)	(0.054)
n	4,232	9,426
White-collar workers in firms with at least 20 employees		
Job-related training	0.059	0.033
(general)	(0.057)	(0.078)
n	2,478	5,207

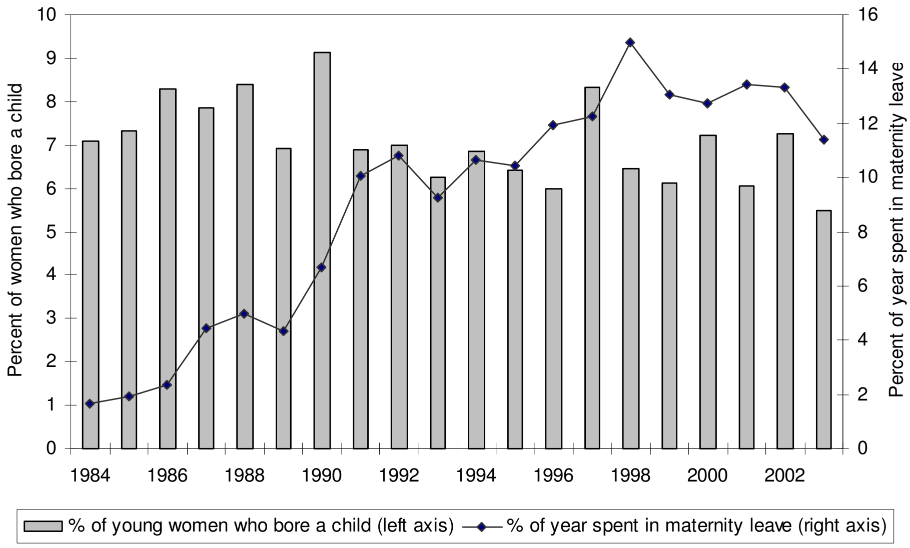
Note: *, ** and *** denote significance at the 10 %, 5 % and 1 % level, respectively. The estimates are based on regressions with the following set of control variables: age, age squared, dummy variables for high school (A-level, *Abitur*) and tertiary degrees, dummy variables for full-time employment, white-collar job, civil-service employment, for being married and for having children.
Source: German Socio-Economic Panel (GSOEP); own calculations.

Figure 1-1: Average length of maternity leave taken



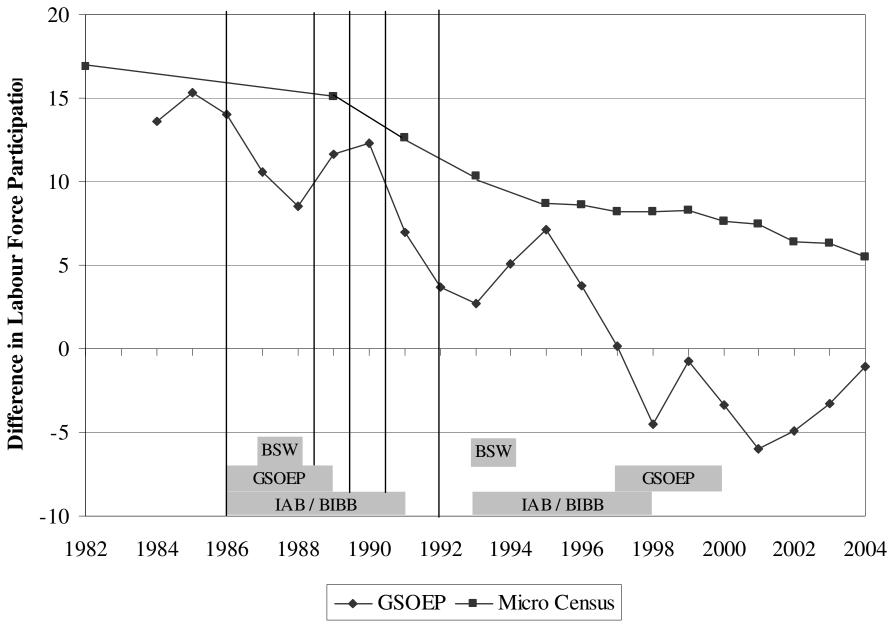
Note: All durations longer than 36 months were censored to 36 months. Vertical lines show the timing of the reforms to increase maternity leave duration. The length of maternity leave is measured in months for women between 20 and 35 years of age who started their maternity leave in the year before the interview. In the top line, we add the durations of official maternity leave and post maternity leave career breaks, which are common in Germany. The lower line only considers official maternity leave for mothers who return to the labor market directly after their official maternity leave.
Source: German Socio-Economic Panel (GSOEP); own calculations.

Figure 1-2: Young women's labor force participation and birth rates



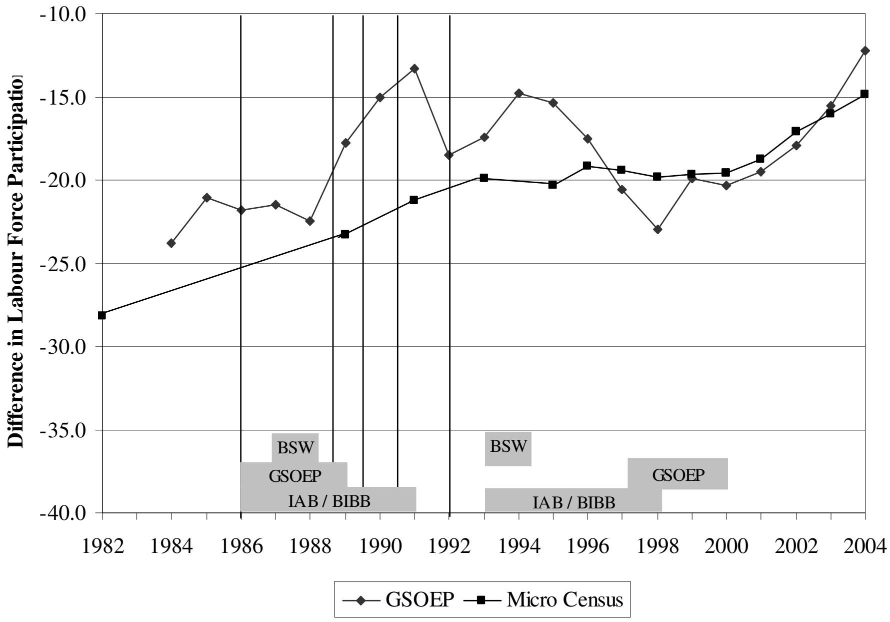
Note: These results refer to all women aged between 20 and 35 years. Percentage rate of year spent in maternity leave gives an idea of how long young women, on average, are absent due to maternity leave each year.
Source: German Socio-Economic Panel (GSOEP); own calculations.

Figure 1-3: Difference between young and older women's labor force participation – Full-time equivalents



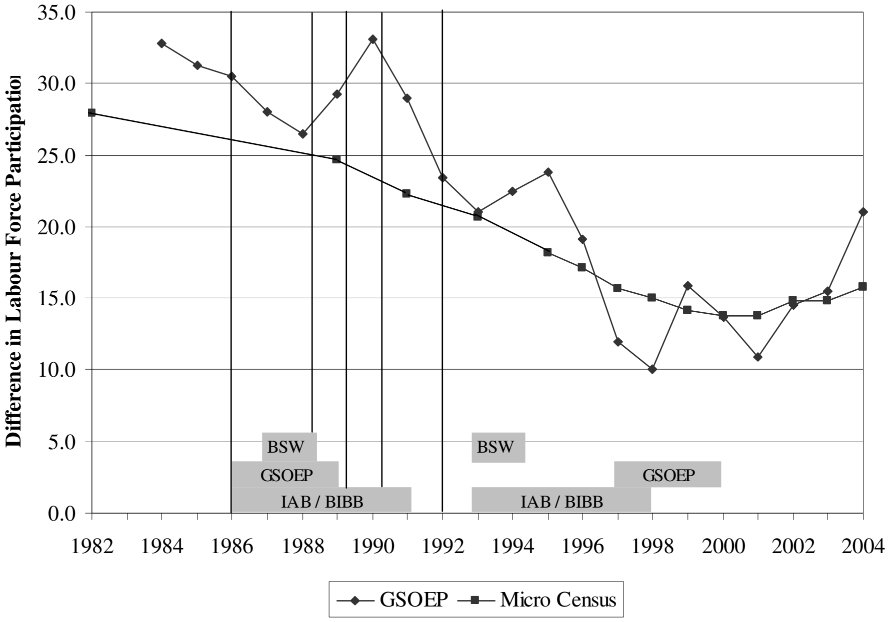
Note: The boxes at the bottom of the graphs indicate the event windows referred to in the training questions in the respective surveys. As mentioned in the text, the BSW refers to job-related training in the previous year, whereas the GSOEP and IAB-BIBB data refer to the previous three and five years, respectively. Vertical lines show the timing of the reforms to increase maternity leave duration.
Source: Micro Census (MZ); German Socio-Economic Panel (GSOEP); own calculations.

Figure 1-4: Difference between young women's and young men's labor force participation – Full-time equivalents



Note: The boxes at the bottom of the graphs indicate the event windows referred to in the training questions in the respective surveys. As mentioned in the text, the BSW refers to job-related training in the previous year, whereas the GSOEP and IAB-BIBB data refer to the previous three and five years, respectively. Vertical lines show the timing of the reforms to increase maternity leave duration.
Source: Micro Census (MZ); German Socio-Economic Panel (GSOEP); own calculations.

Figure 1-5: Difference in difference of young and older persons' labor force participation between men and women – Full-time equivalents



Note: The boxes at the bottom of the graphs indicate the event windows referred to in the training questions in the respective surveys. As mentioned in the text, the BSW refers to job-related training in the previous year, whereas the GSOEP and IAB-BIBB data refer to the previous three and five years, respectively. Vertical lines show the timing of the reforms to increase maternity leave duration.
Source: Micro Census (MZ); German Socio-Economic Panel (GSOEP); own calculations.

Appendix to Chapter 1

Appendix 1-1: Summary statistics

	BSW		GSOEP		IAB-BIBB	
	1988	1994	1989	2000	1991	1998
Training (All)	0.25	0.33	0.30	0.36	0.35	0.42
High school	0.20	0.25	0.20	0.30	0.18	0.26
University	0.15	0.17	0.12	0.20	0.16	0.14
Age	37.2	37.8	37.1	39.0	36.9	38.2
Age between 20 and 35	0.51	0.51	0.53	0.47	0.54	0.48
White-collar Worker	0.53	0.53	0.53	0.61	0.51	0.56
Blue-Collar Worker	0.34	0.34	0.36	0.30	0.40	0.34
Civil Servant	0.12	0.13	0.11	0.10	0.09	0.10
Female	0.40	0.39	0.42	0.44	0.41	0.43
Married	0.68	0.66	0.58	0.55	0.75	0.74
Children	0.41	0.48	0.31	0.32	0.34	0.42
Working full-time	0.83	0.81	0.86	0.81	0.88	0.84
n	3,112	2,147	2,764	5,639	16,682	17,564

Source: Report System Further Education (BSW); German Socio-Economic Panel (GSOEP); Qualification and Careers (IAB-BIBB); own calculations.

Appendix 1-2: Data description

The three datasets used are the Report System [on] Further Education (Berichtssystem Weiterbildung, BSW), the German Socio-Economic Panel (GSOEP), and the Qualification and Careers Survey (Qualifikation und Berufsverlauf, IAB-BIBB).

The BSW is relatively unknown compared to the other datasets. The BSW survey was conducted seven times (1979, 1982, 1985, 1988, 1991, 1994, 1997, 2000 and 2003) by the Federal Ministry for Education and Research (Bundesministerium für Bildung und Forschung); data are provided by the Central Archive for Empirical Social Research, University of Cologne. Each survey year, about 7,000 persons between 19 and 64 years are interviewed orally (this includes employed and non-employed people). The BSW dataset is at present the only regular representative survey containing all kinds of training incidences in Germany.[14] In contrast to the other datasets, questions on training are the focus of this survey. We take the year 1988 as observations before and 1994 as observations after the reform. Questions on job-related training refer to the last 12 months.

The GSOEP is an individual-level dataset with panel structure. It is the largest representative longitudinal study of private households in Germany. The same private households, persons and families have been surveyed annually since 1984. In this dataset we have information on whether a person took part in job-related training in the last three years. Observations before the reform refer to 1989 and observations after the reform to the year 2000. The GSOEP has been conducted since 1984, but questions on job-related training started in 1989 and were only repeated in 1993, 2000, and 2004. We do not use 1993 because in asking for training during the last three years, this wave barely covers the 1992 reform.

The IAB-BIBB data are a representative survey of employed persons, which was conducted in 1985, 1991, and 1998. It focuses on job descriptions and detailed information on

[14] For more information, please see:
http://www.bmbf.de/pub/berichtssystem_weiterbildung_9.pdf

qualification profiles and occupational development. Each survey wave consists of more than 34,000 observations; questions on job-related training refer to the last five years.

Although there are some questions on job-related training in the German Micro Census (Mikrozensus, MZ), this dataset is not suitable for this analysis, because training participation is underrepresented there.[15] As pointed out by Wohn (2007) there are several reasons why training participation in the MZ is underrepresented compared to the BSW training participation. Since the other two datasets (GSOEP and IAB-BIBB) have comparable training incidences to the BSW, we focus on these three datasets in the regression analyses and use the Micro Census data only for descriptive analyses (see Figures 3 to 5).

The choice of datasets is driven by information on job-related training at the individual level both before and after the maternity leave extension of 1992. Because the treatment group comprises all women of childbearing age, actual information on maternity leaves was not required for a dataset to be used here. Nevertheless, problems do arise in the dataset comparison. First, all three datasets measure the outcome variable, job-related training, for a different period of time: the last five years in the IAB-BIBB data, the last three years in the GSOEP, and the last 12 months in the BSW. The second difficulty stems from the needs of our difference-in-differences analysis. Not only does it require training incidence observations before and after the maternity leave extension, but these can only be done properly by focusing on the most drastic reform, that which lengthens maternity leave from 18 to 36 months. However, the post-1992 reform surveys differ enormously in timing: 1994 for the BSW, 1998 for the IAB-BIBB, and 2000 for the GSOEP. Variation also exists in the timing of the pre-1992 reform surveys, which refer to the following years: BSW, 1988; GSOEP, 1989; and IAB-BIBB, 1991. Obviously, these differences must be taken into account. For example, by asking for training in the five years previous to 1991, the prereform survey refers to a period during which three smaller extensions of maternity leave benefits occurred (see the

[15] In our analysis, training participation in the Micro Census was only less than half as high as in the other three datasets.

grey-shaded boxes in Figures 3 to 5). The surveys also differ somewhat in their sample sizes, with the largest, the IAB-BIBB, containing more than 16,000 observations per wave. GSOEP and BSW are smaller, the former with over 2,700 observations in 1989 but more than 5,000 in 2000 because of refreshment samples, and the latter with more than 3,000 and 2,000 observations before and after the reform, respectively.

Chapter 2:

The Effects of a Sick Pay Reform on Absence and on Health-Related Outcomes *

Joint work by
Patrick A. Puhani and Katja A. Sonderhof

* This chapter is based on an earlier discussion paper version: IZA Discussion Paper (2009), No. 4607.

Acknowledgements: This research was supported by the German Research Foundation (DFG) within the project 'Labour Market Effects of Social Policy' which is part of the research initiative 'Flexibility in Heterogeneous Labour Markets'. We are grateful to Christian Dustmann, John Heywood, Jeff Smith, Fan Wu, two anonymous referees and seminar participants at Cornell University; EEA (Barcelona); ERMES, University of Paris II; Hebrew University; Tel Aviv University; University of Hannover; University of Magdeburg; University of Michigan; University of Wisconsin-Milwaukee; and ZEW, Mannheim, for helpful comments. All remaining errors are our own.

2.1 Introduction to Chapter 2

Sick pay and disability insurance programs, while reducing exposure to risk and seeking to promote equity through support of people in need, entail moral hazard problems. Several studies on U.S. and Canadian disability schemes find negative labor supply and positive take-up effects of increased benefit generosity as well as effects of economic conditions on benefit take up (Black, Daniel and Sanders (2002), Gruber (2000), Johnson and Ondrich (1990), Kreider and Riphahn (2000), Meyer, Viscusi and Durbin (1995), Neuhauser and Raphael (2004), on screening, see Campolieti (2004)). Gruber (2000) stresses that the U.S. disability program is 'one of the largest social insurance programs' with an expenditure that amounts to $46 billion, which was 0.13 percent of U.S. GDP in 1998.

Unlike the U.S. or the U.K., several continental European countries require employers to provide sick pay from day one of each sickness spell. In Germany, Europe's largest economy, sick pay is 100 percent of the wage for the first six weeks of sickness. Combined with the high level of employment protection typical of many continental European economies, these regulations make absence hard to sanction.[16] As a consequence, presence at the workplace is – at least in the short run when promotion is disregarded – a form of voluntary cooperation by the worker. Absence from work carries a high cost in terms of workdays lost, with rates ranging from 2.0 percent in the U.S. to 4.2 or 7.2 percent in continental European countries like Germany or France, respectively (Osterkamp (2002)). If the cost of sick pay regulations in Germany were compared to the U.S. disability program, a back-of-the envelope calculation would dwarf the size of the U.S. disability program in terms of percentage of GDP spent in the respective country: if labor contributed two-thirds to the GDP, a reduction in working days lost from the German to the U.S. level would raise the

[16] In Germany and Sweden, a worker can remain absent from work for 2 and 7 days, respectively, without a physician's certificate (Johansson and Palme (2005), Riphahn and Thalmaier (2001)).

GDP by about 2.2 x (2/3) = 1.5 percent, more than eleven times the cost of the U.S. disability program in percentage of U.S. GDP.[17]

In this study, we extend the recent literature using natural experiments to estimate the effects of incentives on absence (Henrekson and Persson (2004), Ichino and Riphahn (2005), Johansson and Palme (2002, 2005), Riphahn (2004), Riphahn and Thalmaier (2001)). However, unlike previous studies, we do not only consider absence from work as an outcome, but also estimate the reform's effects on health-related outcomes like the duration of hospital stays and subjective health indicators and show that moral hazard problems of sick pay extend to inefficient use of the medical system. We further add to the literature by estimating the effects of the introduction (switch on) and the repeal of a reform (switch off) that reduced sick pay in Germany from 100 to 80 percent of the wage. Methodologically, because this reform affected only workers not covered by collective bargaining contracts, we can apply a difference-in-differences identification strategy to German Socio-Economic Panel data so as to distinguish the effects of the reform from time- or group-specific effects. Fixed-effects regressions provide an additional control for unobserved individual heterogeneity.

The relationship between financial incentives and absence is amply shown in earlier papers using regression analysis on observational data (i.e., without natural experiments). Fewer studies, however, use natural experiments to relate the *cost of absence* to its incidence or duration. Ichino and Riphahn (2005), Riphahn and Thalmaier (2001), and Riphahn (2004) exploit probationary periods or time to reach virtually 'undismissable' status as a natural experiment which leads to high employment protection. The authors find that absence rates increase with employment protection in Italy and Germany. For Sweden, Henrekson and Persson (2004) use time series data for 1955–1999 to show that reforms that make sick pay more generous increase absence from work and vice versa. Likewise, Johansson and Palme

17 Admittedly, this number may be somewhat lower if genuinely sick employees going to work are not only less productive but may also decrease the productivity of others through infection, yet 1.5 percent of the GDP is a large enough number to illustrate the potential importance of policies affecting workers' absence.

(2002, 2005) use person-level data to evaluate the Swedish sick pay reform of 1991, which resembled that investigated here for Germany but applied only to blue-collar workers. The authors identify reactions to the incentives created by the reform: both the incidence and the duration of absence decreased when the cost of absence increased.

This present paper investigates the case of the late 1996 German reform that reduced sick pay from 100 to 80 percent during the first 6 weeks of sickness for workers without collective bargaining contracts. However, unlike previous studies using natural experiments to evaluate the incentives linked to sick pay, we can also evaluate the effects of the early 1999 repeal of the reform, which re-set sick pay to 100 percent of the wage rate from day 1. Apart from absence from work (the outcome considered in the previous literature), we also evaluate the reform's effects on use of the medical system and on subjective health indicators. Specifically, we find that for workers aged 20 to 55 years who remained with their firm during the estimation period, the average number of days absent from work fell by 2.4 days per year (according to a fixed-effects estimate). Furthermore, we show that the reform particularly reduced long durations of absence and that part of this decrease (0.4 days) coincides with a reduction in the average number of days spent in hospital, although we cannot find any robust effects of the reform on subjective health indicators. The results also indicate that the switch-on effects of the reform might be slightly smaller than the switch-off effects on absence from work. However, this difference is not statistically significant. Altogether, it seems that the reform reduced the – in international comparison – long and frequent contacts of Germans with their health care system. These contacts are costly both for employers and the health care system, but their reduction due to the sick pay reform seemingly had no statistically robust negative effects on subjective health indicators or in terms of long-term sickness. We also find no negative effects of the reform on liquidity constraints, measured as the perception of financial security in case of sickness.

2.2 Sick pay in Germany

Germany has one of the most generous sick pay regulations among industrialized countries. German federal law dictates that employees reporting sick are entitled to 100 percent of their pay for the first 6 weeks of sickness, to be paid by the employer (Bundesministerium der Justiz (2003), Schmitt (2005)). Only after this period does the percentage reduce to the 70 percent covered by mandatory health insurance (Bundesministerium der Justiz (2008)).[18] Moreover, in contrast to regulations in the U.S., the U.K. or Switzerland, German federal law regulates sick pay for the first few days of illness (Osterkamp (2002)).

As of October 1, 1996, the Christian Democrat and Liberal coalition government reformed the federal law regulating sick pay in Germany so that all employees (whether blue- or white-collar) were entitled to only 80 percent (rather than 100 percent) of their previous wage from day 1 of sickness through the first 6 weeks of absence (Schmitt (2005)).[19] This law, however, was heavily resisted by the trade unions, which prior to 1970 had fought for years to gain 100 percent sick pay for all workers. Hence, the implementation of the new law was followed in 1996 and 1997 by a plenitude of lawsuits (each referring to a particular collective bargaining contract) in which the unions argued that collective bargaining contracts based on the old version of the law were still valid and implied sick pay corresponding to 100

[18] Some employees are subject to more generous sick pay rules arrived at through collective bargaining agreements. For example, public sector employees already in place before July 1, 1994, receive sick pay of 100 percent of their wage for more than 6 weeks depending on their tenure (9, 12 15, 18 and 26 weeks for 2, 3, 5, 8 and 10 years of tenure, respectively). For public sector employees hired after this date, the 6-week rule applies (Clemens et al. (2006)). However, after the first 6 weeks, public sector employers must pay an additional allowance into the 70 percent sick pay covered by the mandatory health insurance. Such allowances in addition to health insurance sick pay after the sixth week of sickness also exist in other sectors of the economy and depend on the specific collective bargaining contract.

[19] Besides reducing the sick pay covered by the employer for the first 6 weeks, the January 1, 1997, changes to the law on mandatory health insurance reduced sick pay from the 7th week onwards (covered by mandatory health insurance) from 80 to 70 percent (Bundesministerium der Justiz (1996)). This type of sick pay is paid for up to 78 weeks within 3 years for a single type of sickness (Bundesministerium der Justiz (2008)). It should also be noted that this reform (the reduction from 80 to 70 percent) had not been reversed by the time of writing. There was also a small reform of hospital stay co-payments. In 1994, co-payments were DEM 14 (€7) per day in Western Germany and DEM 9 (€4.50) in Eastern Germany. In 1997, they were slightly raised to DEM 17 (€8.50) and DEM 14 (€7) in Western and Eastern Germany, respectively. Although the reform of 1997 concurs with the treatment period, the raise of co-payments by €1.50 in Western Germany (€2.50 in Eastern Germany) is minute compared to the cut in wages by 20 percent for each sickness day.

percent of the wage. According to Bispinck and WSI-Tarifarchiv (1997), these lawsuits were generally won. Hence, as of December 1997, over 15 million employees were covered by collective bargaining contracts that guaranteed them sick pay of 100 percent of their wage, which implies full coverage of about 55 percent of all employees (not counting civil servants, who were not affected by the reform). Indeed, according to a 1998 publication by the German Parliament (Deutscher Bundestag), 80 percent of employees were receiving sick pay corresponding to 100 percent of their wage, and the remaining 20 percent were largely those not covered by collective bargaining contracts (Deutscher Bundestag (1998), p. 17).[20]

This group of workers without collective bargaining coverage comprises our treatment group, which we compare to the control group of workers covered by collective bargaining contracts using a difference-in-differences estimation strategy. Nevertheless, some measurement error can be expected in the treatment status for two major reasons. First, some workers in our control group did in fact receive ‘treatment’ because their collective contract did not provide for sick pay covering 100 percent of their wage. Second, more workers received treatment immediately after the reform became effective (October 1, 1996) than by the middle of 1997 or later because it took time for lawsuits to establish that the old rules applied for most workers covered by collected bargaining. Both these sources of measurement (classification) error are likely to lead to an attenuation bias; that is, because estimates of the treatment effect are biased toward zero, the true effects might be larger than those estimates.[21] However, we assessed the second measurement problem by producing estimates using only

[20] We could not find other statistics on the share of employees who still obtained 100 percent of their wage as sick pay. We did contact all major trade unions, but most information they provided referred to regulations in specific contracts rather than statistics on the number of employees covered by different sick pay regimes.

[21] As surveyed in Bound, Bown and Mathiowetz (2001, p. 3725), classification error (measurement error in a binary variable) usually leads to bias towards zero, unless classification error is so prevalent that the sign of the estimate actually changes. The statement by the German parliament (Deutscher Bundestag) above in the text, however, suggests that the overlap between collective bargaining coverage and not being affected by the reform turned out to be almost perfect so that we have to assume that classification error leads to small attenuation bias in our application. In the difference-in-differences context, the classification bias affects the coefficient of the dummy variable of the treatment group indicator. However, the interaction coefficient of interest, namely the coefficient on the interaction between the treatment indicator and the reform period will also be attenuated because only part of the indicated treatment group will actually have been treated. This implies that the true effects are probably even somewhat larger than the ones we estimate.

1998 as the treatment period thus ignoring 1997 (a time of ongoing lawsuits). These estimates were similar to our main results, so we could not find evidence for attenuation bias.

Two years after the late 1996 reduction in sick pay, the right-wing coalition government between the Christian Democrats and the Liberals ended after a regular election installed a left-wing coalition government between the Social Democrats and the Green Party. As a result, on January 1, 1999, only two months after the change of government, the 1996 sick pay reform was repealed. This introduction and then repeal of the reform within such a short period allows us to estimate the effects of reduced sick pay through both the switch-on and switch-off effects of the policy change.

Methodologically, the question arises whether policy endogeneity or anticipation effects may bias our estimates. However, any transitory developments in absence from work or other health related outcomes that might have triggered policy reforms are taken care of by our difference-in-differences estimation strategy (as long as these shocks affected treatment and control groups similarly). Furthermore, an electronic search of a major German newspaper, the Frankfurter Allgemeine Zeitung, for articles on sick pay revealed that a motivation for the reduction in sick pay were not a recent rise in absence rates, but a gradual realization that the German labor market regulations built up over decades had reduced labor market competitiveness. Furthermore, the debate on the reform only heated up after April 6th 1996, when the Minister of Labor proposed changes to sick pay, which is already after our pre-reform years 1994 and 1995. The law was passed on September 13, 1996, only slightly more than 2 weeks before it became effective.

2.3 Data and descriptive statistics

To the best of our knowledge, the German Socio-Economic Panel (GSOEP), in existence since 1984, is the only person-level dataset providing information on both workers' absence from work and worker coverage by collective bargaining contracts. Whereas information on absence, asked as the number of days the worker was absent from work in the previous year, is collected annually, information on a worker's coverage by a collective bargaining contract is only available for the 1995 survey. However, because average tenure in Germany is longer than in the U.K. or the U.S. (in 1998, 10.4 years versus 8.2 and 6.6 years, respectively; Auer and Cazes (2000)), one option for the empirical strategy is to use the 1995 information on coverage by a collective bargaining contract and impute this value for each individual in all other waves. Nevertheless, because an employee may alter the treatment status by changing employer, this procedure may blur the partition of the sample into treatment and control groups to produce a third source of potential attenuation bias in our estimates (see Section 2.2). We therefore restrict the sample to workers who did *not* change employer during the years under consideration (hereafter, 'firm stayers').[22] Specifically, this means that when defining treatment and control groups, we include only workers who responded to the 1995 question on collective bargaining coverage and did not change employer during the 1996/1997/1998 period when reduced sick pay was in place (the treatment period). Appendix 2.1 and Appendix 2.2 detail our selection of the estimation sample for this study.[23]

For the years prior to the reform, we use GSOEP data for 1994 and 1995 (days absent surveyed in 1995 and 1996, respectively) but exclude 1996 data because they could be partly affected by the October 1 implementation of the reform. In addition, 1996 was the beginning

[22] It should be noted that in Germany, in contrast to some other countries, employers agreeing to a collective bargaining contract must apply its terms to *all* workers in the company, not simply to workers that belong to the trade union negotiating the contract. Employers can avoid collective bargaining contracts, however, by leaving the employers' federation. However, if employers had so changed their status, it would be yet another source of attenuation bias.

[23] In order to gauge whether these restrictions generate a selected sample, we regressed indicators for a) being a mover, b) for answering to the question on collective bargaining and on c) leaving the panel survey between 1995 and 1997 (panel attrition is an especially important issue for the fixed-effects estimators) on days of absence in 1995 (before the reform) and other controls. It turns out that a) being a mover and b) answering the question on collective bargaining coverage is not related to pre-reform absence, whereas c) leaving the panel is positively correlated with days of absence. Here, for the age group 20-55, which we mainly focus on, increasing the days of absence by 50 days (50 days is already the 98^{th} percentile of days of absence) raises the probability of leaving the sample by 9.5 percentage points.

of the lawsuits clarifying that previous collective bargaining contracts made the reform ineffective for most workers these contracts covered (see Section 2.2). These exclusions leave 1997 and 1998 as the viable years for examining effects when the reduced sick pay reform was in place (because of the 1997 lawsuits, we also check the sensitivity of our results when only 1998 is considered as the treatment year). Because the reform was repealed on January 1, 1999, GSOEP data referring to the years 1999 and 2000 provide the sample for the post-reform period.

Table 2.1 displays the sample means by reform period (pre-reform, reform, post-repeal) and by coverage by collective bargaining contracts. The sample consists of workers aged between 20 and 64 years who are not self-employed nor students or apprentices. Although the sample size changes across the years due to panel attrition and panel refreshment samples, it is lowest during the reform years because we exclude workers who changed employer during these years. Nevertheless, not only should the rich set of control variables contained in the GSOEP account for attrition based on observables, we also present fixed-effects estimates (see Section 2.4 below) that account for attrition based on unobserved variables as long as their effects are constant over time.

It should be noted, however, that as the outcome variable, we only observe the total number of absence/sickness days in a calendar year, not the number and length of specific sickness spells. Moreover, although the original GSOEP question asks about workdays lost due to illness, the fact that we observe some people reporting sickness durations exceeding the number of working days indicates that the measurement of absence might be a mix of lost workdays and the total number of sick days (including weekends and public holidays).[24]

[24] The Ministry of Health also collects data on absence from the public health insurance system and publishes it on an annual basis. The average absence rates (percent of working days lost due to sickness) are 4.74, 4.79, 5.08, and 4.75, for 1993-1996, respectively, 4.19, 4.13 for 1997 and 1998, and 4.27 and 4.22 for the years 2000 and 2001, respectively. These absence rates, which refer to the population of workers covered by public health insurance (which is the vast majority), correspond to what we observe in the GSOEP data (based on 253 working days and on all workers: 4.99, 5.06, and 4.78 percent in 1993, 1994 and 1995, respectively. Then during the reform these shares are 4.20 and 4.64 percent in 1997 and 1998 and after the reform 4.39 and 4.29 percent in

As illustrated in the upper part of Figure 2.1, the average number of days absent differs between treatment (not covered by a collective agreement) and control (covered workers) groups, indicating that the former generally report fewer days of absence. This finding holds true before, during and after the reform, except for workers under 40 following repeal (see the lower part of Figure 2.1). Nevertheless, the raw means also suggest that the reform did have an effect on absence. That is, whereas the absence gap between treated and control observations prior to the reform was -3.4 days (8.8 days for workers without coverage and 12.2 days for covered workers), this gap widened to -4.7 days during the reform years only to shrink again to -2.0 days after its repeal. The rise and fall of this gap between treated and control observations is even more pronounced when the analytical focus shifts to younger workers. As the lower part of the figure illustrates, younger workers (below 55 or 40 years of age) seem to have reacted more strongly to the reform. For treated workers younger than 55 years of age, the average number of days absent decreased from 8.1 to 6.4 days during the reform period but rose to 9.3 days following repeal. The change in the gap between treated and control observations is even more pronounced, moving from -3.3 pre-treatment to -5.6 during treatment and down to -1.4 after treatment (repeal). For workers younger than 40, these averages are -2.5, -4.3 and +0.9, respectively.

Although these numbers represent raw gaps that do not take observed or unobserved heterogeneity into account, they nevertheless suggest that the reform did have an effect on workers, especially those younger than 55 years of age. Older workers, in contrast, are likely to be less credit constrained and may thus be less sensitive to reduced sick pay. Their absence may also be more strongly driven by genuine health concerns and hence less influenced by financial incentives. We therefore conduct an analysis of the treatment effects for all workers (aged 20 to 64) and then examine restricted age groups.

1999 and 2000, respectively. So despite the slightly different populations, the GSOEP and the Ministry of Health data seem to indicate similar absence rates.

As Table 2.1 shows, treatment and control groups not only differ systematically in their average number of days absent but also in other characteristics. For instance, the treatment group earns lower hourly wages than the control group (by between 4 and 9 percent, depending on the period considered).[25] Moreover, although both groups have roughly the same average age, gender, civil status and health indicators, the treatment group is somewhat more educated and somewhat less likely to be blue collar or work part time.[26] However, the most striking differences between the treatment and control groups are in terms of tenure, firm size, industry and civil service status. That is, workers without collective bargaining coverage (the treatment group) have lower tenure; work in smaller firms; are much more likely to work in services like trade, real estate and business activities; and are generally no civil servants.[27] These differences between the two groups persist across the observation period: there are no *major* compositional changes between the covered and uncovered groups across time. Nevertheless, the regression analysis reported below controls for any compositional changes related to observed or time-constant unobserved characteristics.

Table 2.2 displays the distribution of the outcome variable, the annual number of days a worker was absent from work. In almost all periods, the 4th decile of the absence days' distribution is 0 or 1, meaning that almost half the workers are not absent from work for a single day. Moreover, even though the median number of days absent is 2 in the treatment group and 4 or 5 among the controls, the distribution is highly skewed to the left with the 7th decile between 6 and 12 days, the 9th decile between 20 and 30 days and the 99th percentile

---

[25] Based on the assumption that reduced sick pay might lead to lower (efficiency) wages and thus might have both a direct effect on absence and an indirect effect through the wage rate, we estimated the effects of the sick pay reform on regular wages using standard difference-in-differences models with control variables and fixed-effect estimates. However, contrary to what efficiency wage theory might predict, all estimates of wage effects are insignificant, with most point estimates positive. This observation is consistent with experimental evidence in Dürsch, Oechsller and Vadovic (2008) who find barely any effort response by workers to sick pay. In our study, both these results support the interpretation that changes in the raw wage gap between treatment and control groups can be explained by compositional effects.

[26] Detailed information on the variables contained in the GSOEP is provided in Haisken-DeNew and Frick (2001).

[27] Treated observations indicating that the individual is a civil servant most probably represent classification error in the civil service or the collective bargaining status in the original GSOEP data. The coding error affects only 2 percent of the sample assigned 'treatment' status. We remain conservative by keeping the data as they are, because if these observations were in fact controls, this classification error would generate attenuation bias.

at 98 or more days. Thus, our estimation strategy must take into account the heavy censoring of the outcome distribution at zero.

2.4 Effects of the sick pay reform on absence from work

We begin by estimating linear difference-in-differences models with the following specification:

$$absence_{it} = \alpha + \beta_1 X_{it} + \beta_2 reform_t + \beta_3 nocoverage_i + \tau(reform_t \times nocoverage_i) + \varepsilon_{it} \tag{1}$$

where *absence* denotes the number of days of absence and *reform* is a dummy variable indicating the time period during which the reduced sick pay reform was in place (1997 and 1998) and valued at zero pre-reform and post-repeal. Likewise, *nocoverage* is a dummy variable indicating that a worker was *not* covered by a collective bargaining contract (the treatment group). This *nocoverage* indicator controls for differences in absence rates between the treatment and control groups, which in the absence of any reform are assumed to be constant across time (the identifying assumption of the difference-in-differences estimator).

Time-specific variations in absence affecting both groups similarly are controlled for by the *reform* dummy as well as further time effects.[28] The difference-in-differences estimator is given by τ, which indicates the change in the absence differential between treatment and control groups after sick pay was reduced from 100 to 80 percent. Specification (1) includes no control variables. However, specifications (2)-(4) stepwise add control variables to allow for compositional changes in the two groups across time and improve the efficiency of the

28 These time effects control for macro shocks as long as these affect treatment and control groups similarly. Separate time effects for treatment and control groups would make the treatment effect unidentified. However, because treatment and control groups are distributed differently across industries, we could allow for industry-specific time and treatment effects and thereby allow for different effects of the macroeconomy on treatment and control groups. Because sample sizes shrink too much to estimate the industry-specific effects precisely, we have compared the average treatment effects on the treated, obtained as a weighted average of industry-specific treatment effects, with the main estimates reported in this paper. The differences in the point estimates were mostly minor.

difference-in-differences estimator as long as they can be regarded as exogenous. The first group of variables, included in specification (2), are the regional unemployment rate, the log hourly wage, age, civil status (married, children), gender and some interactions between them. We employ these standard controls from the absence literature because of their likely impact on the incidence of sickness through their effect on the benefits and costs of shirking through absence. Specification (3) then integrates a further set of controls by including education, citizenship, and job and firm characteristics, as well as a dummy for West Germany (see Table 2.1 for details). The full specification (4) adds a last set of controls that refer to '*health at present*' and '*satisfaction with health*' as asked in the GSOEP. If respondents answer these health-related questions truthfully irrespective of their potential shirking behavior and if the reform had no impact on health (see Section 2.5), these variables are valid controls; otherwise, they are endogenous.

Table 2.3 shows the results for the OLS difference-in-differences estimator when both the pre-reform and post-repeal period are simultaneously included as the reference period (here and in the following, we use robust standard errors clustered at the person level). Hence, this phase of the analysis does not distinguish between the reform's switch-on and switch-off effects but rather compares the difference between the treatment and control groups during the reform with that before or after its repeal. This approach increases the sample size and hence the precision of the estimates.

To check the sensitivity of the estimates with respect to the control variables just described (the coefficients of the control variables are reported in Appendix 3.3), in Table 2.3, we report the estimated treatment effects for specification (1) through (4). In this table, we also report the marginal effects at the mean of the difference-in-differences estimates of a count-data model, which is expected to fit the data better because of the dependent variable's count nature. We use the negative binomial model, NEGBIN, where the right hand side index of equation (1) enters an exponential function to model the expected value of absence days;

for an application in the context of absence, see Winkelmann (1999); technical descriptions of the NEGBIN II model that we apply here can be found in Cameron and Trivedi (1998, p. 70ff.) or Winkelmann (2008, p.134ff.); nonlinear difference-in-differences models are discussed in Athey and Imbens (2006). The parameters reported here are the incremental effects of the treatment indicator (the interaction term) at the data mean.[29]

However, because the restriction of the sample to firm stayers may cause systematic attrition not only based on observed characteristics (which we control for in the OLS and NEGBIN estimates) but also based on unobserved characteristics, we control for unobserved heterogeneity by reporting (linear) fixed-effects estimates. More specifically, we identify the policy reform effect using only the 'within individual variation', because the fixed-effects estimator effectively assesses the response to the reform by considering only treated and control individuals observed both *during* the reform *and* in a period without reform.

In terms of the estimate's sensitivity to the inclusion of control variables, controlling for compositional changes matters mostly for the estimated standard errors (and hence the statistical significance of the estimates). The point estimates are rather similar across specifications. Therefore, below we report estimates from the full specification (4) that have lower standard errors.

As regards the different modeling strategies, the differences in the point estimates between OLS and NEGBIN are minor, meaning that despite its theoretical deficiencies, the OLS model seemingly provides a very good approximation of the treatment effect at the mean. However, not surprisingly, most standard errors are somewhat smaller in the NEGBIN model, which is tailored to fit the count data. In addition, even though both the OLS and NEGBIN models suggest that the decrease in sick pay reduced the number of absence days per year by 2, this effect is only statistically significant in the NEGBIN model. Once we

[29] Ai and Norton (2003) derive a correct presentation of the cross derivative and cross difference in nonlinear models with interaction terms. However, this cross difference is not equal to the treatment effect shown in Puhani (2008).

control for unobserved heterogeneity in a (linear) fixed-effects model, the point estimate becomes somewhat smaller (and insignificant) with an estimated reduction of 1.2 days in specification (4). Given that for the treated group the mean number of days absent was 8.8 before the reform, this figure amounts to a reduction in absence of between 14 and 24 percent, which is sizable.

As hinted at in the previous section, based on the raw data, we might expect the effect of the reform to be higher among younger workers. Therefore, in Table 2.4, we provide the difference-in-differences estimates for the different age groups (here and subsequently, for full specification (4) with all control variables). We find that in all models, the point estimates become larger when older workers (aged 56–64) are excluded. More specifically, for the age group 20–55, the fixed-effects estimate shows a significant 2.4 reduction in days absent per year. For the further restricted age group 20–40, at 2.2 days, this reduction is somewhat smaller. We do not provide separate point estimates for older workers, because the precision of these estimates is too low (for the age group 56-64, the standard errors range between 4 and 6 days, so that all estimates are statistically insignificant; the point estimates even change sign). Similarly, if we estimate the effects separately for men and women, the estimates become too imprecise to reach firm conclusions on gender differences: point estimates are negative for both genders and there is no clear pattern for whom the point estimates are larger.

2.4.1 Effects at different points of the distribution

The skewed nature of the distribution of absence days, with a high probability mass at zero, raises the question of whether the reform had a larger effect on longer or on shorter durations of absence. To answer this question, we describe the reform's effect at different parts of the distribution by difference-in-differences quantile regressions (for a technical description see Athey and Imbens (2006), p. 446f.; another application is Song and

Manchester (2007), Koenker (2005), Chapter 2, provides a general introduction to quantile regression). Quantile regression difference-in-differences implies stronger identifying assumptions than does OLS because we must assume that the differences in the distributions (not simply the differences in the means) between treatment and control groups would have remained constant in the absence of reform. Hence, in Table 2.5, we show difference-in-differences quantile estimates by decile, again for the three age groups sampled. Theoretically, the OLS estimate is the mean of the coefficients at all quantiles; however, as the results show, up to the 4th decile, the effect is (virtually) zero. This finding is not surprising given that around 40 percent of all workers in the sample report not having been absent for a single day. The point estimates of the reform's effect on days absent then grow even more negative with increasing deciles. For all workers (i.e., aged 20–64), by the 9th decile, the point estimate is a statistically significant 4.8 days reduction in absence.[30] It should also be noted that the 9th decile of the number of days absent for the treatment group is 23, which corresponds to a sizable reduction in absence of more than 20 percent. In other words, the quantile regressions reveal that it is mainly absence durations of several weeks (cumulated over the year) that are reduced by the reform.

Once the sample is restricted to workers aged 20–55 or 20–40, we obtain statistical significance from the 7th decile onwards, with 7th decile estimates of -0.8 in both cases. Given that the 7th decile of absence days in the treatment group before the reform was 8 days, this figure constitutes a reduction of almost one tenth. The reduction becomes larger at higher deciles (both absolutely and relatively) for the group aged 20–55. In the 20–40 age group, the point estimates at the very high quantiles (95th and 98th) are the largest of all quantiles but are statistically insignificant because of the large standard errors associated with the sensitivity to outliers of quantile regressions for extreme quantiles of the distribution.

[30] The displayed quantile regression estimates, obtained using the econometric software package Stata 10, take sampling weights into account. Standard errors reported for the quantile regressions do not allow for clustering; however, we find that block-bootstrapped standard errors that do take clustering into account differ little from the asymptotic standard errors ignoring clustering in an unweighted regression.

In results not shown here both OLS and probit estimates cannot detect any effect of the reform on the incidence of absence (i.e. a binary indicator of whether a person has been absent for at least one day or not). This is in line with the quantile regression estimates which do not find any effect near the median of the days of absence distribution.

2.4.2 Switch-on and switch-off effects

Because our dataset includes information on absence before and during the reform and after its repeal, we can estimate the effect of the reduction in sick pay (switch on) separately from the subsequent repeal and increase in sick pay (switch off). Doing so has two advantages. First, the difference-in-differences approach used here relies on the identifying assumption that in the counterfactual absence of the reform, the gap in absence days between treatment and control groups would have remained constant. One reason for violating this assumption would be another incident or reform of which the researcher is unaware that might have had a differential impact on absence days for both groups. To dissipate such doubts, research designs that introduce and subsequently take back a reform are very helpful. If both effects have similar values and both indicate that absence is lower with lower sick pay, we can have more confidence that the effects estimated are genuinely caused by the sick pay reform.

In fact, the above-mentioned estimates do not distinguish between the pre-reform and post-repeal periods, which implies that the introduction of the reform has the same impact on absence (in absolute terms) as its repeal. To check this assumption, we estimate the effects of the reform’s implementation (switch on) and repeal (switch off) separately. Table 2.6 presents the switch-on and switch-off estimates separately by age group based on data for the years 1994, 1995, 1997 and 1998 for the switch-on effects and for the years 1997, 1998, 1999 and 2000 for the switch-off effects. Table 2.7 then reports the corresponding quantile regression estimates. The models are specified so that a negative estimate always implies that, as expected, absence is lower during the period of reduced sick pay.

In Table 2.6, all switch-on and switch-off point estimates are negative in absolute value. In the NEGBIN model, all coefficients are statistically significantly different from zero. Hence, we argue that the reform had a genuine effect on days absent. Interestingly, however, when we compare the absolute size of the switch-on and switch-off effects in Table 2.6, we find that in all cases (except for the age group 20-64 in the fixed-effects estimates) the switch-off point estimates are larger than the switch-on effects (this holds at virtually all quantiles, as shown in Table 2.7). However, this difference is only statistically significant in OLS specifications that neither take into account the count data nature of the outcome variable nor unobserved heterogeneity. In the NEGBIN model the point estimate is an insignificant 0.44 days larger for the switch-off effect compared to the switch-on effect, but this difference amounts only to insignificant 0.14 days in the fixed-effects model for workers aged 20-55. Because these differences are not statistically significant, one choice would be to ignore them; however, the difference becomes larger – albeit insignificant – for the 20–40 age group at 1.47 days (marginally insignificant at the 10 percent level) in the NEGBIN, and at 1.37 (insignificant) in the fixed-effects model.

One possible basis for interpreting these larger switch-off point estimates is the experimental and psychological literature on extrinsic versus intrinsic motivation. According to Pinder (2008, p. 81), intrinsic motivation, roughly defined, relates to ‘behavior that is performed for its own sake rather than for the purpose of acquiring any material or social rewards’. The fact that the switch-off effects are larger than the switch-on effects is congruent with experimental evidence from Gächter, Kessler and Königstein (2007) that incentive contracts negatively impact voluntary cooperation among individuals, and that these negative effects persist even after the incentive contract is repealed. It also relates to an ongoing debate in the psychological literature on whether extrinsic motivation may crowd out intrinsic motivation (Pinder, 2008, p. 86ff.). That is, because sick pay before the reform was 100 percent, showing up for work in Germany had an aspect of voluntary cooperation, at least for workers not seeking promotion, and such voluntary cooperation can be linked to intrinsic motivation. Reduced sick pay then added an element of immediate extrinsic motivation that was abolished after the reform was repealed. Hence, in light of Gächter, Kessler and Königstein's (2007) findings, the experience of extrinsic motivation may have crowded out some intrinsic motivation even after repeal. Nevertheless, the extrinsic motivation of reduced sick pay during the reform period was effective in reducing absence, which supports the economist's paradigm that people react to incentives.

2.4.3 Placebo estimates and estimates by calendar year

Because we have two years of observations for each 'regime' (pre-reform, reform, after repeal), we can in theory test the identifying assumption of the difference-in-differences estimator by, for example, testing whether a 'placebo treatment effect' estimate for the year 1995 (pre-reform) with 1994 as the base year (also pre-reform) is equal to zero. In Table 2.8, we therefore define 1994 as the base year and estimate 'treatment effects' for all further years used in the previous estimates: 1995 (pre-reform), 1997, 1998 (both reform) and 1999, 2000 (both post-repeal). If the difference-in-differences identifying assumption is correct, only the grey-shaded coefficients (reform period) should be different from zero.

Table 2.8 shows that standard errors become very large when estimating treatment effects by calendar year (most standard errors are between 1 and 2 days, some are even larger), so that hardly any coefficient is statistically significant. Still, larger negative coefficients are (with few exceptions) observed mainly during the reform period. After the repeal of the reform, some point estimates turn quite large and positive, especially for the year 2000 (two of them even significant), but the standard errors are large as well. We cannot determine whether this finding is due to crowding out of intrinsic motivation as mentioned in the previous subsection or due to a violation of the identifying assumption. In general, we observe a clear decrease in absence with the onset of the reform period and a subsequent increase after the repeal of the reform. This holds both across the defined age group samples and across estimation methods (OLS, NEGBIN, and linear fixed-effects).

2.5 Effects of the sick pay reform on other health-related outcomes

Although reduced sick pay decreased absence from work, it remains unclear whether this means that the reform was beneficial from a welfare perspective. In this section, we show that the reform surprisingly even reduced the average number of days spent in hospital. This suggests that at least part of the absence reduced by the reform was genuinely health related. However, we also show that the reform was not associated with a significant reduction in indicators of subjective health or long-term sickness. Taken together, the reform might have reduced inefficient use of the health care system.

According to the OECD, in 1995, just before the sick pay reform, health expenditure in Germany as a percentage of the GDP was 10.1 percent, higher than in the U.K. (6.9 percent) but lower than in the U.S. (13.3 percent). Life expectancy at birth, however, was rather similar in these three countries (between 75.7 and 76.7 years). The number of doctor visits per year was highest in Germany (6.4), followed by the U.K. (6.1) and the U.S. (3.3); by 2003, these gaps had become even larger, at 7.6, 5.2 and 3.9, respectively. The average number of hospital stays per person was 0.18, 0.21 and 0.12 and the average length of stay for acute care was 11.4, 7.1 and 6.5 days for Germany, the U.K. and the U.S., with Germany having by far the longest average duration of acute care stays. Hence, contacts with the medical system are seemingly more frequent and longer in Germany. Because these OECD figures (for the whole population) correspond roughly to the sample means in the GSOEP (for a sample of workers aged 20–64), we consider three further outcome variables: the number of doctor visits in the last 3 months (asked in the GSOEP), number of days in hospital (including zeros) and number of hospital stays (see Table 2.9 for the sample distributions). We then go on to investigate the reform's effects on two subjective health indicators and an indicator for long-term sickness before we conclude by investigating the reform's effects on satisfaction with financial security in case of sickness.

2.5.1 Effects on usage of the health care system

In Table 2.10, we report difference-in-differences estimates for the three outcomes concerning usage of the health care system by age group. Not only are all point estimates negative, but those for number of days in hospital and number of hospital stays are all statistically significant. Moreover, the fixed-effects estimates for these two variables are similar to the OLS results, implying that the OLS findings are not driven by unobserved heterogeneity. Because of the extreme extent of censoring of the hospital visit outcome variables, we place special emphasis on the NEGBIN estimates. The NEGBIN point estimates are somewhat smaller in absolute value; however, they still remain economically and statistically significant. For the 20–64 age group, compared to a pre-reform treatment group average of 1.35, the reform reduced the average number of days in hospital by 0.41 days (almost one third, 30 percent) on average. Given that it also reduced the number of stays by an estimated 0.045 (41 percent) at least part of the reduction in the number of days hospitalized is explained by the actual elimination of some hospital stays. Although these estimates may seem large, they can be made plausible by doctors' incentives given the low occupancy rates of hospital beds: these ranged between only 76 and 82 percent in Germany in the period 1996 through 2006 (data from the German Hospital Society, Deutsche Krankenhaus Gesellschaft).

2.5.2 *Effects on health indicators*

Given the reform's effects on absence from work and hospital stays, we ask whether the estimated reductions had a detrimental effect on health. Hence we use the two subjective health indicators asked in the GSOEP (*health at present* and *satisfaction with health*) as outcome variables. Subjective health measures have been critically discussed in the literature. On the one hand, economists usually postulate that each person should be the best judge of his or her utility and this may also be true for health (Dolan (2000), p. 1732). In a literature survey, Idler and Benyamini (1997) find that 'global self-rated health is an independent predictor of mortality in nearly all of the studies, despite the inclusion of numerous specific health status indicators' (p. 21). On the other hand, inter-person comparisons of self-rated health seem to be plagued by people's adaptation to changing states of health as well as changing reference groups over the life cycle and with changing health (Groot, 2000).

In addition to subjective health, we check whether the reform had an impact on the incidence of continuous sickness spells lasting for at least six weeks. This is the only indicator in the GSOEP that we could find as an objective proxy for serious illness. The subjective health measures are recorded on Likert scales and have been normed to range between 0 and 1. Control variables are the same as in specification (3) of Table 2.3. We report OLS and fixed-effects estimates, for the whole sampling period and separately for switch-on and switch-off effects. The reform's effects on the subjective health indicators are presented in Table 2.11.[31]

In the table, most of the estimates are statistically insignificant. For *health at present*, two of the switch-off fixed effects estimates are negative and statistically significant at the 10 percent level. However, neither the corresponding estimates for *satisfaction with health* nor the corresponding switch-on estimates are statistically significant. For *satisfaction with health*

[31] We use the same sample as for the estimation of the reform's effects on absence. Hence, in order to be in the sample, a person has to have valid responses in the current and in the consecutive year. The reason is that the information on absence is obtained from retrospective information in the following year's GSOEP questionnaire. Our sample definitions guarantee that the estimates of absence and health effects refer to the same population.

as the outcome, three of the nine OLS estimates are statistically significant and negative, but none of the fixed-effects estimates is statistically significant. Hence, from these estimates, there is no convincing evidence that the sick pay reform had a negative effect on subjective health.

Nevertheless, health may be deteriorating over time: in Table 2.11, we have considered subjective health in the current year (1997 and 1998 for the reform years). In Table 2.12 we investigate whether the reform had an impact on subjective health a year later (so the outcomes for the reform years are measured in 1998 and 1999; for the control years, we also lag the outcomes by one year accordingly). As the table shows, all point estimates are close to zero and none of them is statistically significant. Hence, when considering subjective health indicators for all employees, either in the current or in the following year, there is no robust evidence for the reform to have had any significantly negative effects.

Because the population of employees consists of a lot of people who have not been sick during the entire year, we narrow down the population of interest in the following: first, we consider the reform's effects on subjective health only on employees who state to have visited the doctor at least once during the previous three months (i.e. people who experienced some sort of sickness). Second, we restrict the sample further by considering only employees who have been in hospital during the current year. We then estimate the reform's effects on subjective health for these subpopulations. Again, we distinguish between the effects on subjective health in the current year and between effects on subjective health in the subsequent year to capture any potentially protracted effects. The results for workers who have been to the doctor are presented in Table 2.13 and Table 2.14 and those for workers who have been to hospital in Table 2.15 and Table 2.16, respectively. As can be seen from Table 2.13 and Table 2.14, there is no convincing indication that the sick pay reform decreased subjective health outcomes either in the current or in the following year. There is one negative and statistically significant (at the 10 percent level) coefficient in the fixed effects model: the switch-off effect for the age group 20-40 in the estimate for *health at present*. However, the corresponding estimate for *satisfaction with health* is statistically insignificant. The two OLS estimates which are negative and statistically significant at the 10 percent level in Table 2.13 are close to zero and statistically insignificant when a fixed-effects model is estimated. The one statistically significant fixed-effects estimate in Table 2.14 (effect on *health at present* in the following year) is positive instead of negative and insignificant in the corresponding estimate for *satisfaction with health* as the outcome. When we restrict the sample to workers who have been in hospital (see Table 2.15 and Table 2.16), there are again no statistically significant negative effects on subjective health.[32]

To investigate at least a proxy for an objective health outcome in the GSOEP, we use the question on a continuous sickness spell of 6 weeks or longer. The way this question will be understood by most Germans is referring to being sick as declared by a physician, because for sickness spells longer than 2 days, employees have to provide a doctor's certificate. Note that although we have already shown in the previous section that the reform predominantly reduced longer durations of absence, as demonstrated by the quantile regression estimates, longer duration there meant longer days of absence accumulated over a calendar year, that is, a long duration of absence might be an accumulation of many shorter spells. Here, we look at a *continuous* sickness spell of at least 6 weeks. As Table 2.17 shows, the sick pay reform seems to have decreased, not increased the incidence of long and continuous sickness spells. The estimates indicate a 2 to 3 percentage point reduction in long-term sickness due to the reform and they are statistically significant. One explanation for the reduced incidence of long-term sickness may again be the incentive effects provided by the reform that seem to

[32] Note that in these estimates, the number of observations who are treated during the reform years is reduced to only 29 to 75 workers (depending on the age group considered). This may explain why there is a positive and fairly large estimate of the reform for the age group 20-40 in Table 16. We report no fixed-effects estimates for this age group, because we would only have 10 persons in the treatment group during the treatment period with a within variation in the treatment status that is needed to identify the fixed-effects estimate. For the age group 20-64, inference in the fixed effects estimates is also plagued by the low number of persons with a within variation in the treatment indicator, which is 26 in this case.

have dominated any potentially negative effects on health. To investigate the possibility of negative health effects further, we check whether the incidence of long continuous sickness spells was increased with a delay of one year: these estimates are provided in the lower panel of Table 2.17. Only one of the estimates is statistically significant using OLS (again suggestion a reduction in long-term sickness), whereas the corresponding fixed-effects estimate is virtually zero and statistically insignificant. All other point estimates are also close to zero and statistically insignificant.

Hence, we conclude that there is no convincing evidence that the sick pay reform impaired health outcomes, despite of the fact that it reduced stays in hospital.

2.5.3 Effects on the perception of liquidity constraints in case of sickness

As a last check, we estimate whether the reform changed the employee's 'satisfaction with their financial security in the case of sickness'. Again, this question was asked on a Likert scale in the GSOEP, which we normalize between 0 (very bad) and 1 (very good). During our observation period, this question was only asked in 1997 and 2002, so that we only provide switch-off estimates. The results are presented in Table 2.18. For all workers and for the restricted sample of workers who have visited the doctor in the last three months at least once, none of the point estimates is statistically significant and all point estimates are small. If we restrict the sample further to workers who have been to hospital in the current year, the number of treated persons in the treatment period becomes very small: there are only 59, 49, and 22 such persons for the OLS estimates and only 3, 3, and 1 person with a within variation in the treatment indicator for the age groups 20-64, 20-55 and 20-40, respectively, so that we do not report fixed-effects estimates. We are also cautions in interpreting the estimates based on 59, 49, and 22 treated persons in the sample of people who have been to hospital for the age groups 20-64, 20-55 and 20-40, respectively, and conclude that at least from the sample of people who have been to the doctor as well as for all workers, we cannot find any evidence for negative effects of the reform on the satisfaction with financial security in case of sickness.

2.6 Conclusions of Chapter 2

The economic costs of absence from work can be influenced by economic policy. For example, in contrast to the cases of Switzerland, the U.K. or the U.S., German federal law (as well as statutes in other continental European countries) dictates that employees receive 100 percent of their wages as sick pay from day 1 of their absence spell. However, whereas the literature to date does suggest that such absence is influenced by economic incentives like wages, local unemployment, probationary periods or sick pay, few studies estimate the effects of sick pay on absence by way of natural experiments. Moreover, to the best of our knowledge, ours is the first study to analyze health-related outcomes of sick pay reform and also the first to estimate both the switch-on and switch-off effects; that is, the effects of the reform's implementation and its subsequent repeal by a changed federal government.

The basis of our empirical strategy is a difference-in-differences approach that controls for time and group effects. In some specifications, we also control for unobserved individual heterogeneity by explicitly using the panel nature of the data in a fixed-effects estimation. Overall, we estimate that the reduction in sick pay from 100 percent of the wage to 80 percent decreased absence days by about 2 days per annum on average, which is equal to about one percent of annual working days in Germany (about half the difference between U.S. and German absence rates). As our quantile regressions find, this reduction is primarily driven by a shortening of very long spells. These results are confirmed by separate estimates for switch-on and switch-off effects. Our finding is significant in that if labor contributes two-thirds of the GDP, then the ceteris paribus effect of the reform amounts to an increase in the GDP of about two thirds of a percent through the reduction of absence from work alone.

Besides reducing absence from work, decreased sick pay also reduces reliance on the health care system, which in Germany has almost zero marginal costs to most individuals (the system is mostly public). We find that the reduction in absence due to the reform (by about 2 days) also reduces the average number of days spent in hospital (by not quite half a day, a reduction of 30 percent). Data from the German Hospital Society (Deutsche Krankenhaus Gesellschaft) report hospital costs as a percent of GDP at a fairly steady 2.4 percent. This would imply that the sick pay reform had saved 0.72 percent of GDP through hospital costs alone, and in addition to the two thirds of GDP saved for employers. In sum, the reform might have saved up to 1.38 percent of GDP. Although costs saved might be lower if the reduction in absence and hospital stays referred to less than average productivity days at work and less than average costs per day in hospital, even half a percent of GDP saved would be sizable. The policy relevance of these results is reinforced by the fact that we did not find any remarkable effect of the reform on subjective health indicators nor on long-term sickness, indicating that the reform might have reduced the inefficient use of the medical system.

References for Chapter 2

Ai, C. and E.C. Norton (2003) Interaction terms in logit and probit models, Economics Letters, Vol. 80(1), 123–129.

Athey, S. and G. Imbens (2006) Identification and inference on nonlinear difference-in-differences models, Econometrica, Vol. 74(2), 431–497.

Auer, P. and S. Cazes (2000) The resilience of the long-term employment relationship: Evidence from the industrialized countries, International Labour Review, Vol. 139(4), 379–408.

Bispinck, R. and WSI-Tarifarchiv (1997) Entgeltfortzahlung im Krankheitsfall, Eine tarifpolitische Bilanz ein Jahr nach der gesetzlichen Neuregelung [Sick pay: Taking stock from a collective bargaining perspective one year after the new law]. Elemente Qualitativer Tarifpolitik No. 31, WSI Informationen zur Tarifpolitik, Düsseldorf.

Black, D., K. Daniel and S. Sanders (2002) The impact of economic conditions on participation in disability programs: Evidence from the coal boom and bust, American Economic Review, Vol. 92(1), 27-50.

Bound, J., Ch. Brown and N. Mathiowetz (2001) Measurement Error in Survey Data. In Heckman, J.J., Leamer, E. (eds.) Handbook of Econometrics, Vol. 5, Amsterdam: Elsevier.

Bundesministerium der Justiz [Federal Ministry of Justice]. (1996) Bundesgesetzblatt, Teil I [Federal Law Journal, Part I], No. 55, Bundesanzeiger Verlagsgesellschaft, Bonn.

Bundesministerium der Justiz [Federal Ministry of Justice:]. (2003) Gesetz über die Zahlung des Arbeitsentgelts an Feiertagen und im Krankheitsfall (Entgeltfortzahlungsgesetz) [Law on Payment on Public Holidays and in the Case of Sickness (the Continuation of Pay Act)], Retrieved January 29, 2009 from http://www.gesetze-im-internet.de/bundesrecht/entgfg/gesamt.pdf

Bundesministerium der Justiz [Federal Ministry of Justice]. (2008) Sozialgesetzbuch (SGB) Fünftes Buch (V) [Social Law, Book 5], Retrieved January 29, 2009 from http://www.gesetze-im-internet.de/bundesrecht/sgb_5/gesamt.pdf

Cameron, A.C. and P.K. Trivedi (1998) Regression Analysis of Count Data, Cambridge University Press, Cambridge.

Campolieti, M. (2004) Disability insurance benefits and labor supply: Some additional evidence, Journal of Labor Economics, Vol. 22(4), 863-889.

Clemens, H., O. Scheuring, W. Steingen and F. Wiese (2006) Kommentar zum Bundes-Angestelltentarifvertrag (BAT) [Comment on the collective bargaining contract of federal employees], R. Boorberg Verlag, Edition Moll.

Deutscher Bundestag [German Parliament], (1998) Drucksache 14/45 (circular), Available at http://dip.bundestag.de/btd/14/000/1400045.pdf

Dolan, P. (2000) The Measurement of Health-Related Quality of Life for Use in Resource Allocation Decisions in Health Care, in A. Culyer and J.P. Newhouse (eds.): Handbook of Health Economics, Vol. 1B, Amsterdam: Elsevier.

Dürsch, P., J. Oechssler and R. Vadovic (2008) Sick pay provision in experimental labor markets, Discussion Paper No. 476, University of Heidelberg.

Gächter, S., E. Kessler and M. Königstein (2007) Performance incentives and the dynamics of voluntary cooperation (mimeo), University of Nottingham.

Groot, W. (2000) Adaptation and Scale of Reference Bias in Self-Assessments of Quality of Life, Journal of Health Economics, Vol. 19(3), 403-420.

Gruber, J. (2000) Disability insurance benefits and labor supply, Journal of Political Economy, Vol. 108(6), 1162-1183.

Haisken-DeNew J.P. and J.R. Frick (2001) DTC Desktop Companion to the German Socio-Economic Panel Study (SOEP), available at http://www.diw.de/documents/dokumentenarchiv/17/diw_01.c.38951.de/dtc.409713.pdf

Henrekson, M. and M. Persson (2004) The effects on sick leave of changes in the sickness insurance system, Journal of Labor Economics, Vol. 22(1), 87–113.

Ichino, A. and R.T. Riphahn (2005) The effect of employment protection on worker effort: Absenteeism during and after probation, Journal of the European Economic Association, Vol. 3(1), 120–143.

Idler, E.L. and A. Benyamini (1997) Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies, Journal of Health and Social Behavior, Vol. 38(1), 21-37.

Johansson, P. and M. Palme (2002) Assessing the effect of public policy on worker absenteeism, Journal of Human Resources, Vol. 37(2), 381–409.

Johansson, P. and M. Palme (2005) Moral hazard and sickness insurance, Journal of Public Economics, Vol. 89(9-10), 1879–1890.

Johnson, W.G. and J. Ondrich (1990) The Duration of Post-Injury Absences from Work, Review of Economics and Statistics, Vol. 72(4), 578-586.

Koenker, R. (2005) Quantile Regression, Cambridge University Press, Cambridge.

Kreider, B. and R.T. Riphahn (2000) Explaining Applications to the U.S. Disability System, Journal of Human Resources, Vol. 35(1), 82-115.

Meyer, B.D., W.K. Viscuse and D.L. Durbin (1995) Workers' compensation and injury duration: Evidence from a natural experiment, American Economic Review, Vol. 85(3), 322-340.

Neuhauser, F. and S. Raphael (2004) The effect of an increase in worker's compensation benefits on the duration and frequency of benefit receipt, Review of Economics and Statistics, Vol. 86(1), 288-302.

Pinder, C.C. (2008) Work Motivation in Organizational Behavior, 2nd edition, Psychology Press, New York.

Puhani, P.A. (2008) The treatment effect, the cross difference, and the interaction term in nonlinear ‘difference-in-differences’ models, IZA Discussion Paper No. 3478, Bonn.

Osterkamp, R. (2002) Dice Reports: Work lost due to illness – An international comparison, CESifo Forum, Vol. 4, 36–40.

Riphahn, R.T. (2004) Employment protection and effort among German employees. Economics Letters, Vol. 85(3), 353–357.

Riphahn, R.T. and A. Thalmaier (2001) Behavioural effects of probation periods: An analysis of worker absenteeism, Jahrbücher für Nationalökonomie und Statistik, Vol. 221(2), 178–201.

Schmitt, J. (2005) Entgeltfortzahlungsgesetz, Kommentar [The Continuation of Pay Act: Commentary], Verlag C.H. Beck, Munich.

Song, J.G. and J. Manchester (2007) New evidence on earnings and benefit claims following changes in the Retirement Earnings Test in 2000, Journal of Public Economics, Vol. 91(3-4), 669–700.

Winkelmann, R. (1999) Wages, firm size and absenteeism, Applied Economic Letters, Vol. 6(6), 337–341.

Winkelmann, R. (2008) Econometric Analysis of Count Data, 5^{th} edition, Springer Verlag, Heidelberg.

Tables and Figures of Chapter 2

Table 2-1: Sample means by year and treatment status

	Treated			Control		
	1994/95	1997/98	1999/00	1994/95	1997/98	1999/00
Days absent	8.8	8.0	9.7	12.2	12.7	11.7
Hourly wage	2.47	2.65	2.65	2.56	2.69	2.70
Regional unemployment rate	10.6	12.6	11.3	10.4	12.4	11.3
Civil status indicators						
Age	38.8	42.9	41.6	40.5	43.5	42.7
Married	0.60	0.64	0.57	0.63	0.68	0.65
Female	0.44	0.45	0.48	0.42	0.42	0.41
Children younger than 16	0.38	0.35	0.35	0.37	0.35	0.34
Female × children younger than 16	0.15	0.16	0.16	0.14	0.12	0.11
Female × married	0.25	0.28	0.26	0.25	0.27	0.26
Educational attainment						
Higher education - University degree	0.13	0.14	0.20	0.12	0.12	0.15
Higher education - no degree	0.21	0.19	0.22	0.24	0.26	0.26
Apprenticeship	0.53	0.54	0.48	0.48	0.48	0.46
No apprenticeship	0.13	0.12	0.11	0.15	0.15	0.13
Job and firm characteristics						
Temporary work contract	0.04	0.02	0.06	0.06	0.03	0.05
Working fulltime	0.80	0.81	0.79	0.85	0.86	0.85
Blue-collar worker	0.32	0.26	0.28	0.38	0.37	0.35
White-collar worker	0.67	0.72	0.70	0.50	0.51	0.54
Civil servant	0.02	0.01	0.02	0.12	0.12	0.11
Citizenship/region						
German	0.92	0.93	0.94	0.92	0.91	0.91
West Germany	0.77	0.79	0.80	0.82	0.82	0.82
Firm size						
Firm size (1–19)	0.41	0.40	0.34	0.14	0.12	0.14
Firm size (20–199)	0.32	0.35	0.31	0.28	0.29	0.28
Firm size (200–1,999)	0.15	0.14	0.19	0.26	0.27	0.27
Firm size (>2,000)	0.12	0.11	0.15	0.32	0.32	0.30
Tenure						
Tenure (<1 year)	0.06	0.00	0.17	0.03	0.00	0.08
Tenure (1–3 years)	0.31	0.06	0.20	0.16	0.04	0.09
Tenure (3–5 years)	0.18	0.20	0.11	0.14	0.11	0.08
Tenure (5–10 years)	0.19	0.37	0.25	0.21	0.31	0.25
Tenure (10–15 years)	0.07	0.13	0.11	0.12	0.12	0.14
Tenure (15–20 years)	0.06	0.06	0.06	0.11	0.13	0.12
Tenure (>20 years)	0.12	0.18	0.11	0.23	0.29	0.24
Industry						
Agriculture, hunting and forestry	0.02	0.02	0.02	0.01	0.00	0.01
Mining and quarrying	0.00	0.00	0.00	0.01	0.00	0.00
Manufacturing	0.29	0.32	0.28	0.29	0.31	0.30
Electricity, gas and water supply	0.01	0.00	0.01	0.02	0.02	0.02
Construction	0.10	0.08	0.06	0.08	0.06	0.06
Wholesale and retail trade	0.24	0.21	0.23	0.11	0.09	0.10
Transport and communication	0.04	0.03	0.03	0.07	0.06	0.05
Financial intermediation	0.03	0.04	0.05	0.04	0.05	0.05

Table 2-1: Sample means by year and treatment status (continued)

	Treated			Control		
	1994/95	1997/98	1999/00	1994/95	1997/98	1999/00
Industry						
Real estate and business activities	0.14	0.18	0.19	0.03	0.02	0.04
Public administration and defense	0.02	0.01	0.02	0.14	0.15	0.15
Education	0.01	0.01	0.02	0.07	0.08	0.06
Health and social work	0.06	0.05	0.07	0.11	0.11	0.12
Other social and personal service	0.04	0.04	0.03	0.03	0.04	0.04
Satisfaction with health						
Very poor	0.01	0.01	0.00	0.01	0.01	0.01
Poor	0.06	0.09	0.09	0.07	0.06	0.07
Satisfactory	0.27	0.31	0.26	0.28	0.30	0.30
Good	0.45	0.45	0.48	0.44	0.47	0.45
Very good	0.21	0.14	0.17	0.20	0.16	0.17
Health at present						
Very poor	0.01	0.01	0.01	0.01	0.01	0.01
Poor	0.10	0.13	0.12	0.11	0.10	0.11
Satisfactory	0.32	0.34	0.32	0.31	0.34	0.36
Good	0.49	0.44	0.45	0.47	0.47	0.45
Very good	0.08	0.08	0.10	0.10	0.08	0.07
n	2,227	1,056	1,620	8,024	5,044	5,731

Source: German Socio-Economic Panel (GSOEP), own calculations.

Table 2-2: Percentiles of absence days by period and treatment status

	1994 / 1995 (Pre-reform)		1997 / 1998 (Treatment Period)		1999 / 2000 (Repeal)	
Percentile	Treated (no coll. agreement)	Control (coll. agreement)	Treated (no coll. agreement)	Control (coll. agreement)	Treated (no coll. agreement)	Control (coll. agreement)
30	0	0	0	0	0	0
40	0	1	0	0	0	1
50	2	5	2	4	2	4
60	5	8	4	7	5	6
70	8	12	6	10	8	10
80	14	16	10	15	12	15
90	23	30	20	30	21	28
95	40	49	30	50	36	44
96	42	60	30	60	42	53
97	51	65	40	75	52	64
98	65	90	50	110	80	90
99	105	125	98	165	117	124
100	210	365	365	365	365	365
Mean	8.8	12.2	8.0	12.7	9.7	11.7
n	2,227	8,024	1,056	5,044	1,620	5,731

Source: German Socio-Economic Panel (GSOEP), own calculations.

Table 2-3: Difference-in-differences estimates

	OLS	NEGBIN	FE
Specification (1)	-1.82	-1.97	-1.19
	(1.44)	(1.50)	(1.27)
Specification (2)	-1.92	-2.07	-1.18
	(1.42)	(1.28)	(1.25)
Specification (3)	-1.74	-1.94*	-1.28
	(1.37)	(1.11)	(1.24)
Specification (4)	-1.99	-2.07**	-1.24
	(1.33)	(0.91)	(1.22)
n	23,702	23,702	23,702

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. The specifications are distinguished by the set of control variables: specification (1) includes no controls; specification (2) adds state unemployment, log hourly wage, civil status indicators, gender and some interaction terms to account for compositional changes: specification (3) adds education, citizenship, job and firm characteristics, and a dummy for West Germany; specification (4) extends the set of control variables by adding reported health status and satisfaction with health.
Source: German Socio-Economic Panel (GSOEP), own calculations.

Table 2-4: Difference-in-differences estimates for restricted age groups

	OLS	NEGBIN	FE
Age 20-64	-1.99	-2.07**	-1.24
n=23,702	(1.33)	(0.91)	(1.22)
Age 20-55	-2.85**	-2.30***	-2.35**
n=21,451	(1.24)	(0.83)	(1.10)
Age 20-40	-2.56**	-2.04***	-2.24**
n=12,097	(1.14)	(0.75)	(0.98)

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively.
Source: German Socio-Economic Panel (GSOEP), own calculations.

Table 2-5: Difference-in-differences quantile regression estimates

Percentile	**Age 20-64**	Pre-treatment percentile of treatment group	**Age 20-55**	Pre-treatment percentile of treatment group	**Age 20-40**	Pre-treatment percentile of treatment group
40	0.02 (0.21)	0	-0.07 (0.14)	0	-0.04 (0.15)	0
50	-0.20 (0.32)	2	-0.39 (0.29)	2	0.01 (0.59)	3
60	-0.28 (0.45)	5	-0.28 (0.53)	5	-0.71 (0.72)	5
70	-0.79 (0.56)	8	-0.77* (0.45)	8	-0.76** (0.36)	8
80	-1.33 (0.99)	14	-1.32** (0.58)	14	-1.18*** (0.36)	14
90	-4.83** (2.05)	23	-4.85*** (1.69)	21	-3.94** (1.53)	20
95	-9.30*** (2.86)	40	-7.66*** (2.06)	35	-4.01 (2.52)	30
98	-10.49* (6.04)	65	-15.84** (6.30)	60	-7.31 (6.49)	50
OLS	-1.99 (1.33)	-	-2.85** (1.24)	-	-2.56** (1.14)	-
n	23,702		21,451		12,097	

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. Percentiles of 30 and lower are zero because more than 30% of the sample did not report a single day of absence. 'pre-treatment percentile of treatment group' refers to the corresponding percentile of the treatment group in the pre-treatment period 1994 and 1995.
Source: German Socio-Economic Panel (GSOEP), own calculations.

Table 2-6: Switch-on versus switch-off difference-in-differences estimates

		Age 20-64		Age 20-55		Age 20-40	
		Switch-on	**Switch-off**	**Switch-on**	**Switch-off**	**Switch-on**	**Switch-off**
OLS		-1.39 (1.43)	-3.04** (1.53)	-1.97 (1.29)	-4.46*** (1.52)	-1.35 (1.11)	-5.49*** (1.86)
Difference (off-on) (p-value)		1.65 (0.23)		2.49* (0.07)		4.14** (0.02)	
	n	16,351	13,451	14,865	11,959	8,552	6,338
NEGBIN		-2.03** (1.03)	-2.05** (0.91)	-2.19** (0.93)	-2.63*** (0.80)	-1.55* (0.84)	-3.02*** (0.84)
Difference (off-on) (p-value)		0.02 (0.95)		0.44 (0.54)		1.47 (0.10)	
	n	16,351	13,451	14,865	11,959	8,552	6,338
FE		-1.39 (1.45)	-1.05 (1.39)	-2.29* (1.34)	-2.43* (1.33)	-1.66 (1.29)	-3.03* (1.55)
Difference (off-on) (p-value)		-0.34 (0.82)		0.14 (0.93)		1.37 (0.50)	
	n	23,702		21,451		12,097	

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. The fixed-effects switch-on and switch-off estimates are obtained from a single fixed-effects regression equation.
Source: German Socio-Economic Panel (GSOEP), own calculations.

Table 2-7: Switch-on versus switch-off quantile regression difference-in-differences estimates

	Aged 20–64		Aged 20–55		Aged 20–40	
Percentile	**Switch-on**	**Switch-off**	**Switch-on**	**Switch-off**	**Switch-on**	**Switch-off**
40	0.31	-0.05	-0.00	-0.28	0.27	-0.67***
	(0.17)	(0.22)	(0.21)	(0.22)	(0.23)	(0.01)
50	0.10	-0.65**	-0.20	-0.88***	0.41**	-1.49***
	(0.28)	(0.29)	(0.55)	(0.28)	(0.20)	(0.37)
60	0.25	-0.93**	0.08	-1.10**	-0.32	-2.42***
	(0.29)	(0.39)	(0.56)	(0.51)	(0.61)	(0.21)
70	-0.49	-1.57***	-0.51	-1.68***	-0.18	-2.30***
	(0.45)	(0.38)	(0.45)	(0.46)	(0.37)	(0.49)
80	-0.94	-2.81***	-1.19	-2.19***	0.10	-3.44***
	(0.84)	(0.72)	(0.85)	(0.38)	(0.63)	(0.31)
90	-3.86**	-4.33**	-3.94**	-5.31***	-2.05*	-7.71***
	(1.95)	(1.75)	(1.90)	(1.29)	(1.08)	(1.94)
95	-7.74***	-10.16***	-7.28**	-10.75***	-4.37	-8.72***
	(2.65)	(2.56)	(3.05)	(1.84)	(3.34)	(2.85)
98	-10.68	-8.20	-9.62*	-13.05*	-2.80	-12.63***
	(8.66)	(6.34)	(10.91)	(6.88)	(10.44)	(4.77)
OLS	-1.39	-3.04**	-1.97	-4.46***	-1.35	-5.49***
	(1.43)	(1.53)	(1.29)	(1.52)	(1.11)	(1.86)

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. Percentiles of 30 and lower are zero because more than 30% of the sample did not report a single day of absence.
Source: German Socio-Economic Panel (GSOEP), own calculations.

Table 2-8: 'Treatment effects' by calendar year (base year 1994) – including placebo estimates

	OLS	NEGBIN	FE
Age 20-64			
1995 * no coll. agr.	-1.20	1.84	-0.91
	(1.50)	(1.60)	(1.33)
1997 * no coll. agr.	-1.94	-1.69	-2.17
	(1.81)	(1.13)	(1.53)
1998 * no coll. agr.	-2.45	-0.63	-1.52
	(2.29)	(1.80)	(2.28)
1999 * no coll. agr.	0.90	0.77	-0.32
	(1.68)	(1.39)	(1.54)
2000 * no coll. agr.	0.46	1.66	-1.42
	(2.63)	(1.71)	(2.34)
Age 20-55			
1995 * no coll. agr.	-0.31	2.48	-1.12
	(1.30)	(1.58)	(1.35)
1997 * no coll. agr.	-1.95	-1.27	-2.94*
	(1.60)	(1.06)	(1.51)
1998 * no coll. agr.	-2.76	-0.91	-2.90**
	(1.79)	(1.58)	(1.80)
1999 * no coll. agr.	1.54	1.28	-0.76
	(1.48)	(1.36)	(1.51)
2000 * no coll. agr.	2.85	3.45*	-0.21
	(2.71)	(1.89)	(2.51)
Age 20-40			
1995 * no coll. agr.	-0.68	1.49	-2.11
	(1.41)	(1.13)	(1.69)
1997 * no coll. agr.	-2.01	-1.12	-3.15*
	(1.53)	(1.09)	(1.61)
1998 * no coll. agr.	-1.58	-0.7	-2.60
	(1.58)	(1.33)	(1.72)
1999 * no coll. agr.	1.20	1.62	-1.76
	(1.57)	(1.62)	(1.88)
2000 * no coll. agr.	6.77	5.18**	2.37
	(4.57)	(2.60)	(4.05)

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively.
Source: German Socio-Economic Panel (GSOEP), own calculations.

Table 2-9: Percentiles of other health-related outcomes

	Doctor visits (last 3 months)	Days in hospital	Number of hospital stays
30	0	0	0
40	1	0	0
50	1	0	0
60	2	0	0
70	2	0	0
80	3	0	0
90	6	0	0
95	10	7	1
96	10	10	1
97	10	12	1
98	12	15	1
99	17	24	2
100	90	220	20
Mean	2.41	1.16	0.12
n	23,701	23,680	23,612

Source: German Socio-Economic Panel (GSOEP), own calculations.

Table 2-10: Effects on other health-related outcomes

	OLS	NEGBIN	FE	Pre-reform mean
Age 20–64 (n=23,702)				
Doctor visits (last 3 months)	-0.26	-0.21	0.03	2.2
	(0.18)	(0.15)	(0.15)	
Days in hospital	-0.65**	-0.41***	-0.62*	1.35
	(0.30)	(0.11)	(0.32)	
Number of hospital stays	-0.061***	-0.045***	-0.065**	0.111
	(0.023)	(0.012)	(0.028)	
Age 20–55 (n=21,451)				
Doctor visits (last 3 months)	-0.25	-0.23	-0.01	2.21
	(0.19)	(0.16)	(0.16)	
Days in hospital	-0.67**	-0.37***	-0.73**	1.12
	(0.28)	(0.10)	(0.34)	
Number of hospital stays	-0.068***	-0.046***	-0.065**	0.108
	(0.024)	(0.012)	(0.031)	
Age 20–40 (n=12,097)				
Doctor visits (last 3 months)	-0.42*	-0.34**	-0.26	1.91
	(0.22)	(0.17)	(0.19)	
Days in hospital	-0.53**	-0.25*	-0.68**	0.65
	(0.23)	(0.15)	(0.29)	
Number of hospital stays	-0.079***	-0.050***	-0.085*	0.084
	(0.031)	(0.012)	(0.047)	

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. The ‘pre-reform mean’ refers to the treatment group in the pre-treatment period 1994 and 1995.
Source: German Socio-Economic Panel (GSOEP), own calculations.

Table 2-11: Effects on subjective health indicators (full sample)

	OLS	OLS – switch on	OLS – switch off	FE	FE – switch on	FE – switch off	Pre-reform mean
Health at present							
Age 20-64	-0.008 (0.013)	-0.000 (0.014)	-0.022 (0.014)	-0.007 (0.011)	0.003 (0.012)	-0.021* (0.012)	0.63
n	23,702	16,351	13,451	23,702	23,702		
Age 20-55	-0.011 (0.014)	-0.005 (0.014)	-0.024 (0.015)	-0.006 (0.011)	0.005 (0.013)	-0.020 (0.013)	0.63
n	21,451	14,865	11,959	21,451	21,451		
Age 20-40	0.003 (0.020)	0.018 (0.021)	-0.034 (0.022)	-0.002 (0.018)	0.020 (0.020)	-0.031* (0.018)	0.66
n	12,097	8,552	6,338	12,097	12,097		
Satisfaction with health							
Age 20-64	-0.021* (0.011)	-0.019 (0.012)	-0.025** (0.012)	-0.011 (0.008)	-0.008 (0.009)	-0.015 (0.010)	0.69
n	23,702	16,351	13,451	23,702	23,702		
Age 20-55	-0.019 (0.013)	-0.018 (0.014)	-0.023* (0.014)	-0.005 (0.008)	-0.002 (0.009)	-0.010 (0.010)	0.70
n	21,451	14,865	11,959	21,451	21,451		
Age 20-40	-0.009 (0.018)	-0.005 (0.020)	-0.023 (0.018)	-0.009 (0.011)	-0.006 (0.014)	-0.012 (0.013)	0.71
n	12,097	8,552	6,338	12,097	12,097		

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. Both indicators range between 0 and 1 with 1 indicating very good health. *Health at Present* is coded in 5, *Satisfaction with Health* in 11 different values. The fixed-effects switch-on and switch-off estimates are obtained from a single fixed-effects regression equation. The 'pre-reform mean' refers to the treatment group in the pre-treatment period 1994 and 1995.
Source: German Socio-Economic Panel (GSOEP), own calculations.

Table 2-12: Effects on subjective health indicators in the following year (full sample)

	OLS	OLS – switch on	OLS – switch off	FE	FE – switch on	FE – switch off	Pre-reform mean
Health at present (next year)							
Age 20-64	0.008 (0.012)	0.004 (0.013)	0.017 (0.014)	0.010 (0.008)	0.008 (0.009)	0.013 (0.010)	0.63
n	23,674	16,328	13,438	23,674	23,674		
Age 20-55	0.005 (0.012)	-0.002 (0.014)	0.018 (0.015)	0.011 (0.009)	0.010 (0.009)	0.014 (0.011)	0.64
n	21,426	14,844	11,948	21,426	21,426		
Age 20-40	0.013 (0.017)	0.011 (0.018)	0.019 (0.023)	0.015 (0.012)	0.019 (0.014)	0.010 (0.017)	0.67
n	12,085	8,543	6,334	12,085	12,085		
Satisfaction with health (next year)							
Age 20-64	-0.005 (0.011)	-0.010 (0.012)	0.004 (0.013)	0.001 (0.008)	0.001 (0.009)	0.000 (0.009)	0.69
n	23,674	16,328	13,438	23,674	23,674		
Age 20-55	-0.007 (0.012)	-0.014 (0.013)	0.005 (0.014)	0.000 (0.008)	-0.001 (0.010)	0.002 (0.009)	0.69
n	21,426	14,844	11,948	21,426	21,426		
Age 20-40	-0.010 (0.017)	-0.014 (0.018)	-0.004 (0.020)	-0.012 (0.012)	-0.012 (0.015)	-0.012 (0.014)	0.72
n	12,085	8,543	6,334	12,085	12,085		

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. Both indicators range between 0 and 1 with 1 indicating very good health. *Health at Present (next year)* is coded in 5, *Satisfaction with Health (next year)* in 11 different values. The fixed-effects switch-on and switch-off estimates are obtained from a single fixed-effects regression equation. The 'pre-reform mean' refers to the treatment group in the pre-treatment period 1994 and 1995.
Source: German Socio-Economic Panel (GSOEP), own calculations.

Table 2-13: Effects on subjective health indicators for the sample of people with positive number of doctor visits

	OLS	OLS – switch on	OLS – switch off	FE	FE – switch on	FE – switch off	Pre-reform mean
Health at present							
Age 20-64	-0.012	-0.009	-0.020	-0.005	0.005	-0.017	0.61
	(0.018)	(0.019)	(0.020)	(0.017)	(0.019)	(0.018)	
n	15,664	10,790	8,906	15,664	15,664		
Age 20-55	-0.015	-0.012	-0.024	-0.010	-0.001	-0.021	0.61
	(0.020)	(0.021)	(0.022)	(0.018)	(0.020)	(0.020)	
n	13,903	9,631	7,724	13,903	13,903		
Age 20-40	-0.015	0.002	-0.053	-0.030	-0.010	-0.059*	0.64
	(0.032)	(0.032)	(0.032)	(0.033)	(0.036)	(0.032)	
n	7,637	5,411	3,977	7,637	7,637		
Satisfaction with health							
Age 20-64	-0.028*	-0.029*	-0.028	-0.009	-0.007	-0.011	0.67
	(0.016)	(0.017)	(0.017)	(0.011)	(0.014)	(0.013)	
n	15,664	10,790	8,906	15,664	15,664		
Age 20-55	-0.025	-0.027	-0.023	0.000	0.003	-0.003	0.67
	(0.018)	(0.019)	(0.019)	(0.012)	(0.015)	(0.014)	
n	13,903	9,631	7,724	13,903	13,903		
Age 20-40	-0.014	-0.011	-0.019	-0.018	-0.015	-0.023	0.69
	(0.027)	(0.030)	(0.025)	(0.018)	(0.024)	(0.016)	
n	7,637	5,411	3,977	7,637	7,637		

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. Both indicators range between 0 and 1 with 1 indicating very good health. *Health at Present* is coded in 5, *Satisfaction with Health* in 11 different values. The fixed-effects switch-on and switch-off estimates are obtained from a single fixed-effects regression equation. The ‘pre-reform mean’ refers to the treatment group in the pre-treatment period 1994 and 1995.
Source: German Socio-Economic Panel (GSOEP), own calculations.

Table 2-14: Effects on subjective health indicators in the following year for the sample of people with positive number of doctor visits in the current year

	OLS	OLS – switch on	OLS – switch off	FE	FE – switch on	FE – switch off	Pre-reform mean
Health at present (next year)							
Age 20-64	0.006	-0.000	0.017	0.013	0.014	0.013	0.60
	(0.014)	(0.015)	(0.018)	(0.012)	(0.013)	(0.014)	
n	15,652	10,780	8,898	15,652	15,652		
Age 20-55	0.006	-0.003	0.022	0.014	0.013	0.015	0.60
	(0.016)	(0.017)	(0.019)	(0.013)	(0.014)	(0.016)	
n	13,893	9,623	7,717	13,893	13,893		
Age 20-40	0.016	0.014	0.025	0.033*	0.045**	0.015	0.64
	(0.021)	(0.021)	(0.030)	(0.018)	(0.018)	(0.024)	
n	7,633	5,409	3,974	7,633	7,633		
Satisfaction with health (next year)							
Age 20-64	-0.012	-0.017	-0.004	0.003	0.012	-0.009	0.66
	(0.014)	(0.015)	(0.017)	(0.010)	(0.012)	(0.012)	
n	15,652	10,780	8,898	15,652	15,652		
Age 20-55	-0.014	-0.021	-0.003	0.005	0.013	-0.006	0.66
	(0.015)	(0.017)	(0.019)	(0.011)	(0.013)	(0.013)	
n	13,893	9,623	7,717	13,893	13,893		
Age 20-40	-0.009	-0.009	-0.007	0.012	0.029	-0.013	0.68
	(0.022)	(0.024)	(0.028)	(0.016)	(0.018)	(0.019)	
n	7,633	5,409	3,974	7,633	7,633		

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. Both indicators range between 0 and 1 with 1 indicating very good health. *Health at Present (next year)* is coded in 5, *Satisfaction with Health (next year)* in 11 different values. The fixed-effects switch-on and switch-off estimates are obtained from a single fixed-effects regression equation. The 'pre-reform mean' refers to the treatment group in the pre-treatment period 1994 and 1995.
Source: German Socio-Economic Panel (GSOEP), own calculations.

Table 2-15: Effects on subjective health indicators for the sample of people with positive number of days in hospital

	OLS	OLS – switch on	OLS – switch off	FE	FE – switch on	FE – switch off	Pre-reform mean
Health at present							
Age 20-64	-0.036 (0.043)	-0.044 (0.047)	-0.025 (0.045)	0.045 (0.061)	0.113* (0.067)	0.016 (0.064)	0.56
n	2,183	1,533	1,242	2,183	2,183		
Age 20-55	-0.040 (0.055)	-0.066 (0.057)	-0.001 (0.056)	0.062 (0.083)	0.070 (0.086)	0.057 (0.093)	0.58
n	1,890	1,334	1,041	1,890	1,890		
Age 20-40	0.034 (0.058)	0.009 (0.056)	0.073 (0.072)	-	-	-	0.62
n	1,034	726	555				
Satisfaction with health							
Age 20-64	-0.055 (0.049)	-0.071 (0.051)	-0.026 (0.049)	0.041 (0.045)	0.042 (0.057)	0.040 (0.049)	0.63
n	2,183	1,533	1,242	2,183	2,183		
Age 20-55	-0.047 (0.064)	-0.073 (0.064)	-0.000 (0.062)	0.031 (0.058)	0.024 (0.070)	0.036 (0.031)	0.63
n	1,890	1,334	1,041	1,890	1,890		
Age 20-40	0.106** (0.042)	0.066 (0.050)	0.163*** (0.053)	-	-	-	0.66
n	1,034	726	555				

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. Both indicators range between 0 and 1 with 1 indicating very good health. *Health at Present* is coded in 5, *Satisfaction with Health* in 11 different values. The OLS results for the age group 20-40 are only based on 29 persons who are in the treatment group in the treatment period. Hence, statistical inference on these coefficients may be invalid, so that we do not take the statistical significance of the positive coefficients for 'Satisfaction with Health' seriously. We do not report fixed-effects estimates for this age group because we only have 10 persons with a within variation in the treatment indicator. For the age groups 20-64 and 20-55 the number of persons with a within-variation in the treatment indicator is 26 (13 switch on and 17 switch off) and 21 (11 switch on and 13 switch off), respectively. Hence, these estimates have to be taken with a grain of salt. The fixed-effects switch-on and switch-off estimates are obtained from a single fixed-effects regression equation. The 'pre-reform mean' refers to the treatment group in the pre-treatment period 1994 and 1995.
Source: German Socio-Economic Panel (GSOEP), own calculations.

Table 2-16: Effects on subjective health indicators in the following year for the sample of people with positive number of days in hospital in the current year

	OLS	OLS – switch on	OLS – switch off	FE	FE – switch on	FE – switch off	Pre-reform mean
Health at present (next year)							
Age 20-64	0.002	0.010	-0.002	0.006	-0.066	0.038	0.52
	(0.045)	(0.048)	(0.052)	(0.056)	(0.068)	(0.056)	
n	2,181	1,532	1,241	2,181	2,181		
Age 20-55	-0.008	-0.015	0.020	-0.021	-0.079	0.015	0.54
	(0.053)	(0.054)	(0.059)	(0.054)	(0.072)	(0.054)	
n	1,888	1,333	1,040	1,890	1,890		
Age 20-40	0.053	0.048	0.109	-	-	-	0.60
	(0.064)	(0.060)	(0.071)				
n	1,033	725	555				
Satisfaction with health (next year)							
Age 20-64	-0.005	-0.002	0.004	-0.059	-0.062	-0.057	0.59
	(0.041)	(0.044)	(0.050)	(0.053)	(0.065)	(0.054)	
n	2,181	1,532	1,241	2,181	2,181		
Age 20-55	-0.012	-0.022	0.022	-0.068	-0.079	-0.061	0.60
	(0.047)	(0.050)	(0.057)	(0.056)	(0.063)	(0.061)	
n	1,888	1,333	1,040	1,890	1,890		
Age 20-40	0.001	-0.010	0.075	-	-	-	0.65
	(0.057)	(0.061)	(0.071)				
n	1,033	725	555				

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. Both indicators range between 0 and 1 with 1 indicating very good health. *Health at Present (next year)* is coded in 5, *Satisfaction with Health (next year)* in 11 different values. The OLS results for the age group 20-40 are only based on 29 persons who are in the treatment group in the treatment period. Hence, statistical inference on these coefficients may be invalid. We do not report fixed-effects estimates for this age group because we only have 10 persons with a within variation in the treatment indicator. For the age groups 20-64 and 20-55 the number of persons with a within-variation in the treatment indicator is 26 (13 switch on and 17 switch off) and 21 (11 switch on and 13 switch off), respectively. Hence, these estimates have to be taken with a grain of salt. The fixed-effects switch-on and switch-off estimates are obtained from a single fixed-effects regression equation. The 'pre-reform mean' refers to the treatment group in the pre-treatment period 1994 and 1995.
Source: German Socio-Economic Panel (GSOEP), own calculations.

Table 2-17: Effects on continuous sickness for at least 6 weeks

	OLS	OLS – switch on	OLS – switch off	FE	FE – switch on	FE – switch off	Pre-reform mean
6 continuous weeks ill							
Age 20-64	-0.024**	-0.018	-0.027**	-0.022*	-0.029**	-0.014	0.032
	(0.012)	(0.013)	(0.013)	(0.011)	(0.013)	(0.013)	
n	23,702	13,451	16,351	23,702	23,702		
Age 20-55	-0.032***	-0.035***	-0.031***	-0.034***	-0.036***	-0.031**	0.028
	(0.010)	(0.012)	(0.011)	(0.011)	(0.013)	(0.012)	
n	21,451	11,959	14,856	21,451	21,451		
Age 20-40	-0.021**	-0.029*	-0.019*	-0.015	-0.012	-0.018	0.020
	(0.010)	(0.015)	(0.010)	(0.009)	(0.012)	(0.013)	
n	12,097	6,338	8,552	12,097	12,097		
6 continuous weeks ill (next year)							
Age 20-64	-0.007	-0.009	-0.005	-0.002	-0.002	-0.002	0.034
	(0.011)	(0.015)	(0.011)	(0.010)	(0.012)	(0.013)	
n	23,702	13,451	16,351	23,702	23,702		
Age 20-55	-0.012	-0.019	-0.008	-0.008	-0.006	-0.010	0.031
	(0.011)	(0.016)	(0.011)	(0.011)	(0.012)	(0.013)	
n	21,451	11,959	14,856	21,451	21,451		
Age 20-40	-0.016	-0.042*	-0.003	-0.002	0.013	-0.022	0.016
	(0.010)	(0.023)	(0.008)	(0.009)	(0.012)	(0.016)	
n	12,097	6,338	8,552	12,097	12,097		

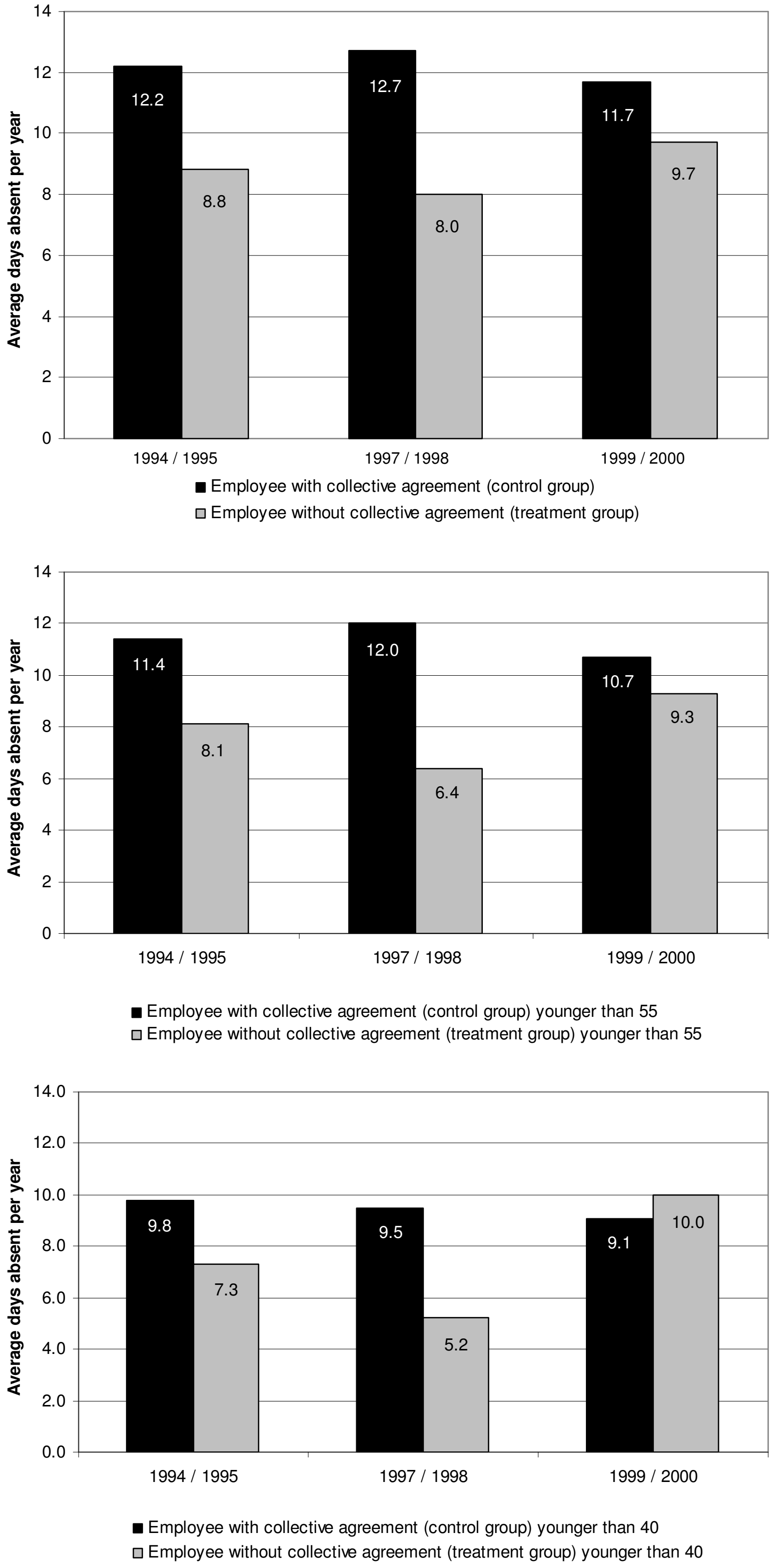
Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. The fixed-effects switch-on and switch-off estimates are obtained from a single fixed-effects regression equation. The 'pre-reform mean' refers to the treatment group in the pre-treatment period 1994 and 1995.
Source: German Socio-Economic Panel (GSOEP), own calculations.

Table 2-18: Effects on satisfaction with financial security in case of sickness

		OLS switch-off	FE switch-off	Pre-reform mean
All workers				
	20-64	0.014	-0.013	0.66
		(0.020)	(0.019)	
	n	5,559	5,559	
	20-55	0.027	-0.000	0.65
		(0.020)	(0.018)	
	n	4,930	4,930	
	20-40	0.007	0.000	0.63
		(0.031)	(0.030)	
	n	2,388	2,388	
Doctor visits > 0				
	20-64	0.014	-0.033	0.64
		(0.026)	(0.030)	
	n	3,579	3,579	
	20-55	0.032	-0.010	0.64
		(0.027)	(0.030)	
	n	3,103	3,103	
	20-40	0.001	-0.000	0.62
		(0.045)	(0.050)	
	n	1,407	1,407	
Days in hospital > 0				
	20-64	0.105**	-	0.63
		(0.052)		
	n	662		
	20-55	0.097*	-	0.63
		(0.055)		
	n	560		
	20-40	-0.150**	-	0.64
		(0.067)		
	n	262		

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. The indicator ranges between 0 and 1 with 1 indicating a high satisfaction. The original variable is coded in 11 different values. The OLS result for the age groups 20-64, 20-55 and 20-40 for those who have been in hospital is only based on 59, 49 and 22 persons who are in the treatment group in the treatment period, respectively. Hence, statistical inference, especially on the negative point estimate for the age group 20-40 may be invalid, so that we do not take the statistical significance of this estimate seriously. We do not report fixed-effects estimates for the restriction on persons who were in hospital because we only have 3, 3, and 1 person with a within variation in the treatment indicator for the age groups 20-64, 20-55, and 20-40, respectively. The 'pre-reform mean' refers to the treatment group in the pre-treatment period 1994 and 1995.
Source: German Socio-Economic Panel (GSOEP), own calculations.

Figure 2-1: Number of days absent over time by treatment status and age group



Note: We only observe the total number of days absent by calendar year, not the length of single spells of absence. The sample includes only firm stayers.
Source: German Socio-Economic Panel (GSOEP), own calculations.

Appendix to Chapter 2

Appendix 2-1: Sample selection

Year	Sample size (including all years)	Individual is in the sample this year	Individual is also in the sample the following year	Including only employed persons between 20 and 64 years of age	No missings for questions on absence	No missings for questions on income and other explanatory variables
1994	56,150	13,417	12,520	6,288	6,040	5,134
1995	56,150	13,768	12,851	6,526	6,278	5,576
1997	56,150	13,283	12,180	5,931	5,658	4,964
1998	56,150	14,670	13,373	6,394	6,160	5,428
1999	56,150	14,085	13,035	6,443	6,196	5,263
2000	56,150	24,586	21,233	10,083	9,690	8,527
n	336,900	93,809	85,192	41,665	40,022	34,892

Source: German Socio-Economic Panel (GSOEP), own calculations.

Appendix 2-2Selection of treatment and control groups

Year	Treated	Control	Movers	Rest	n
1994	1,021	3,702	0	411	5,134
1995	1,206	4,322	0	48	5,576
1997	585	2,691	962	726	4,964
1998	471	2,353	1,104	1,500	5,428
1999	845	2,956	0	1,462	5,263
2000	775	2,775	0	4,977	8,527
n	4,903	18,799	2,066	9,124	34,892

Note: To be part of either the treatment or control group in this study, an individual must have answered the question on collective bargaining in 1995. Hence, the number of observations is highest for both treated and control individuals in 1995. Panel attrition then works both backward and forward in time. So that observations can be classified into treatment and control, a worker must not have changed employer between 1996 and 1998 (i.e., until the end of the treatment period). Workers that have changed (termed ‘movers’) are deleted from the sample. If, however, an individual answered the question on collective bargaining coverage in 1995 but changed employer before 1995 or in 1999/2000, we retain that employee in the sample. The last column, labeled ‘rest’, includes workers who did not answer the question on collective bargaining in 1995, meaning that they cannot be classified as either treated or control and are therefore deleted from the sample. The allocation to the treatment or control group here is based on the 1995 information on collective bargaining coverage. It should also be noted that misclassification outside the treatment period is harmless because neither the treatment nor the control group was treated either before or after the repeal of the reform. Thus, keeping all persons who answered the 1995 question on collective bargaining coverage may improve precision in the repeated cross-section difference-in-differences estimates. In the fixed-effects estimates, the coefficient on treatment is driven only by observations present at least once in the treatment period and at least once in a non-treatment period.

Appendix 2-3: Full estimation results

	OLS	NEGBIN	FE
No collective agreement	-1.02	-0.93	
	(0.79)	(0.65)	
Year 1995	0.58	0.79*	0.20
	(0.58)	(0.47)	(0.54)
Year 1997	0.04	0.35	0.35
	(1.18)	(0.68)	(1.08)
Year 1998	1.80	1.29*	1.42
	(1.27)	(0.71)	(1.09)
Year 1999	-0.02	0.18	-0.16
	(0.81)	(0.58)	(0.76)
Year 2000	1.05	0.47	
	(0.96)	(0.59)	
No coll. agreem. × Year of Reform	-1.99	-2.07**	-1.24
	(1.33)	(0.91)	(1.22)
Unemployment rate	0.04	-0.05	-1.33
	(0.18)	(0.10)	(1.54)
Hourly wage	-1.74*	-1.58*	-0.11
	(0.98)	(0.82)	(0.29)
Civil status indicators			
Age	-0.14	-0.27	-2.71***
	(0.39)	(0.22)	(0.86)
Age squared	0.00	0.00	0.04***
	(0.00)	(0.00)	(0.01)
Married	0.34	-0.06	-1.27
	(1.46)	(0.81)	(1.67)
Female	1.94	-3.14	
	(9.89)	(6.29)	
Children younger than 16	-2.12**	-1.08*	0.79
	(0.86)	(0.63)	(1.80)
Female × children younger than 16	-2.21	-1.68*	-1.40
	(1.83)	(0.99)	(2.65)
Female × Married	-0.09	0.31	1.88*
	(0.55)	(0.35)	(1.06)
Female × age	0.00	0.00	-0.02*
	(0.01)	(0.00)	(0.01)
Female × age squared	3.28**	1.60	0.11
	(1.32)	(1.14)	(2.18)
Education (ref. Apprenticeship)			
Higher education (university degree)	-1.88**	-2.41***	-1.47
	(0.92)	(0.65)	(1.22)
Higher education (no degree)	-0.14	0.07	3.34**
	(0.78)	(0.58)	(1.34)
No degree	0.95	0.27	4.92*
	(0.99)	(0.60)	(2.88)
Job and firm characteristics			
Temporary work contract	0.41	-0.10	-1.51
	(1.74)	(1.09)	(2.70)
Working fulltime	3.91***	3.24***	2.58*
	(0.81)	(0.51)	(1.49)

Appendix 2-3: Full estimation results (continued)

	OLS	NEGBIN	FE
Job and firm characteristics			
Blue-collar worker	4.18***	4.57***	3.26**
	(0.78)	(0.64)	(1.58)
Civil servant	3.15	3.22**	-0.63
	(2.10)	(1.33)	(2.88)
Citizenship			
German	-1.81	-1.17	1.86
	(1.12)	(0.80)	(3.29)
West-Germany	0.94	-0.20	-3.06
	(1.37)	(0.87)	(2.74)
Firm size (ref. 1-19)			
Firm size (20-199)	2.42***	2.04***	-0.06
	(0.77)	(0.67)	(1.37)
Firm size (200-1999)	2.17**	2.29***	-1.63
	(0.86)	(0.73)	(1.51)
Firm size (>2000)	3.45***	3.36***	-1.39
	(0.88)	(0.79)	(1.72)
Tenure (ref. < 1 year)			
Tenure (1-3 years)	-0.08	-0.06	3.30*
	(1.65)	(1.45)	(2.00)
Tenure (3-5 years)	1.21	1.13	4.66***
	(1.66)	(1.60)	(1.46)
Tenure (5-10 years)	0.03	0.60	4.22***
	(1.56)	(1.54)	(1.55)
Tenure (10-15 years)	0.01	-0.37	3.63**
	(1.65)	(1.46)	(1.65)
Tenure (15-20 years)	-0.26	0.01	3.93**
	(1.68)	(1.55)	(1.80)
Tenure (>20 years)	0.46	0.03	5.91***
	(1.72)	(1.53)	(2.23)
Industry (ref. manufacturing)			
Agriculture, hunting and forestry	-2.81*	-0.78	-2.21
	(1.47)	(1.47)	(2.10)
Mining and quarrying	5.64	7.80	17.82
	(7.77)	(8.80)	(17.60)
Electricity, gas and water supply	-1.66	-0.46	-1.64
	(1.22)	(1.12)	(2.02)
Construction	1.07	1.22	1.78
	(1.10)	(0.84)	(1.64)
Wholesale & retail trade	0.33	0.65	0.31
	(0.88)	(0.84)	(1.96)
Transport and communication	4.81**	3.56***	1.46
	(2.18)	(1.34)	(2.11)
Financial intermediation	-0.36	-0.60	-1.67
	(0.89)	(0.81)	(2.29)
Real estate and business activities	0.59	1.22	0.90
	(1.04)	(0.95)	(1.64)
Public administration and defence	0.59	1.29	-0.67
	(1.43)	(0.87)	(1.93)
Education	-0.26	0.59	1.24
	(1.49)	(1.07)	(2.88)

Appendix 2-3: Full estimation results (continued)

	OLS	NEGBIN	FE
Industry (ref. manufacturing)			
Health and social work	2.84**	2.50***	0.04
	(1.23)	(0.95)	(1.79)
Other social & personal service	2.28*	2.74*	1.96
	(1.28)	(1.56)	(1.77)
Health at present (ref. satisfactory)			
Very poor	38.52***	21.65***	30.70***
	(6.40)	(4.06)	(6.07)
Poor	11.54***	6.69***	8.16***
	(1.52)	(1.02)	(2.16)
Good	-2.47***	-2.66***	-1.06
	(0.58)	(0.43)	(0.75)
Very good	-5.02***	-4.66***	-1.73*
	(0.86)	(0.56)	(0.96)
Satisfaction with health (ref. satisfactory)			
Very poor	7.59	4.84**	1.76
	(6.44)	(2.35)	(6.75)
Poor	4.32**	2.24***	3.50*
	(1.79)	(0.85)	(2.10)
Good	-1.96***	-2.14***	-1.94***
	(0.56)	(0.41)	(0.72)
Very good	-1.22	-1.60***	-1.15
	(0.84)	(0.60)	(0.88)
n	23,702	23,702	23,702
R^2	0.11		0.04

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively.
Source: German Socio-Economic Panel (GSOEP), own calculations.

Chapter 3:

The Relationship between Overweight and Wages for Women in Germany: Empirical Evidence for Discrimination

Acknowledgment: This project was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft) as part of the research initiative 'Flexibility in Heterogeneous Labour Markets'. I am especially grateful to Patrick Puhani for guidance and advice. I also thank Falko Tabbert, Philip Yang, Olaf Hübler, seminar participants at the University of Hannover and participants at the research initiative's meeting in Berlin for helpful comments. All remaining errors are my own.

3.1 Introduction to Chapter 3

Obesity seems to be one of the most severe health problems in industrialized countries in the last decades. Among countries belonging to the European Union, the 'International Association for the Study of Obesity' ranks Germany at top position with respect to having the highest share of overweight and obese people. According to self-reported data from a large German individual-level dataset, the German Micro Census, the share of overweight and obese women is 40 percent while the share of overweight or obese men is 57 percent.[33] Overweight and obese persons face significant health limitations, like cardiovascular diseases or Type 2 diabetes mellitus.[34] In Germany, around 25 percent of the adults suffer from cardiovascular diseases (e.g. high blood pressure) which are one of the most common consequences of overweight.

Besides health limitations, overweight persons might also face discrimination in their private and in their business environment. Reduction of well-being, a negative body image, social exclusion, and decreased concentration ability are potential consequences of obesity which could lead to a reduced quality of life. This might also have an impact on productivity and therefore result in lower wages or less success in the labor market. For the U.S., several studies find a negative impact of obesity on labor market outcomes, especially on wages. This negative effect of a higher body weight on earnings is especially true for overweight white women, who earn significantly less than their healthy-weight counterparts (Cawley (2004), Averett and Korenman (1996)). This result is partly reconfirmed in European studies (for Denmark, Belgium, Ireland, Italy, Greece, Spain, Portugal, Austria and Finland) by Fahr (2006), Brunello and D`Hombres (2006) and Garcia and Quintana-Domeque (2006) based on the European Community Household Panel.

[33] As stated in Cawley and Danziger (2004), these numbers are likely to be underreported.
[34] See: http://www.iotf.org/cardiovascular.asp

This paper finds that these patterns are also true for Germany: overweight and obese women earn significantly less than women of healthy weight, while there seems to be no relationship between overweight and wages for men. In contrast to most papers, the focus of this study is to uncover through which channel the negative relationship between wages and obesity can be explained. I test two hypotheses that could be possible explanations for the gap in wages between overweight women and women of healthy weight. The first hypothesis states that lower wages are due to reduced productivity of overweight women (productivity hypothesis); while the second hypothesis claims that it is due to discrimination (discrimination hypothesis). In order to find support for either the productivity or the discrimination hypotheses, I set up four different subgroup designs and estimate the effect of weight on wages for each subgroup. At first, I sort women according to the gender-composition of their coworker into jobs that are 'male-dominated', 'female-dominated' or 'male-female-balanced'. Secondly, women are categorized by whether their job is interactive (with contact to customers) or non-interactive. Thirdly, correlations between overweight and wages are estimated for employed and self-employed women. Lastly, the dataset is divided into young and older women. Since the correlation between weight and wages differs significantly between subgroups, results clearly support the discrimination hypothesis.

The remainder of this paper is organized as follows. Section 3.2 gives a literature review on studies on obesity and on the effect of overweight on wages. Section 3.3 introduces the dataset and discusses the four hypotheses in detail. Results are presented in Section 3.4 and can be summarized as follows: obese women earn 2.4 percent less than women of healthy weight; while women belonging to the heaviest 10 percent earn 4.3 percent lower wages. The subgroup analyses reveal that women with mainly male coworkers face a higher penalty for being overweight than overweight women in jobs with female coworkers or in male-female-balanced jobs. In addition, overweight women in interactive jobs receive much higher wage reductions than overweight women in non-interactive jobs. Moreover, only overweight employed women face lower wages, while the obesity-effect on wages in non-existing in self-employment. Finally, young women get much higher wage reductions when overweight than older women. Section 3.5 summarizes and concludes the paper.

3.2 Literature review

In the last 20 years, the study of obesity and labor market outcomes has come into the interest of economists. Obesity seems to be one of the most severe health problems in industrialized countries and it also has economic consequences. In the last decades, several studies emerged, bringing together obesity and labor market outcomes such as wages or employment. Most studies focus on the U.S. since the prevalence of obesity is extremely high there. But also European countries have rising obesity rates, with Germany heading the countries belonging to the European Union with respect to the share of overweight and obese persons.[35] The literature on obesity can be divided into two parts: Studies that detect the prevalence and potential reasons for rising obesity rates and those focusing on the impact of obesity on labor market outcomes such as employment and wages.

For the U.S., Chou et al. (2004) show, that prevalence of obesity has been relatively constant between 1960 and 1980, while is has doubled between 1980 and 2000. But what are the reasons for the enormous increase of obesity in the last 30 years? Chou et al. (2004) find that lower fast food prices, higher per capita number of restaurants, risings cigarette prices and anti-smoking campaigns are the most important factors for rising obesity rates. Similar results are found by Lakdawalla and Philipson (2002). They show that reduced food prices and declining physical activity from agricultural innovations and technological changes account for a significant increase of obesity in the United States. Other studies on the determinants of obesity (Conley and Glauber (2005), Costa-Font and Gil (2004), Robert and

[35] Source: International Association for the Study of Obesity.

Reither (2004), Sobal and Stunkard (1989), Zhang and Wang (2004)) find that a higher body weight is associated with a lower socioeconomic status; although Zhang and Wang (2004) show that this trend has decreased over the last 30 years.

Empirical literature has a clear focus on the impact of overweight and obesity on wages (Register and Williams (1990), Gortmaker et al. (1993), Averett and Korenman (1996), Pagán and Dávila (1997), Behrman and Rosenzweig (2001), Cawley (2004), Baum and Ford (2004)). Most of these U.S. studies find a significant reduction in wages for overweight and obese white women. Results for men and for black and Hispanic women are not clear.

Studies for European countries also find negative consequences of a higher body weight: Johansson et al. (2007) provide evidence that waist circumference is negatively related to wages for women in Finland. Paraponaris (2005) shows that overweight and obese people spend significantly more time in unemployment in France. Like in U.S. studies, Sargent and Blanchflower (1994) find for Great Britain a strong and negative association of overweight and wages for women, but no effect for men. Other European studies find mixed results based on the European Household Panel (for Spain, Greece, Italy, Portugal, Austria, Ireland, Denmark, Belgium and Finland): Brunello and D'Hombres (2006) show that men get a higher wage deduction when overweight or obese (except for Finland and Portugal), while Fahr (2006) finds that lower wages are associated with higher body weight for women but not necessarily for men. The only German study by Cawley et al. (2005), based on GSOEP data, finds a negative correlation between overweight and wages for women, although these results do not hold in an IV estimation.

Besides finding an association between overweight and wages, many studies try to estimate a *causal* effect of overweight on wages. If there were no unobserved factors, which are correlated with weight, an OLS regression would estimate a causal effect of weight on wages. But since there could be unobserved factors (for example discipline or self-esteem) that are correlated with weight, OLS results become biased. In order to solve this endogeneity problem, recent studies use instrumental variable estimation. Pagán and Dávila (1997) use family poverty level, health limitations and self-esteem as instruments, but by using a Hausman specification test, they cannot reject the hypothesis of weight being not endogenous in a wage regression. As pointed out by Cawley (2004), this is probably due to a correlation of the instruments with the error term in the wage regression.

Most studies using IV estimation to address the endogeneity problem take weight of a family member as an instrument for a respondent's own weight. Cawley et al. (2005), Cawley (2000), and Brunello and D'Hombres (2006) use the weight of children or parents as instruments. IV estimates generally go into the same direction as OLS results (coefficients are mostly larger than for OLS estimates), but some become insignificant due to reduced sample sizes and much larger standard errors. Cawley (2004) takes the weight of a sibling as an instrument for a respondent's weight. His finding, that white women (in contrast to white men) receive lower wages when they have a higher BMI can be confirmed through this IV estimation, although a Hausman specification test cannot reject the hypothesis that OLS and IV results are equal. Therefore, he concludes that OLS should be preferred over IV since it has lower standard errors. Moreover, any potential endogeneity of weight does not seem to have an impact on OLS results.

Since a person shares the same genes with his or her children and siblings, the weight of children and siblings seems to be the most convincing instrument for a person's own weight in this context. But still, one cannot be sure whether the exogeneity condition is met. If a mother's weight is correlated with the child's weight through genetics, she might also pass on other characteristics that are correlated with labor market success (discipline, motivation). In this case, the weight of a child is not a valid instrument for the mother's weight. Analogously, if siblings share the same genes when it comes to weight, they might also have the same genes in other personal characteristics that affect labor market outcomes. Again, in this case, the weight of a sibling cannot be used as an instrument of a persons own weight.

In contrast to papers that concentrate on estimating the impact of weight on wages, this paper does not try to estimate a causal effect, since there is no natural experiment to evaluate in this context and no suitable instrument in this dataset. But even if there was a causal effect of weight on wages, the question of whether overweight persons are being discriminated or whether lower wages are due to reduced productivity of overweight women could not be answered. Therefore, this paper focuses on this question and results clearly favor the discrimination hypothesis over the productivity hypothesis.

3.3 Data and identification strategy

This study is carried out on basis of the German Micro Census. It is a large German dataset; consisting of a one-percent sample of the entire German population (the scientific community receives a 70 percent sample of that one percent). Only more recent waves contain information on weight and height. Therefore, trends in overweight and obesity over time cannot be shown.[36] For this analysis, I take a pooled sample of observations for the years 2003 and 2005. The dependent variable is the logarithm of the monthly net wage. Wages are reported in 24 categories, with the highest limit at a net of 18,000 Euros per month, which puts only 142 individuals (0.04 percent of all women and 0.13 percent of all men) into a right-censored category. To estimate the relationship between weight and wages, I use interval regressions to account for wages being reported in intervals. As in other studies on this topic, analyses will be carried out separately for men and women. But since I will not be able to find any effects for men, further analyses will be carried out for women only.

Persons older than 55 years of age are excluded, since many employers offer early retirement programs to their employees. In this case, an employee receives lower wages while

[36] For more information, please see: http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/EN/press/abisz/Mikrozensus__e,templateId=renderPrint.psml

working, retires earlier than in the age of 65 and still receives a lower salary until the retirement payments start at the usual retirement age of 65. In this case, information on wages would be biased. Compared to other datasets that have information on weight and height, the Micro Census dataset has the advantage that it has enough observations to conduct separate analysis for subgroups, which is described in detail in the next section. The sample of employed women between 20 and 55 years of age consists of more than 63,000 observations; the sample of men is a bit larger with more than 75,000 observations since more men than women join the labor force. In the full sample of all persons between 20 and 55 years of age, 20 percent of all men and 26 percent of all women do not participate in the labor market. The fact that more women tend not to participate in the labor market might lead to problems of sample selection. For example, if women of higher body weight are less likely to be employed, one had to determine whether these women do not want to work or whether they do not find a job (which might be due to discrimination). But in this sample, there is no significant difference in weight or BMI of women who participate in the labor market and those who do not, although age seems to matter. While the difference in BMI between employed and non-employed women is virtually zero (0.2 BMI units) for women between 20 and 35 years of age, the difference in weight becomes larger (1.3 BMI units) for women older than the age of 35, although both differences are not statistically significant different from zero.

Descriptive statistics of the average height and different measures of weight for employed persons (excluding self-employed and unemployed persons) are shown in Table 3.1. While the average woman is 1.67m tall and weighs 66.1kg, her BMI of 23.7 is in the recommended range. The average BMI of men is 25.9 which is classified as slightly overweight (corresponding to an average height and weight of 1.80m and 83.4kg, respectively). In this paper, standard definitions of BMI (defined as (weight in kg)/(height in m)^2) and the classification into four clinical categories ‘underweight’ (BMI lower than 18.5), 'healthy weight' (BMI between 18.5 and 25), 'overweight' (BMI between 25 and 30), and 'obese' (BMI higher than 30) are used. Due to the fact that height and weight are self-reported, height might be overstated while weight is likely to be understated.[37] To account for this problem, I will not only rely on OLS regressions with BMI or dummies for the clinical categories as explanatory variables, but also introduce the deciles of the distribution of BMI as explanatory dummy variables. By sorting persons into the deciles of the distribution of BMI, they are classified relatively to the other respondents. If all respondents follow the same pattern by reporting a lower weight and overstating their height, the reported measures of BMI do not matter, since only the position of a person relative to the other respondents is of interest. In the tables, results for all three regression types (correlation with BMI, with clinical categories of weight and with deciles of BMI) are presented.

The main interest of this paper is to answer the question through which channel lower wages for obese women can be explained. Therefore, I introduce two hypotheses. First hypothesis: Overweight persons are less productive due their corpulence. If they are paid according to their productivity, it would be fair that they earn less (**Productivity Hypothesis**). Second hypothesis: Employers perceive overweight and obese persons to be less productive because of prejudices against them. According to Roehling (1999), the most common prejudices against overweight persons are that they are lazy, less conscientious, less competent, emotionally unstable, or have less self-discipline or self-control. If these prejudices are not true, it would be discrimination if employers pay lower wages to overweight persons (**Discrimination Hypothesis**). Moreover, there might be unobserved factors (like self-esteem) that influence both, weight and wages. But since the true (causal) effect is hard to estimate in the absence of a natural experiment or a suitable instrument for body weight, this cannot be tested. Therefore, I concentrate on finding evidence to support

[37] Cawley and Danziger (2004) account fort his problem by using reference weight and height measures in their study. They show that female current and former welfare-recipients report 8 to 12 pounds lower body weight and about 0.7 inches taller height. Nevertheless, this procedure to account for misreporting cannot be used for German data, since we do not have reference measures of weight and height.

either the productivity hypothesis or the discrimination hypothesis, using four different subgroup designs.

First, I divide the dataset of employed women by gender-dominance in their workplace. Based on more than 340 occupations, those jobs with more than 70 percent males or females are labeled 'male-dominated jobs' or 'female-dominated jobs', respectively. The rest are called 'male-female balanced jobs'. If overweight and obese women were less productive than women of healthy weight (productivity hypothesis), one would expect similar effects of weight on wages irrespective of gender-dominance in the workplace, since they should be less productive in all working environments. Nevertheless, it might be the case that being slightly overweight could be an advantage in some male-dominated jobs such as manufacturing jobs or jobs that demand physical power, where it bit more body mass might be useful (e.g. farmer or construction worker). In this case, one would expect the effect of weight on wages to be smaller in male-dominated jobs (if the productivity hypothesis holds). Results would be in favor of the discrimination hypothesis, if effects were larger for male-dominated jobs, since not only most of the coworkers are male, but also the supervisor is likely to be male and empirical studies find that men have more prejudices against overweight women than women do (Harris et al. (1991)), and are therefore more likely to discriminate against overweight women.

Second, two categories of jobs are built: interactive and non-interactive jobs. Interactive jobs are jobs that require communication skills, the ability to work with others (including coworkers and customers), while non-interactive jobs do not require these skills. Based on the task-based approach by Spitz-Oener (2006) interactive tasks include negotiating, lobbying, coordinating, organizing, teaching or training, selling, buying, advising customers, advertising, entertaining or presenting, and employing or managing personnel. Since the Micro Census dataset does not have information on which tasks are included in each job, I take the Qualification and Careers Survey (Qualifikation und Berufsverlauf, IAB-BIBB) dataset to determine whether or not a job is an interactive job.[38] Results can be re-matched to the Micro Census since in both datasets jobs are categorized by a standard classification. Spitz-Oener (2006) defines five groups of tasks (non-routine analytic, non-routine interactive, routine cognitive, routine manual, and non-routine manual tasks), while in this paper, I am only interested in whether a job is classified as an interactive job. Following Spitz-Oener (2006), a task measure is calculated as the number of activities in the task category performed by each person divided by the total number of activities in the task category.[39] In this paper, a job is classified as an 'interactive' job if more than 70 percent of the respondents in this job report to do interactive tasks at least sometimes and if their task measure is higher than 0.5 (indicating that people do at least 50 percent of all interactive tasks).[40] Following this definition, nearly 30 percent of all employed women work in interactive jobs. According to the productivity hypothesis, overweight and obese women are supposed to face about the same wage reductions in interactive and non-interactive jobs if they were less productive than women of healthy weight. Higher wage cuts for overweight and obese women in interactive jobs would in turn favor the discrimination hypothesis, indicating that overweight women in jobs with contact to customers (interactive jobs) face higher wage reductions than overweight women in non-interactive jobs.

Third, I divide the dataset of all women into employed and self-employed women. By definition, self-employed women cannot be discriminated by their employer in terms of receiving lower wages. If overweight and obese women were less productive than women of healthy weight (productivity hypothesis), similar effects of weight on wages for both groups are expected: employed and self-employed women. But if overweight women are being discriminated by their employer and therefore face lower wages (discrimination hypothesis),

[38] The IAB-BIBB dataset focuses on job descriptions and detailed information on qualification profiles and occupational development. More information is available at http://www.gesis.org/Datenservice/Themen/38Beruf.htm.

[39] For a more detailed description of tasks and task-measures, please see Spitz-Oener (2006) p.242f.

[40] Results remain robust against using different definitions of ‚interactive jobs'.

one would expect to estimate a negative effect of weight on wages for employed women and no such effect for self-employed women. Nevertheless, selection could be a problem in this case. For example, Garcia and Quintana-Domeque (2006) find in their study on several European countries that obese women have a higher probability to be self-employed in countries like Greece, Ireland, and Italy, which might be because they do not find an employment relationship due to their obesity. On the other hand, there is no correlation between self-employment and obesity in other European countries (Austria, Belgium, Denmark, Finland, Portugal, and Spain). Also for Germany (see Appendix 3.3), these findings do not seem to hold: there is no difference in average weight or BMI between employed and self-employed women. In contrast, self-employed women have, on average, a lower BMI (23.4 compared to 23.7), although this difference is not statistically significantly different from zero. Summary statistics of employed and self-employed women can be found in Appendix 3.3 (first and second column).

Fourth and lastly, the dataset of employed women is divided into young and older women. Young women are defined to be between 20 and 39 years of age and older women between 40 and 55. Both groups have roughly the same number of observations. If one thinks of discrimination against overweight women, we would expect larger coefficients for younger women, since they could not demonstrate their competence yet, while older overweight women might already achieved a higher position and proved that possible prejudices against them are wrong. Again, if the productivity hypothesis was true, the effect of overweight on wages for young and older women should be about the same (it might also be larger for older women, since their accumulated lifetime-productivity is lower). If the effect of weight on wages is higher for young women, the discrimination hypothesis might fit better.

3.4 Results

Summary statistics of all relevant variables are shown in Table 3.2. Descriptive statistics already illustrate that wages are lower for overweight or obese women, while this does not hold for overweight or obese men. In general, overweight and obesity is associated with lower education, which is also found in Zhang and Wang (2004). It is a bit surprising that heavier women have a higher probability to be married, while Averett and Korenman (1996) show that for women, a higher BMI is associated with a lower probability to be married. This is probably due to the fact that overweight and obese persons are, on average, older than healthy- or underweight persons and that these summary statistics do not control for age effects.

First results test the hypothesis that a higher weight is correlated with lower wages. Previous studies for different countries find that this relationship only holds for women while there is no such effect for men (Johansson et al. (2007), Paraponaris (2005), Cawley (2004), Averett and Korenman (1996), Sargent and Blanchflower (1994)). Germany does not seem to be an exception; results are found in Table 3.3 and Table 3.4. The first column in Table 3.3 shows the correlation between weight and wages using BMI as a measure of weight without any further control variables. The effect for women is large and negative, while there is a positive relationship between BMI and wages for men. Both effects are highly significant. If there were no other factors that had an influence on wages, women would earn 0.8 percent lower wages for each higher unit of BMI. In order to increase the BMI by one unit, a person has to gain about 3 kilograms. This means that a woman who is 30kg heavier than the average woman earns 8 percent less. In contrast, a 30kg heavier man would earn 3.5 percent *more* than a man having an average BMI. But since there are many factors, besides BMI, that influence wages, further control variables are added stepwise. Most importantly, educational controls are included in column 2. Schooling is included in three categories (lower, intermediate and higher education – ‘Hauptschule’, ‘Realschule’, and ‘Gymnasium’) and further vocational education in four categories (no apprenticeship, apprenticeship, master craftsman, university degree). For women, the effect becomes much smaller, which means that heavier women tend to have a lower education. This is different for men: the effect becomes larger when education is controlled for, which shows that heavier men have, on average, a higher education.[41] But keeping in mind that a man of average weight is already overweight, this is not very surprising. In the third column, personal characteristics are added (age and age squared, dummy for having children, marital status, nationality and state of residence). The last column shows the estimates for the full set of control variables, adding job characteristics such as dummies for temporary work contract and for fulltime employment, tenure, tenure squared, 14 industry dummies, dummies for blue- and white-collar workers, civil servants, firm size, usual hours of work per week and usual hours of work squared. The coefficient for women remains unchanged, while the effect for men becomes smaller and insignificant as personal and job characteristics are included. For men there seems to be no relationship between BMI and wages, but women with a higher BMI earn significantly less. An increase in weight of about 30kg (or 10 units of BMI) is associated with 2.6 percent lower wages for women. In the following, all results will include the full set of control variables and only the coefficients of the weight variables are shown.

Results for the effects of all other control variables on wages are shown in Appendix 3.1. As expected, schooling is positively related to wages, people in the eastern part of Germany earn less, wages increase with age and tenure, white-collar workers and civil servants earn more than blue-collar workers, and larger firms pay higher wages. Most of the other control variables also have expected signs.

Besides BMI as explanatory variable to measure a person's physical appearance, I also use dummies for the clinical categories 'underweight', healthy weight', 'overweight' and 'obese' as regressors, with healthy weight as base category. Since measures of weight and

[41] The term 'effect' should not be interpreted as a causal effect of weight on wages, but rather in the sense of a correlation.

height are self-reported, the third alternative to measure ‘weight’ is to use dummies for the 10 deciles of the distribution of BMI as explanatory variables. By doing so, not the person’s reported BMI matters, but the BMI of a person in relation to other persons is of interest. If all data on weight and height are biased in the same way (by understating weight and overstating height), this procedure is a good indicator for whether heavier persons earn less. Table 3.4 shows results for the different measures of weight, separately for men and women. The table contains coefficients of three different regressions using BMI, the clinical categories of weight, or the deciles of BMI as explanatory variables of interest. Obviously, there is virtually no effect for overweight or obese men in all three regressions, but effects for women become much stronger for heavier women. This can be seen by comparing the coefficients of ‘overweight’ and ‘obese’. While the wage reduction for overweight women is 1.3 percent, it is almost twice as much for obese women (2.4 percent). By using the deciles of the distribution of BMI as regressors, it can be observed that higher values of BMI (the highest 20 percent) are associated with much higher wage reductions (about 4 percent lower wages) than for those women between the 40^{th} and the 80^{th} percentile (about 2.5 percent lower wages).

For women, there is no relationship between underweight and wages. This finding could be interpreted as supportive of the discrimination hypothesis: Assuming that, from a health point of view, underweight is as unhealthy as overweight. Then, both underweight and overweight women should receive lower wages if they were less productive than women of healthy weight. However, since underweight is socially accepted (if not preferred, since it represents the current ideal of beauty), we only observe lower wages for overweight women, which supports the discrimination hypothesis. In contrast to women, underweight seems to be strongly negatively related to wages for men. Being male and underweight is associated with about 3-times the wage reduction of being female and obese. The problem with these numbers is that only 0.6 percent (or n=416) men are classified as underweight. Therefore, it might be better to compare men and women in the lowest (highest) deciles of BMI since each decile has more than 7,000 and 6,000 observations for men and women, respectively. The negative effect of belonging to the thinnest 10 percent for men is as high as being in the highest decile for women (more than 4 percent lower wages in both cases). But since this paper focuses on the relationship between overweight and wages, I will rely on results for women only, since there is no such correlation for men. Potential discrimination of underweight men is not part of this study.

3.4.1 Results by gender dominance in the workplace

To come back to find evidence for the productivity or the discrimination hypothesis, previous results for employed women are compared to results estimated separately by gender dominance in the workplace. As motivated above, I expect roughly the same coefficients in all three categories or lower effects for male-dominated jobs if the productivity hypothesis holds, because there should be no difference in productivity between these three working environments. Contrary, I interpret results as supportive of the discrimination hypothesis if coefficients of male-dominated jobs are larger, because literature on discrimination reveals that men are more likely to discriminate against overweight women than women (Harris et al. (1991)). Summary statistics (in Appendix 3.2, first three columns) show that there are no major differences between the four clinical weight categories in male-dominated, female-dominated, or male-female-balanced jobs; which can also be seen when comparing the average values of BMI at 23.8, 23.7, and 23.5, respectively. Thus, there seems to be no sorting of overweight women into a particular working environment.

Table 3.5 provides the results by gender dominance in the workplace. For comparison, results for all employed women are repeated in the first column. The coefficient of BMI in male-dominated jobs is almost twice as large as the average effect. Moreover, it is 1.5 times and 2.5 times larger than for male-female balanced and female-dominated jobs, respectively.

The same patterns are true for the clinical categories 'overweight' and 'obese': Correlations are strongest for male-dominated jobs. The reason why large effects (for example for obese women in male-dominated jobs) are not statistically significant is probably due to the fact that only few women work in male-dominated jobs (by definition), so that sample size is much smaller here. Nevertheless, taking deciles of BMI as regressors reveals that from the 70th percentile onwards effects for women in male-dominated jobs are two to three times larger than the coefficients of these deciles in female-dominated and in male-female balanced jobs. Thus, results can be interpreted as first support of the discrimination hypothesis, since effects are much stronger in male-dominated jobs.

3.4.2 Results: interactive vs. non-interactive jobs

Following the arguments above, there should be no difference in wage cuts for overweight women in jobs with contact to customers (interactive jobs) and those without customer contact (non-interactive jobs). In this classification, problems might arise if overweight women tend to sort themselves into non-interactive jobs while women of healthy weight prefer working in interactive jobs. Summary statistics in Appendix 3.2 (last two columns) show that women in non-interactive jobs tend to be heavier than women in interactive jobs. However, differences are not statistically significant, so that any sorting effects will be neglected. Moreover, differences in weight between women in interactive and non-interactive jobs could also be explained by differences in socioeconomic status. As stated in the literature (Zhang and Wang (2004)), social status is negatively related to weight. Appendix 3.2 reveals that women in non-interactive jobs have, on average, a lower status (lower education, more blue-collar worker) than women in interactive jobs, which is correlated with a higher probability to be overweight.

Table 3.6 reveals that there are huge differences in wage reductions for overweight women in interactive jobs compared to overweight women in non-interactive jobs. The effect of an increase in BMI of 10 points is more than three times higher for women in interactive jobs. They face 5.1 percent lower wages while women in non-interactive jobs only receive a wage reduction of 1.6 percent for each 10-point increase in BMI. The same patterns hold when taking weight categories as explanatory variables: Obese women in interactive jobs earn 5.1 percent lower wages; underweight women working in interactive jobs earn 4.6 percent more than women of normal weight. There are no such effects for women in non-interactive jobs. When turning to the deciles of BMI distribution, the effect is strongest for women in interactive job belonging to the heaviest ten percent. Nevertheless, effects for women in non-interactive jobs are also strong, all of them being negative and most of them statistically significant. Fewer significant results for women in interactive jobs are probably due to smaller samples size for women in interactive jobs.

However, the most striking difference is that the thinnest ten percent receive higher wages in interactive jobs while they receive lower wages in non-interactive jobs. One explanation for this finding could be that belonging to the thinnest ten percent is not healthy (and therefore penalized in non-interactive jobs), nevertheless being very thin comes close to the current ideal of beauty which is why women in interactive jobs (with contact to customers) might receive higher wages. Many jobs with contact to customers are jobs in sales, so that an attractive look which is close to the current ideal of beauty (namely, very thin), can be productivity-enhancing and thus rewarded by paying a ‘beauty premium’.[42] The question is, whether results in Table 3.6 can be interpreted as discrimination by the employer or whether results are driven by customer discrimination. In this case, customers prefer to buy from more attractive salespeople, which in turn makes them more productive than overweight persons. Nevertheless, there are two reasons, why lower wages for overweight women are due

[42] This term was introduced by Hamermesh and Biddle (1994), who find that attractive persons earn higher wages while plain people earn less.

to discrimination by the employer and not due to customer discrimination (which could be interpreted as lower productivity of overweight women in interactive jobs). First, being overweight is not necessarily related to a less attractive appearance. Therefore, it cannot be concluded that thin people receive a 'beauty premium' while overweight persons face a 'homely-penalty'. Second, when focusing on self-employed women, the only type of discrimination they could possibly face is customer discrimination. The next section shows that this is not the case. Therefore, I interpret results in Table 3.6 as supportive of the discrimination hypothesis.

3.4.3 Results: employed vs. self-employed women

In this section the dataset of all working women is divided into employed and self-employed women. If overweight women receive lower wages due to lower productivity, there should be no difference in the weight-coefficients between employed and self-employed women, since in this case overweight self-employed women should also be less productive than self-employed women of healthy weight and therefore supposed to earn lower wages. If the discrimination hypothesis holds, one might expect a negative effect of weight on wages for employed women only. There should be no effect for self-employed women since there is no employer who might discriminate them because of prejudices against heavier persons.

In general, self-employed women differ from employed women when it comes to wages. The share of self-employed women with very low or very high wages is much higher, whereas there are less self-employed women with intermediate wages. Summary statistics of wages and the explanatory variables for employed and self-employed women are shown in the first two columns of Appendix 3.3 (first two columns). Self-employed women are, on average, older than employed women. Moreover, they are better educated, more likely to be married, more likely to work fulltime and have more working hours per week than employed women. Controlling for these differences, Table 3.7 shows clearly that heavier self-employed women do not receive lower wages than their healthy weight counterparts, while significant differences are found for employed women.

For self-employed women, the point estimate for BMI is positive, though not significantly different from zero. The absolute value of the coefficient of BMI for employed women is of similar size: a 10-point increase of BMI is associated with 2.6 percent lower wages for employed women, while it is associated with 2.2 percent higher wages for self-employed women (not significant). Results for the clinical categories also go into this direction: while employed obese and overweight women earn 2.4 and 1.3 percent less than employed women of healthy weight, respectively, obese and overweight self-employed women earn more than self-employed women of healthy weight, although the effect for self-employed women is not significant. These results also favor the discrimination hypothesis, indicating than overweight women are discriminated by their employer in terms of receiving lower wages than women of healthy weight.

3.4.4 Results: young vs. older women

In a last step, I divide the dataset of employed women into young and older women. Different results for these two groups might also be an indicator for whether the productivity or the discrimination hypothesis fits better. As motivated in Section 3.3, one would expect larger effects of weight on wages for younger women if the discrimination hypothesis is favorable, while coefficients are supposed to be of equal size (or larger for older women) in order to support the productivity hypothesis. Summary statistics of young and older women can be found in column 3 and 4 of Appendix 3.3. As expected, older women are on average notably heavier than young women. Nevertheless, this does not seem to result in larger coefficients for older women. Table 3.8 provides the results for young and older women for all three measures of body weight.

Supporting the discrimination hypothesis, results are indeed stronger for young than for older women. In both groups, I find a significant negative association between weight and wages. The coefficient of BMI indicates that young women receive 2.9 percent lower wages for a 10-point increase of BMI, which is almost 50 percent larger than the effect for older women (at 2 percent for a 10-point higher BMI). For the clinical category ‘obese’ effects for young and older women are about equal. A reason for the insignificance of the ‘obese’-dummy for younger women might be larger standard errors since there are only few women who are young and obese (in the sample are twice as many obese older women than obese young women). Interestingly, the clinical category ‘overweight’ is not associated with lower wages for older women, while there is a large and significant effect for young women (2.4 percent lower wages for overweight young women compared to young women of healthy weight). This could be due to the fact that older women have, on average, a higher BMI. If all older women were overweight, there would be no discrimination among them. But in this sample, the average BMI of older women is 24.5 which is classified as ‘healthy weight’. The coefficients for the deciles also indicate that there are larger effects for young women. Women above the 7th decile in the BMI distribution receive 3.2 to 4.8 percent lower wages when young, while results show between 0.8 and 3.2 percent lower wages for the heaviest 30 percent older women.

Table 3.9 combines the classification into young and older women with the other three classifications (into employed vs. self-employed women, by gender dominance in the workplace into interactive and non-interactive jobs). Previous results hold, since the expected results to support the discrimination hypothesis are again more pronounced for young women than for older women in all groups: The correlation of BMI and wages for young self-employed women becomes much larger and remains positive (0.011 for young and self-employed women compared to -0.0006 for older self-employed women). Coefficients for young women in male-dominated job are also much larger in absolute value than for older women in male-dominated jobs (6 percent wage reduction for a 10-point increase in BMI compared to 3.3 percent). This puts young women who work in male-dominated jobs in a position with most discrimination against overweight and obese women. Comparing the effects of overweight on wages in interactive and non-interactive jobs between young and older women, it is found that young women in interactive job face much higher wage reductions than all other age-job combinations (7.6 percent lower wages compared to 1.8 to 3.5 percent for the other three categories). All in all, it can be summarized that all four subgroup designs favor the discrimination hypothesis over the productivity hypothesis.

3.5 Conclusions of Chapter 3

This paper analyses the relationship between weight and wages, especially for women, who seem to face higher wage reductions when they are overweight or obese. I find that obese women receive 2.4 percent lower wages than women of healthy weight, while women who are in the top 10 percent of the body mass index get 4.3 percent lower wages than thinner women. Based on papers that find a causal effect of overweight on wages for (white) women, this paper asks the question whether observed differences in wages between overweight women and women of healthy weight are due to reduced productivity of heavier women or due to discrimination. I set up four different subgroup designs to find support for either the productivity or the discrimination hypothesis. In all four subgroup designs, estimates are in favor of the discrimination hypothesis: First, overweight or obese women in jobs with mainly male coworkers seem to experience higher wage reductions than overweight or obese women with mainly female coworkers or in male-female balanced jobs. To find support for the productivity hypothesis, one would expect roughly the same coefficients in all working environments. Second, women in interactive jobs (e.g. jobs with a lot of contact to customers, clients or co-workers) face higher wage cuts when being overweight than women in non-interactive jobs. Third, in contrast to overweight employed women, overweight self-employed women do not experience any wage reductions. Fourth, the association between lower wages and higher weight is stronger for young women in contrast to older women, although one would expect the same wage reductions if overweight women were generally less productive. To conclude, all four subgroup designs favor the discrimination hypothesis over the productivity hypothesis, indicating that overweight women receive lower wages although they are not less productive than women of normal weight.

References for Chapter 3

Averett, S. and S. Korenman (1996) The Economic Reality of the Beauty Myth, Journal of Human Resources, Vol. 31(2), 304-330.

Baum, C.L. and W.F. Ford (2004) The Wage Effects of Obesity: A Longitudinal Study, Health Economics, Vol. 13(9), 885–899.

Behrman, J.R. and M.R. Rosenzweig (2001) The Returns to Increasing Body Weight, PIER Working Paper No. 01-052.

Brunello, G. and B. D`Hombres (2006) Does Body Weight Affect Wages? Evidence from Europe, Università Degli Studi Di Padova, Dipartimento di Scienze Economiche 'Marco Fanno', Working Paper No. 27.

Cawley, J. (2000) Body Weight and Women's Labor Market Outcomes, NBER Working Paper No. 7841.

Cawley, J. (2004) The Impact of Obesity on Wages, Journal of Human Resources, Vol. 39(2), 451-474.

Cawley, J. and S. Danziger (2004) Morbid Obesity and the Transition From Welfare to Work, Journal of Policy Analysis and Management, Vol. 24(4), 727-743.

Cawley, J., M.M. Grabka, and D.R. Lillard (2005) A Comparison of the Relationship Between Obesity and Earnings in the U.S. and Germany, Schmollers Jahrbuch, Vol. 125(1), 119-129.

Chou, S.-Y., M. Grossman, and H. Saffer (2004) An economic analysis of adult obesity: results from the Behavioral Risk Factor Surveillance System, Journal of Health Economics, Vol. 23(3), 565–587.

Conley, D. and R. Glauber (2005) Gender, Body Mass and Economic Status, NBER Working Paper No. 11343.

Costa-Font, J. and J. Gil (2004) Social interactions and the contemporaneous determinants of individuals' weight, Applied Economics, Vol. 36(20), 2253–2263.

Fahr, R. (2006) The Wage Effects of Social Norms: Evidence of Deviations from Peers' Body-Mass in Europe, IZA Discussion Paper No.2323.

Garcia, J. and C. Quintana-Domeque (2006) Obesity, Employment and Wages in Europe, Advances in Health Economics and Health Services Research, Vol. 17, 187–217.

Gortmaker, S.L., A. Must, J.M. Perrin, A.M. Sobol, and W.H. Dietz (1993) Social and Economic Consequences of Overweight in Adolescence and Young Adulthood, New England Journal of Medicine, Vol. 329(14), 1008-1012.

Hamermesh, D.S. and J.E. Biddle (1994) Beauty and the Labor Market, American Economic Review, Vol. 84(5), 1174-1194.

Harris, M.B., L.C. Walters, and S. Waschull (1991) Gender and Ethnic Differences in Obesity-Related Behaviors and Attitudes In a College Sample, Journal of Applied Social Psychology, Vol. 21(19), 1545-1566.

Johansson, E., P. Böckerman, U. Kiiskinen, and M. Heliövaara (2007) The Effect of Obesity on Wages and Employment: The Difference between Having a High BMI and Being Fat, Meddelanden Working Papers No. 528.

Lakdawalla, D. and T. Philipson (2002) The Growth of Obesity and Technological Change: A Theoretical and Empirical Examination, NBER Working Paper No. 8946.

Pagán, J.A. and A. Dávlia (1997) Obesity, Occupational Attainment, and Earnings, Social Science Quarterly, Vol. 8(3), 756-770.

Paraponaris, A. (2005) Obesity, Weight Status and Employability: Empirical Evidence from a French National Survey, Economics & Human Biology, Vol. 3(2), 241-258.

Register, C.A. and D.R. Williams (1990) Wage Effects of Obesity among Young Workers, Social Science Quarterly, Vol. 71(1), 130-141.

Robert, S.A. and E.N. Reither (2004) A multilevel analysis of race, community disadvantage, and body mass index among adults in the U.S., Social Science & Medicine, Vol. 59(12), 2421-2434.

Roehling, M.V. (1999) Weight-based discrimination in employment: Psychological and legal aspects, Personnel Psychology, Vol. 52(4), 969-1016.

Sargent, J.D., and D.G. Blanchflower (1994) Obesity and Stature in Adolescence and Earnings in Young Adulthood. Analysis of a British Cohort, Archives of Pediatrics and Adolescent Medicine, Vol. 148(7), 681-687.

Sobal and Stunkard (1989) Socio-economic Status and Obesity: A review of the Literature, Psychological Bulletin, Vol. 105(2), 260-275.

Spitz-Oener (2006) Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure, Journal of Labor Economics, Vol. 24(2), 235-270.

Zhang, Q. and Y. Wang (2004) Trends in the Association between Obesity and Socioeconomic Status in U.S. Adults: 1971 to 2000, Obesity Research, Vol. 12(10), 1622-1632.

Tables of Chapter 3

Table 3-1: Summary statistics (I)

	Women	Men
BMI	23.7	25.9
Weight in kg	66.1	83.4
Height in cm	167	180
Underweight	0.04	0.01
Healthy weight	0.67	0.45
Overweight	0.21	0.43
Obese	0.08	0.12
n	63,388	74,416

Source: Micro Census 2003 and 2005, own calculations.

Table 3-2: Summary statistics (II)

	Women		Men	
	healthy- or underweight	overweight or obese	healthy- or underweight	Overweight or obese
Net wage				
Wage (300-700)	0.20	0.23	0.03	0.02
Wage (700-1100)	0.26	0.26	0.12	0.11
Wage (1100-1700)	0.44	0.43	0.52	0.52
Wage (1700-2900)	0.08	0.06	0.21	0.23
Wage (2900-4000)	0.02	0.01	0.08	0.08
Wage (>4000)	0.01	0.00	0.04	0.04
Education				
Lower secondary school	0.20	0.29	0.29	0.39
Intermediate secondary school	0.46	0.49	0.35	0.35
Higher secondary school	0.33	0.22	0.36	0.26
No further education	0.01	0.01	0.01	0.01
Apprenticeship	0.81	0.87	0.76	0.83
University degree	0.18	0.11	0.23	0.16
Personal Characteristics				
Age	37.8	41.8	37.2	41.2
Married	0.56	0.65	0.52	0.69
Children	0.33	0.28	0.34	0.37
German nationality	0.95	0.96	0.93	0.94
Living in West-Germany	0.80	0.74	0.83	0.80
Living in East-Germany	0.20	0.26	0.17	0.20
Job characteristics				
White-collar worker	0.80	0.72	0.54	0.47
Blue-collar worker	0.13	0.23	0.38	0.45
Civil servant	0.07	0.05	0.07	0.07
Temporary work contract	0.08	0.07	0.08	0.05
Fulltime employed	0.62	0.61	0.96	0.97
Hours worked per week	31.9	31.7	39.1	39.5
Tenure	8.76	10.28	9.29	11.79
Firm size (1-10)	0.26	0.24	0.15	0.13
Firm size (11-49)	0.27	0.28	0.25	0.25
Firm size (> 50)	0.48	0.48	0.60	0.61
n	44,856	18,532	34,067	41,349

Source: Micro Census 2003 and 2005, own calculations.

Table 3-3: BMI as explanatory variable, stepwise including control variables

	Dep. Variable: net ln(wage)			
Women (n=63,388)	**1.**	**2.**	**3.**	**4.**
BMI	-0.0080***	-0.0020***	-0.0023***	-0.0026***
	(0.0008)	(0.0008)	(0.0007)	(0.0006)
controlled for education		x	x	x
controlled for personal characteristics			x	x
controlled for job characteristics				x
Men (n=75,416)	**1.**	**2.**	**3.**	**4.**
BMI	0.0035***	0.0086***	0.0016***	0.0009
	(0.0008)	(0.0007)	(0.0006)	(0.0005)
controlled for education		x	x	x
controlled for personal characteristics			x	x
controlled for job characteristics				x

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. The estimates stepwise increase the set of control variables. In the first column, no control variables are included. The second column controls for education: lower, intermediate, higher education, university degree and apprenticeship. In the third column, personal characteristics are added: age and age squared, children, marital status, nationality and state, while the last column shows the estimates for the full set of control variables, adding job characteristics such as temporary work contract, fulltime working, tenure, tenure squared, industry, dummy for white-collar worker, civil servant, firm size, usual hours of work per week and usual hours of work squared.
Source: Micro Census 2003 and 2005, own calculations.

Table 3-4: Results for women and men

	Women	Men
BMI	-0.0026***	0.0009
	(0.0006)	(0.0005)
Obese – BMI higher than 30	-0.024***	0.004
	(0.008)	(0.006)
Overweight – BMI between 25 and 30	-0.013**	0.005
	(0.006)	(0.004)
Underweight – BMI lower than 18.5	0.000	-0.078***
	(0.013)	(0.026)
BMI - 10. Percentile	-0.007	-0.044***
	(0.011)	(0.009)
BMI - 20. Percentile	-0.009	-0.020**
	(0.011)	(0.008)
BMI - 40. Percentile	-0.023**	-0.011
	(0.010)	(0.008)
BMI - 50. Percentile	-0.015	-0.004
	(0.011)	(0.009)
BMI - 60. Percentile	-0.032***	-0.014*
	(0.011)	(0.008)
BMI - 70. Percentile	-0.025**	-0.008
	(0.010)	(0.009)
BMI - 80. Percentile	-0.022**	-0.013
	(0.011)	(0.008)
BMI - 90. Percentile	-0.037***	-0.016*
	(0.011)	(0.008)
BMI - 100. Percentile	-0.043***	-0.014
	(0.010)	(0.009)
n	63,388	75,416

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. The estimates are based on regressions with the following set of control variables: Age and age squared, education (lower, intermediate, higher education), university degree, apprenticeship, children, marital status, nationality, state, temporary work contract, fulltime working, tenure, tenure squared, industry, dummy for white-collar worker, civil servant, firm size, usual hours of work per week and usual hours of work squared.
Source: Micro Census 2003 and 2005, own calculations.

Table 3-5: Results based on gender-dominance in the job

	Women (all)	Women in male-dominated jobs	Women in female-dominated jobs	Women in male-female balanced jobs
BMI	-0.0026***	-0.0048**	-0.0019**	-0.0032***
	(0.0006)	(0.0020)	(0.0007)	(0.0011)
Obese – BMI higher than 30	-0.024***	-0.033	-0.019*	-0.034**
	(0.009)	(0.029)	(0.011)	(0.016)
Overweight – BMI between 25 and 30	-0.013**	-0.045**	-0.008	-0.017
	(0.006)	(0.021)	(0.008)	(0.011)
Underweight – BMI lower than 18.5	0.000	0.038	-0.012	0.006
	(0.013)	(0.052)	(0.015)	(0.022)
BMI - 10. Percentile	-0.007	-0.017	-0.018	0.001
	(0.011)	(0.035)	(0.014)	(0.019)
BMI - 20. Percentile	-0.009	-0.017	-0.006	-0.020
	(0.011)	(0.038)	(0.013)	(0.019)
BMI - 40. Percentile	-0.023***	-0.064*	-0.013	-0.051***
	(0.010)	(0.034)	(0.014)	(0.018)
BMI - 50. Percentile	-0.015	-0.077**	-0.020	-0.002
	(0.011)	(0.035)	(0.013)	(0.019)
BMI - 60. Percentile	-0.032***	-0.058	-0.033**	-0.008
	(0.011)	(0.039)	(0.014)	(0.018)
BMI - 70. Percentile	-0.025**	-0.070*	-0.025*	-0.033*
	(0.010)	(0.036)	(0.014)	(0.019)
BMI - 80. Percentile	-0.022**	-0.097***	-0.014	-0.021
	(0.011)	(0.035)	(0.014)	(0.019)
BMI - 90. Percentile	-0.037***	-0.085**	-0.026*	-0.040**
	(0.011)	(0.040)	(0.014)	(0.018)
BMI - 100. Percentile	-0.043***	-0.080**	-0.042***	-0.052**
	(0.010)	(0.036)	(0.013)	(0.019)
n	63,388	5,631	37,152	20,605

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. An allocation into male and female-dominated job is based on the proportion of men and women in the different jobs. If more than 70 percent of the employees in a job category are male or female, this job is labeled male or female-dominated, respectively. The remaining jobs are labeled male-female balanced jobs. This division is based on a job classification into 342 jobs. The estimates are based on regressions with the following set of control variables: Age and age squared, education (lower, intermediate, higher education), university degree, apprenticeship, children, marital status, nationality, state, temporary work contract, fulltime working, tenure, tenure squared, industry, dummy for white-collar worker, civil servant, firm size, usual hours of work per week and usual hours of work squared.
Source: Micro Census 2003 and 2005, own calculations.

Table 3-6: Results: interactive vs. non-interactive jobs

	Interactive job	Non-interactive job
BMI	-0.0051***	-0.0016**
	(0.0012)	(0.0007)
Obese – BMI higher than 30	-0.051**	-0.015
	(0.020)	(0.009)
Overweight – BMI between 25 and 30	-0.008	-0.014**
	(0.012)	(0.007)
Underweight – BMI lower than 18.5	0.046*	-0.017
	(0.024)	(0.015)
BMI - 10. Percentile	0.039**	-0.029**
	(0.019)	(0.013)
BMI - 20. Percentile	0.007	-0.017
	(0.019)	(0.013)
BMI - 40. Percentile	-0.018	-0.027**
	(0.018)	(0.013)
BMI - 50. Percentile	-0.012	-0.019
	(0.019)	(0.013)
BMI - 60. Percentile	-0.039*	-0.032**
	(0.020)	(0.013)
BMI - 70. Percentile	-0.017	-0.030**
	(0.020)	(0.012)
BMI - 80. Percentile	-0.008	-0.028**
	(0.020)	(0.013)
BMI - 90. Percentile	-0.024	-0.043***
	(0.021)	(0.012)
BMI - 100. Percentile	-0.061***	-0.038***
	(0.021)	(0.012)
n	18,501	44,887

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. The estimates are based on regressions with the following set of control variables: Age and age squared, education (lower, intermediate, higher education), university degree, apprenticeship, children, marital status, nationality, state, temporary work contract, fulltime working, tenure, tenure squared, industry, dummy for white-collar worker, civil servant, firm size, usual hours of work per week and usual hours of work squared.
Source: Micro Census 2003 and 2005, own calculations.

Table 3-7: Results: employed vs. self-employed women

	Women (employed)	Women (self-employed)
BMI	-0.0026***	0.0022
	(0.0006)	(0.0050)
Obese – BMI higher than 30	-0.024***	0.042
	(0.008)	(0.079)
Overweight – BMI between 25 and 30	-0.013**	0.011
	(0.006)	(0.047)
Underweight – BMI lower than 18.5	0.000	-0.049
	(0.013)	(0.098)
BMI - 10. Percentile	-0.007	0.003
	(0.011)	(0.077)
BMI - 20. Percentile	-0.009	-0.068
	(0.011)	(0.076)
BMI - 40. Percentile	-0.023***	-0.131*
	(0.010)	(0.074)
BMI - 50. Percentile	-0.015	-0.149**
	(0.011)	(0.075)
BMI - 60. Percentile	-0.032***	-0.025
	(0.011)	(0.075)
BMI - 70. Percentile	-0.025**	-0.107
	(0.010)	(0.079)
BMI - 80. Percentile	-0.022**	-0.066
	(0.011)	(0.076)
BMI - 90. Percentile	-0.037***	-0.025
	(0.011)	(0.082)
BMI - 100. Percentile	-0.043***	-0.052
	(0.010)	(0.081)
n	63,388	5,172

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. The estimates are based on regressions with the following set of control variables: Age and age squared, education (lower, intermediate, higher education), university degree, apprenticeship, children, marital status, nationality, state, temporary work contract, fulltime working, tenure, tenure squared, industry, dummy for white-collar worker, civil servant, firm size, usual hours of work per week and usual hours of work squared.
Source: Micro Census 2003 and 2005, own calculations.

Table 3-8: Results: young vs. older women

	Women (20-39 years of age)	Women (40-55 years of age)
BMI	-0.0029***	-0.0020**
	(0.0009)	(0.0008)
Obese – BMI higher than 30	-0.021	-0.022**
	(0.014)	(0.011)
Overweight – BMI between 25 and 30	-0.024***	-0.003
	(0.009)	(0.008)
Underweight – BMI lower than 18.5	-0.005	0.004
	(0.014)	(0.026)
BMI - 10. Percentile	-0.007	-0.009
	(0.013)	(0.018)
BMI - 20. Percentile	-0.016	0.007
	(0.013)	(0.017)
BMI - 40. Percentile	-0.025*	-0.021
	(0.014)	(0.016)
BMI - 50. Percentile	-0.030**	0.003
	(0.014)	(0.016)
BMI - 60. Percentile	-0.041***	-0.017
	(0.014)	(0.016)
BMI - 70. Percentile	-0.007	-0.036**
	(0.015)	(0.015)
BMI - 80. Percentile	-0.032**	-0.008
	(0.015)	(0.015)
BMI - 90. Percentile	-0.045***	-0.025*
	(0.016)	(0.015)
BMI - 100. Percentile	-0.048***	-0.032**
	(0.015)	(0.015)
n	30,670	32,718

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. The estimates are based on regressions with the following set of control variables: Age and age squared, education (lower, intermediate, higher education), university degree, apprenticeship, children, marital status, nationality, state, temporary work contract, fulltime working, tenure, tenure squared, industry, dummy for white-collar worker, civil servant, firm size, usual hours of work per week and usual hours of work squared.
Source: Micro Census 2003 and 2005, own calculations.

Table 3-9: All results by age group

	1. by employment status		2. by gender dominance in the workplace			3. by contact to customer	
	Employed women	**Self-employed women**	**Male-dominated jobs**	**Female-dominated jobs**	**Male-female balanced jobs**	**Interactive jobs**	**Non-interactive jobs**
BMI	-0.0029***	0.0106	-0.0060**	-0.0021*	-0.0035**	-0.0076***	-0.0035**
Young women (20-39)	(0.0009)	(0.0098)	(0.0028)	(0.0011)	(0.0017)	(0.0017)	(0.0017)
n	30,670	1,886	2,935	17,560	10,175	10,048	20,622
BMI	-0.0020**	-0.0006	-0.0033	-0.0014	-0.0029**	-0.0020	-0.0018**
Older women (40-55)	(0.0008)	(0.0058)	(0.0026)	(0.0010)	(0.0014)	(0.0018)	(0.0009)
n	32,718	3,286	2,696	19,592	10,430	8,453	24,265

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. The estimates are based on regressions with the following set of control variables: Age and age squared, education (lower, intermediate, higher education), university degree, apprenticeship, children, marital status, nationality, state, temporary work contract, fulltime working, tenure, tenure squared, industry, dummy for white-collar worker, civil servant, firm size, usual hours of work per week and usual hours of work squared.
Source: Micro Census 2003 and 2005, own calculations.

Appendix to Chapter 3

Appendix 3-1: Results including all controls

	Women	Men
BMI	-0.0026***	0.0009
	(0.0006)	(0.0005)
Education and personnel characteristics		
Lower secondary school	-0.067***	-0.051***
	(0.007)	(0.005)
Higher secondary school	0.065***	0.052***
	(0.008)	(0.007)
No apprenticeship	-0.029	-0.070***
	(0.021)	(0.017)
Master craftsman	0.084***	0.086***
	(0.008)	(0.006)
University degree	0.214***	0.262***
	(0.011)	(0.009)
Children (<3)	-0.172***	0.083***
	(0.017)	(0.007)
Children (3-5)	0.024**	0.104***
	(0.012)	(0.007)
Children (6-9)	0.033***	0.087***
	(0.010)	(0.006)
Children (10-14)	0.027***	0.075***
	(0.007)	(0.005)
Children (>14)	-0.025***	0.017***
	(0.005)	(0.005)
Married	-0.162***	0.116***
	(0.006)	(0.005)
Nationality: German	0.035*	0.089***
	(0.018)	(0.013)
Nationality: Non-German, EU	0.043	0.088***
	(0.029)	(0.021)
Age	0.039***	0.018***
	(0.002)	(0.002)
Age squared	-0.000***	-0.000***
	(0.000)	(0.000)
State - Reference: North Rhine-Westphalia		
Schleswig-Holstein	0.000	0.004
	(0.016)	(0.013)
Hamburg	0.044**	-0.059***
	(0.019)	(0.019)
Lower Saxony	-0.025**	-0.008
	(0.010)	(0.008)
Bremen	-0.004	-0.039**
	(0.020)	(0.020)
Hesse	0.014	0.005
	(0.011)	(0.008)

Appendix 3-1: Results including all controls (continued)

	Women	Men
State - Reference: North Rhine-Westphalia		
Rhineland-Palatinate	-0.013	0.023***
	(0.012)	(0.009)
Baden-Wurttemberg	0.008	0.034***
	(0.009)	(0.007)
Bavaria	-0.011	0.017**
	(0.009)	(0.007)
Saarland	-0.050**	-0.051***
	(0.019)	(0.013)
Berlin	-0.005	-0.110***
	(0.011)	(0.012)
Brandenburg	-0.083***	-0.244***
	(0.013)	(0.011)
Mecklenburg-West Pomerania	-0.112***	-0.242***
	(0.017)	(0.016)
Saxony	-0.124***	-0.263***
	(0.010)	(0.009)
Saxony-Anhalt	-0.137***	-0.266***
	(0.011)	(0.009)
Thuringia	-0.154***	-0.294***
	(0.014)	(0.012)
Job characteristics		
White-collar worker	0.152***	0.156***
	(0.007)	(0.005)
Civil servant	0.359***	0.266***
	(0.014)	(0.010)
Firm size (1-10)	-0.067***	-0.053***
	(0.007)	(0.006)
Firm size(>50)	0.057***	0.093***
	(0.006)	(0.005)
Tenure	0.005***	0.011***
	(0.001)	(0.001)
Tenure squared	0.000	0.000***
	(0.000)	(0.000)
Temporary work contract	-0.079***	-0.131***
	(0.010)	(0.010)
Working fulltime	0.102***	0.229***
	(0.010)	(0.023)
Hours worked per week	0.037***	0.013***
	(0.001)	(0.002)
Hours worked per week squared	-0.000***	-0.000
	(0.000)	(0.000)

Appendix 3-1: Results including all controls (continued)

	Women	Men
Industry - Reference: Manufacturing		
Agriculture, hunting and forestry	-0.127***	-0.142***
	(0.024)	(0.013)
Electricity, gas and water supply	0.094***	0.016
	(0.030)	(0.016)
Construction	0.016	-0.029***
	(0.022)	(0.006)
Wholesale & retail trade	-0.078***	-0.083***
	(0.009)	(0.007)
Hotel and restaurant industry	-0.135***	-0.223***
	(0.016)	(0.020)
Transport and communication	0.009	-0.089***
	0014)	(0.008)
Financial intermediation	0.054***	0.053***
	(0.012)	(0.012)
Real estate and business activities	-0.001	-0.010
	(0.011)	(0.010)
Public administration and defense	-0.033***	-0.102***
	(0.010)	(0.008)
Education	0.037***	-0.108***
	(0.011)	(0.012)
Health and social work	-0.044***	-0.126***
	(0.009)	(0.010)
Other social and personal service	-0.075***	-0.074***
	(0.013)	(0.012)
n	63388	75416

Source: Micro Census 2003 and 2005, own calculations.

Appendix 3-2: Summary statistics for subgroups (I)

	Male-dominated jobs	Female-dominated jobs	Male-female-balanced jobs	Women in interactive jobs	Women in non-interactive jobs
Net wage					
Wage (300-700)	0.12	0.26	0.15	0.21	0.22
Wage (700-1100)	0.22	0.28	0.23	0.23	0.26
Wage (1100-1700)	0.48	0.40	0.48	0.42	0.43
Wage (1700-2900)	0.14	0.05	0.11	0.10	0.07
Wage (2900-4000)	0.03	0.01	0.03	0.03	0.02
Wage (>4000)	0.02	0.00	0.01	0.01	0.01
Weight					
Underweight	0.04	0.04	0.04	0.04	0.04
Healthy weight	0.65	0.67	0.68	0.71	0.66
Overweight	0.22	0.21	0.21	0.18	0.22
Obese	0.09	0.08	0.07	0.06	0.08
Education					
Lower secondary school	0.21	0.26	0.18	0.21	0.24
Intermediate secondary school	0.40	0.52	0.40	0.40	0.49
Higher secondary school	0.39	0.22	0.41	0.40	0.27
No further education	0.02	0.01	0.01	0.01	0.01
Apprenticeship	0.73	0.90	0.73	0.75	0.85
University degree	0.25	0.09	0.26	0.24	0.14
Personal Characteristics					
Age	38.5	39.1	38.8	38.4	39.6
Married	0.53	0.61	0.56	0.56	0.61
Children	0.30	0.33	0.29	0.30	0.33
German nationality	0.96	0.95	0.95	0.95	0.95
Living in West-Germany	0.72	0.79	0.79	0.79	0.78
Living in East-Germany	0.28	0.21	0.21	0.21	0.22
Job characteristics					
White-collar worker	0.62	0.83	0.71	0.74	0.71
Blue-collar worker	0.32	0.13	0.16	0.08	0.18
Civil servant	0.06	0.03	0.12	0.05	0.06
Temporary work contract	0.09	0.07	0.09	0.07	0.08
Fulltime employed	0.79	0.56	0.68	0.66	0.61
Hours worked per week	35.8	30.4	33.3	33.5	31.5
Tenure	8.55	8.74	10.20	8.48	9.43
Firm size (1-10)	0.16	0.33	0.15	0.38	0.27
Firm size (11-49)	0.23	0.29	0.24	0.22	0.27
Firm size (> 50)	0.61	0.38	0.61	0.39	0.47
n	5,631	37,152	20,605	18,501	44,887

Source: Micro Census 2003 and 2005, own calculations.

Appendix 3-3: Summary statistics for subgroups (II)

	Employed women	Self-employed women	Young women (20 - 39)	Older women (40 - 55)
Net wage				
Wage (300-700)	0.21	0.30	0.20	0.23
Wage (700-1100)	0.26	0.16	0.27	0.24
Wage (1100-1700)	0.44	0.28	0.45	0.42
Wage (1700-2900)	0.07	0.12	0.06	0.09
Wage (2900-4000)	0.02	0.08	0.01	0.02
Wage (>4000)	0.01	0.06	0.00	0.01
Weight				
Underweight	0.04	0.04	0.06	0.02
Healthy weight	0.67	0.71	0.73	0.61
Overweight	0.21	0.18	0.16	0.26
Obese	0.08	0.07	0.05	0.10
Education				
Lower secondary school	0.23	0.20	0.16	0.30
Intermediate secondary school	0.47	0.37	0.50	0.45
Higher secondary school	0.30	0.43	0.34	0.26
No further education	0.01	0.01	0.01	0.01
Apprenticeship	0.83	0.68	0.83	0.82
University degree	0.16	0.30	0.16	0.16
Personal Characteristics				
Age	39.0	42.2	31.2	46.8
Married	0.59	0.70	0.45	0.73
Children	0.32	0.39	0.41	0.22
German nationality	0.95	0.94	0.94	0.97
Living in West-Germany	0.78	0.81	0.80	0.76
Living in East-Germany	0.22	0.19	0.20	0.24
Job characteristics				
White-collar worker	0.78	0	0.80	0.75
Blue-collar worker	0.16	0	0.14	0.18
Civil servant	0.06	0	0.05	0.07
Temporary work contract	0.08	0	0.11	0.05
Fulltime employed	0.62	0.68	0.67	0.56
Hours worked per week	31.8	36.5	32.7	30.9
Tenure	9.20	8.29	6.03	12.42
Firm size (1-10)	0.25	0.95	0.26	0.25
Firm size (11-49)	0.27	0.04	0.26	0.28
Firm size (> 50)	0.48	0.01	0.48	0.47
n	63,388	5,172	30,670	32,718

Source: Micro Census 2003 and 2005, own calculations.

Chapter 4:

Maternal Labor Supply and Childhood Overweight: The Role of Birth Order and Sibling Relationships

Acknowledgment: This project was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft) as part of the research initiative 'Flexibility in Heterogeneous Labour Markets'. I am especially grateful to Patrick Puhani and Falko Tabbert for supportive suggestions and advice. I also thank Philip Yang and Timo Trimborn for helpful comments. All remaining errors are my own.

4.1 Introduction to Chapter 4

In most industrialized countries the share of overweight people has been rising over the last decades. While adulthood obesity is recognized as one of the most severe health problems with negative economic consequences, the interest in childhood obesity is relatively new. This is probably due to the fact that childhood obesity has been rising dramatically over the past years. Prevalence of overweight schoolchildren aged between 5 and 6 years has increased by 45 percent between 1982 and 1997 in the state of Bavaria (Kalies et al. (2002)).[43] Compared to the U.S. the share of overweight children is somewhat lower in Germany: while the prevalence of obese children aged between 6 and 11 was 19 percent in the U.S. in 2003 (National Center for Health Statistics, 2007) it was around 7 percent during that time in Germany.

As shown by Serdula et al. (1993) overweight children are more likely to become overweight adults than children of healthy weight. Therefore, one has to address the problem of childhood obesity and identify potential *causes* and *consequences*. Addressing the consequences, it is found that overweight children are more likely to be teased in school, which might have an impact on the development of their self-esteem, their educational achievements, and their social interaction with other children or classmates. Besides psychological consequences overweight children are facing severe physical side effects. Many chronic illnesses, once common among elderly persons, now appear among children. Examples for these illnesses are arthritis, orthopedic problems, or Type 2 diabetes mellitus, which was formally known as 'Adult-Onset Diabetes'. In later life, additionally to physical effects, obese women have a lower probability to be married than women of healthy weight (Averett and Korenman (1996)). Economic consequences of adulthood obesity are investigated by Cawley (2004) who finds a negative impact of obesity on wages for women.

[43] There is no longitudinal dataset on childhood obesity available for Germany. Studies only refer to regional samples, for example schoolchildren in Bavaria.

The causes for increasing rates of childhood obesity cannot be determined as easily, although the physiological mechanism seems clear: a child becomes overweight if the daily calorie intake exceeds the amount of calories burned. This explanation provides two channels to become overweight: 1) they eat more calories (food channel) or 2) they burn fewer calories (activity channel) - or most likely a mixture of both. Possible reasons for the first channel are the availability of high-fat and high-calorie food in schools and in kindergartens or general trends that people eat out more often and spend less time cooking fresh and healthy food at home. The second channel might be explained by the fact that children tend to spend more time on playing computer games or watching television instead of playing outside. Moreover, due to a continuously improving transportation infrastructure getting to school has become easier for children living in suburban or rural areas. While 50 years ago, children had to walk to school or go by bicycle for several miles, today many children either take a bus or they are dropped off at school directly by their parents, which decreases their level of activity and thus their calories burned.

Furthermore, changes in parental behavior have to be taken into account, as they influence a child's overweight status through food and activity channels - at least in the early years of a child's development. In this context, the most significant change in parental behavior in the last decades seems to be the increasing share of employed women. Due to less supervision, it is easier for children to eat unhealthy snacks or to play computer games during the afternoon. Thus, this might be a potential reason for the rapidly increasing share of overweight children over the last decades. Furthermore, if mothers are working longer hours, they spend less time cooking and rely more often on prepared food or fast food (Cawley and Liu (2007)). Accordingly, studies on the effect of maternal employment find a positive correlation between hours worked by the mother and the probability of the child being overweight for the U.S. (Anderson et al. (2003), Cawley and Liu (2007), Ruhm (2004)). However, fewer studies focus on countries different than the U.S.

This study adds to the literature in two ways: First, in contrast to most large datasets data used in this study contain detailed information on siblings (birth order, number of siblings and age differences between siblings). Therefore, I am able to estimate whether the correlation between maternal employment and overweight children varies between children of different birth ranks or between children with large age differences to their siblings. Studies on the impact of birth order on personality traits find that there are significant differences in personalities between children of different birth ranks. Sulloway (1995, 2001) shows that firstborns are more responsible and act in a way to meet their parents' expectations, whereas lastborns are found to be more rebellious. Because of these differences in personality traits, children might react differently to a working mother in terms of what they eat, how much they eat and how they spend their afternoon. Anecdotal evidence supports this theory by confirming that firstborns differ from laterborns when it comes to responsibility, grown-up behavior and reliability. Second, this paper uses German data. Germany is an interesting case in this context, since in contrast to most countries there is no full-day school available for most pupils. Therefore, children return from school in the early afternoon, which makes them spend more time without supervision if the mother works fulltime.

The remainder of this paper is structured as follows. Section 4.2 gives an overview of the economic literature on overweight, the association between overweight and birth rank, and the relationship between maternal employment and childhood obesity. Section 4.3 introduces the dataset and motivates the hypotheses on birth order and the relationship between maternal employment and overweight children. In section 4.4 results are presented. I find a strong correlation between maternal employment and overweight children. Moreover, age differences between siblings and birth order are important factors to determine whether a child is overweight if the mother is working. Lastborns and children with large age differences to their siblings have a much higher probability to be overweight with increasing maternal employment than firstborns or children with siblings of similar age. Section 4.5 concludes the paper.

4.2 Previous research

When analyzing causes and economic consequences of overweight, most studies focus on adults.[44] Recent literature also focuses on children: For the U.S. and the U.K., Case et al. (2002), Currie and Stabile (2003) and Currie et al. (2007) find a strong and positive relationship between socioeconomic status of the parents and good child-health. One health-related outcome that is of particular interest is childhood obesity, since the share of overweight children is increasing dramatically and overweight children are likely to be the next generation of overweight adults. Guo et al. (2000) analyze consequences of childhood obesity; they find that being overweight as a child does not only have health-related but also psychological consequences. In their study, Guo et al. (2000) ask children which attributes they associate with overweight children. The answers were characteristics such as ‘lazy, dirty, stupid, ugly, cheats and lies’. Overweight children have to cope with these prejudices; therefore, it is harder for them to find friends. Out of boredom and due to reduced social activities they tend to eat even more which might lead to a vicious circle.[45] Moreover, as stated by Guo et al. (2000), overweight children have a lower ability to exercise and therefore less motivation to do so. This again leads to even more weight due to fewer calories burned by doing sports.

[44] Studies analyzing the determinants of obesity (e.g. Sobal and Stunkard (1989), Zhang and Wang (2004)) find, that a higher body weight is associated with a lower socioeconomic status; whereas Zhang and Wang (2004) show that this trend has decreased over the last 30 years. For the U.S., Cawley and Danziger (2004) find that being overweight or obese is negatively correlated with employment, hours worked and earnings for current or former welfare recipients. Nevertheless, most interest is paid to effects of overweight on wages. Most studies find a negative relationship between overweight and wages for women (e.g. Register and Williams (1990), Averett and Korenman (1996), Cawley (2004)).

[45] Physical and psychological consequences of childhood overweight are summarized in Loke (2002).

Anderson and Butcher (2006) study causes and consequences of childhood obesity and find that changes in the food market, in schools and child care settings, and in the role of parents (especially mothers who tend to work more than in the past) play important roles in explaining increasing rates of childhood obesity. Studies for the U.S. find a significant positive correlation between maternal employment and children’s weight problems (Anderson et al. (2003), Fertig et al. (2006)), especially when concentrating on mothers with a higher socioeconomic status or education. Cawley and Liu (2007) and Fertig et al. (2006) use time allocation data to show through which mechanism maternal employment and overweight children are related. Fertig et al. (2006) identify supervision and nutrition as the main channels while Cawley and Liu (2007) show that employed mothers spend less time cooking, eating, and playing with their children and attribute this to the correlation between maternal employment and overweight children. Anderson et al. (2003) estimate the causal effect of mothers' employment on children’s overweight. By taking state child care regulations, wages of child care workers, welfare benefit levels, the status of welfare reform in the state and the annual unemployment rate in the state as instruments for maternal employment, the effect on overweight status of the child remains comparable to the probit results, but insignificant due to larger standard errors. They find that both, fixed-effects and instrumental variable results do not differ too much from probit results, which might indicate that problems of unobserved heterogeneity or endogeneity were only a minor problem in their specification.

Studies for countries different than the U.S. come to similar findings: Takahashi et al. (1999) find a positive correlation between maternal labor supply and overweight children in Japan, while Chia (2008) finds this relationship for Canada. Moreover, she identifies channels explaining this relationship by showing that weekly hours worked by the mother are associated with an increase in probability that the child watches three or more hours of television per day. During the course of my research, I did not find any German study relating overweight of children to their mothers' working activity. One might even expect larger effects in countries like Germany, where children return from school in the early afternoon or at lunchtime. In case the mother works fulltime, children often have to prepare their own lunch or rely on snacks. On the other hand, the availability of fast food and unhealthy packaged food is probably higher in the U.S. Moreover, if a child stays at home in the afternoon instead of being in school, there is more time for activities, such as doing sports. Therefore, the association between maternal labor supply and overweight children could also be weaker in Germany than in the U.S. In the end, it remains an empirical question for which country stronger effects can be found.

Another new aspect of this paper is that it takes detailed sibling information (birth order and age difference between siblings) as explanatory variables. These sibling relationships seem to be an important factor in explaining overweight, which has not been documented in the strand of literature investigating the effect of maternal employment on overweight children.[46] Moreover, this paper analyzes whether the correlation between mother's working activities and childhood overweight varies by birth order or by age differences between children. One reason for this hypothesis is that birth order has an influence on personality (Sulloway (2001)) which leads to different behavioral patterns that might have an impact on overweight.[47] In context of maternal employment and overweight children, I find that birth order and age differences play an important role: the correlation between the mother's working activities and overweight children is strongest for last-born children and children with large age differences to their siblings.

[46] There exists small literature on the correlation between birth order or family size and childhood overweight. Ravelli and Belmont (1979) find a negative correlation between number of siblings and the risk of being obese for Dutch males. Konziel and Kolodziej (2001) show that there is a negative relationship between overweight and birth order for girls in three sibling families, while older literature (Howell (1948), Zonta et al. (1975)) does not find a correlation between birth order and childhood overweight.
[47] Other studies find a negative relationship between birth order and educational attainment (Kantarevic and Mechoulan (2006), Booth and Kee (2009)).

4.3 Data and research design

This study uses German Micro Census data, a large individual-level dataset, which consists of a one-percent sample of the entire German population (the scientific community receives a 70 percent sample of that one percent). Information on weight and height are available for the years 1999, 2003 and 2005. I concentrate on children older than three years of age, since it is not uncommon for German mothers to stay at home with their children until they reach the age of three. Maternity leave regulations in Germany are very generous for mothers: they can stay at home for up to three years and then return to their previous employer. The Micro Census dataset of children aged between 3 and 14 years consists of more than 50,000 observations for the pooled sample.

The dependent variable is an indicator variable whether the child is overweight (including obesity). Following the literature, I do not use BMI as a measure of children's overweight, but whether the BMI is above the 90th percentile of the BMI distribution of an age- and sex-specific reference population surveyed in 17 regional studies between 1985 and 1998.[48] Explanatory variables are: age and age squared of mother and child, gender dummy, nationality, state of residence, availability of 'full-day schools' in the state, dummies for mother's weight categories, mother's schooling and education, household income, marital status, number of siblings, birth order, and age differences between siblings.

Besides these control variables, the variable of interest is whether the employment status of the mother is significantly related to the probability that her child is overweight. In this study, different measures of employment status are used: dummies for working fulltime (at least 31 hours per week) and for part-time employment (working between 20 and 30 hours per week), mother's weekly working hours and dummies for working hours (in sets of 10). In order to estimate a causal effect of maternal employment on childhood obesity, few papers use instrumental variable regression. The reason why maternal employment could be

[48] For Germany, these cut-off values are compiled by Kromeyer-Hauschild et al. (2001).

endogenous in an equation estimating a child's probability to be overweight is that there might be unobserved factors influencing a child's overweight status that are correlated with mother's working activities. Most common instruments for maternal employment are local child care regulations or local economic conditions such as local unemployment rate, percentage of the local labor force that is female, percentage of labor force employed in services, the status of welfare reform in the state, child care regulations, wages of child care worker (Anderson et al. (2003), Baum (2003), James-Burdumy (2005)). Nevertheless, the Mirco Census dataset does not contain local information (on district or community level); the smallest unit of an observation is on state level. Using state unemployment rate or percentage of female workers in the state as instrument for maternal employment does not seem very convincing since there is enough variation over the 16 German states. Therefore, I rely on OLS results, keeping in mind that these results might be biased if, for example, working mothers differ form mothers who do not work in a way that is related to their children's weight (and cannot be controlled for).[49] As mentioned in Anderson et al. (2003), one could think of mothers who work more hours to be generally less attentive to their children's health, irrespective of their work effort. In this case, OLS results would be biased. But since probit and IV results in Anderson et al. (2003) are very similar, the bias using probit (or OLS) is likely to be not too large.

In a first step, I estimate the correlation between family characteristics (such as birth order, number of siblings, age differences between siblings) and overweight for children between 3 and 14 years of age. The second step is to analyse whether there exists an association between maternal labor supply and childhood overweight. Addressing this question, I focus on three points: 1) Is there a difference between single mothers and families? 2) Are there different effects for children of different birth ranks? 3) Do age differences between siblings matter? Concentrating on the first point, I divide the dataset into single

[49] OLS results are very similar to probit results. In Appendix 4.1 both results are contrasted.

mothers and families. Here, I expect the correlation between maternal employment and overweight children to be stronger for single mothers; as shown in Cawley and Liu (2007) and Fertig et al. (2006) childhood overweight can be explained by a lack of supervision, bad nutrition and less time spent with children. Since working single mothers probably have less time to spend with their children and less time to cook the correlation is expected to be stronger for them.

Addressing points 2) and 3), I divide the dataset by birth rank and by age differences between siblings, which is the main contribution of this paper. Most studies on the relationship between maternal employment and overweight children do not take detailed family characteristics such as birth order or age differences between siblings into account.[50] Nevertheless, these characteristics strongly correlate with overweight status since birth order affects personality. The correlation between birth order and personality and its potential effect on overweight if the mother is working is discussed in detail in the next sections.

4.3.1 The correlation between birth order and personality

Why should one expect different effects of maternal employment on a child's probability to be overweight for children of different birth orders? As stated in the literature, birth order has an influence on personality. This correlation leads to different behavioral patterns for children of different birth ranks. My hypothesis is that due to these different behavioral patterns, children react differently if their mother spends less time with them due to longer working hours. Some children might easily cope with these circumstances because they learned to spend time on their own or have a sibling of similar age to play with, others might have a problem if they get less attention, which might lead to stress eating or to other

[50] Anderson et al. (2003) and Fertig et al. (2006) control for number of siblings and for being first born, while other birth ranks are not included. Ruhm (2004) and Cawley and Liu (2007) only control for number of siblings and do not include any information on birth order.

behaviors that are correlated with gaining weight such as watching television or playing computer games.

Sulloway (1995, 2001) finds that birth order has an influence on ‘The Big Five’ personality traits, which are extraversion, agreeableness, neuroticism (or emotional instability), openness and conscientiousness. Therefore it is likely to assume that children with different personalities have different behavioral patterns, which lead to different behaviors when staying at home if their mother is working. As stated by Sulloway (1995, p.77), firstborns score higher on ‘conscientiousness’ which means that they are more amenable to their parents' wishes, values, and standards, including many behavioral elements that reflect conformity to parental values. Moreover, they have a stronger identification with parents and authorities in general. In contrast, lastborns score higher on ‘openness’, which stems directly from their lesser identification with parental authority. According to Sulloway (1995, 2001), openness also entails traits like being daring, untraditional, and rebellious. While Sulloway (1995, 2001) focuses on differences between firstborns and laterborns, Feiner et al. (2003) studies differences between only children and firstborns. They come to similar findings for firstborns: they score higher on conscientiousness, extraversion, and agreeableness. Moreover, they find that only children differ significantly from firstborns, which could not be shown by Sulloway (2001). Differences between firstborns and only children might be explained by the fact that in contrast to only children, firstborns have to cope with the fact that they have to share their parents’ attention and behave in a responsible or more grown-up way as soon as their younger sibling arrives, which might lead to personality traits such as ‘conscientiousness’.

4.3.2 Personality and behavioral pattern if the mother is working

Supporting Sulloway's findings, what would children do if their mother is working fulltime and they return from school in the early afternoon? Following Sulloway, one would expect firstborns to behave in a way to meet their parents' expectations, for example by eating what they are supposed to eat, doing their homework etc. In contrast, if we think of rebellious lastborns, one would not expect them to eat their vegetables, but rather to eat unhealthy snacks, watch TV or engage in other unhealthy behaviors that might lead to overweight.[51] Middleborns are characterized as somehow 'in-between'. They did not get the full attention from their parents before birth of another child (as firstborns), yet they are not as spoiled or pampered as lastborns. For this reason, there might be differences between the correlation between maternal employment and childhood overweight for children of different birth ranks. I expect the effect to be stronger for laterborns, since they score lower on conscientiousness than first-born children.

Whether the correlation between maternal employment and overweight children is stronger for only children than for children with siblings is not clear. While Sulloway (1995, 2001) does not find a difference in personality between only children and firstborns, Feiner et al. (2003) find significant differences in the 'Big Five' personality traits. Moreover, there remain many prejudices against only children, for example that they are 'spoiled, selfish, bossy and lonely' (Blake (1981), Chang and Holmberg (2008)). Irrespective of whether these prejudices are found to be true or not, I would expect only children to react differently to a working mother than children with siblings. Compared to firstborns, they do not have to take responsibility for younger siblings, while compared to laterborns, they do not have anyone to take responsibility for. Moreover, they do not have siblings as playmates during the afternoon; and they probably tend to be more by themselves than children with siblings which could lead

[51] Argys et al. (2006) show that lastborns and middleborns generally have a higher probability to engage in risky behaviors such as smoking, drinking, marijuana use, sexual activity, and crime than firstborns or only children.

to more hours of watching TV, playing computer games or increased snacking behavior. On the other hand, only children are raised surrounded by adults, which may lead to more grown-up and responsible behavior. Thus, whether or not only children have a higher probability to be overweight if their mother is working compared to children with siblings remains an empirical question.

Besides birth order, more sibling relationships can be taken into account by focusing on age differences between siblings. Therefore, I divide the dataset into only children, children with siblings of similar age (age difference less than 3 years), and children with much older or much younger sibling (age difference more than 3 years). My hypothesis is that the effect of maternal employment on childhood overweight is stronger for only children or children with much younger or much older siblings (who are therefore grown-up similarly to only children). Due to the lack of playmates of similar age in the household, they might have a higher probability to stay indoors and engage in more inactive activities such as playing computer games or watching television. Moreover, it could be that employed mothers 'spoil' children who have to stay alone at home during the afternoon by giving sweets or unhealthy snacks as compensation. Correspondingly, I expect the correlation between maternal employment and childhood overweight to be less pronounced for children with siblings of similar age.

4.4 Results

Sample means of all variables are shown in Table 4.1. For most variables, there is not much variation between children of different birth ranks. Nevertheless, there are some differences: mothers of only children have a higher probability to work full-time, are less likely to be married and have a lower household income than mothers with more than one child (which is probably related to the fact that most single mothers have only one child). To account for these differences, I restrict the sample to married mothers and to families with two children in later regressions. In a first step, I estimate the correlation between sibling relationships (such as birth order and age differences between siblings) and a child's overweight status. The sample includes children between 3 and 14 years of age.[52] However, information on younger and older siblings refers to the complete household, including siblings of all ages.

The relationship between sibling relationships and the likelihood of a child to be overweight is shown in Table 4.2. Control variables are added stepwise from specification (1) to (5). In the first specification only characteristics of the child are included (gender, age, age squared, nationality, state of residence, year dummies, and the percentage rate of children visiting full-day schools in the state). The second specification additionally includes characteristics of the mother, such as age and age squared, indicator variables of whether she is underweight, overweight or obese and a dummy for being married. The third column adds indicator variables for four levels of educational attainment of the mother (1. low education and no further degree, 2. low education plus apprenticeship, 3. high school degree plus apprenticeship, 4. high school plus university degree). In the fourth column, mother's occupational status is added in four stages (1. not working, 2. in education, apprenticeship or training, 3. self-employed without employees, blue-color workers 4. self-employed with employees, white-collar workers, civil servants), while specification 5 additionally controls for household income. In Table 4.2, as in all other tables, the standard errors are robust, clustered on the mother's identification code, since there might be more than one child per mother. Moreover, all estimates are weighted using children's sampling weight.

The variables of interest are all kinds of sibling relationships or family characteristics. The first panel estimates the relationship between birth order and overweight. Birth order is included in four categories: Only children, firstborns with younger siblings (as reference

[52] All results remain robust against taking different age groups of children (not shown in the paper). Point estimates vary a bit in size; however, the story remains unchanged.

category), middle-born and last-born children. Middleborns build the smallest category, since this implies that there are at least three children in the household. The correlations with overweight are strong and point estimates become only a bit smaller, as more control variables are included. The fact that all point estimates are positive shows that children of all other birth ranks have a higher probability to be overweight than first-born children with younger siblings. Whereas middle-born children have a 1.7 percentage point higher probability to be overweight, only children and lastborns have a 4.1 and 4.7 percentage point higher probability, respectively. Figure 4.1 illustrates this relationship graphically using the raw data without any control variables.

In the second panel, the relationship between age differences between siblings and overweight of the child is estimated. It can be seen clearly throughout the different specifications that age differences between siblings are positively correlated with overweight. Only children are included in these estimations, setting the age difference between siblings to 30 years, which is about one generation. Results indicate that, for example, having a ten year older or younger sibling is associated with a 4.7 percentage point higher probability to be overweight compared to a child with a two year older or younger sibling. Given that 16.5 percent of all children in the sample are overweight, this result is sizeable. Figure 4.2 clearly demonstrates this relationship graphically.

Results of Table 4.2 indicate that sibling relationships, such as birth order and age differences between children, are highly correlated with a child's probability to be overweight and therefore have to be included in a regression estimating the correlation between maternal employment and overweight children. Moreover, there could be differences between children of different birth ranks or with larger or smaller age difference to their siblings when estimating the relationship between mother's labor supply and a child's probability to be overweight. Section 4.4.3 focuses on this point, after estimating the overall correlation between maternal employment and overweight children in Section 4.4.1 and turning to the role of the father in Section 4.4.2.

4.4.1 The relationship between maternal employment and overweight children

The first column of Table 4.3 shows that there is a correlation between a mother's working behavior and the probability of the child to be overweight. While working less than 10 hours per week is associated with a likelihood of being overweight of about 15 percent, nearly 20 percent of the children are overweight if their mother works more than 40 hours per week.[53] It could be the case that the relationship between family characteristics and overweight can be explained by the fact that a mother's working behavior differs with family size. For example, if a mother has only one child, her probability to work more hours might be higher than if she had three children. The causal effect of maternal employment on overweight children (as found in Anderson et al. (2003)) could be one explanation why only children have a higher probability to be overweight, namely because their mothers have a higher probability to work more hours. However, this can only explain part of the relationship between maternal employment and overweight children. Descriptive statistics in Table 4.3 reveal that, for example, first-born children have a lower probability to be overweight in all stages of their mother's employment level whereas only children and lastborns are heavier regardless of their mother's working activities. The positive correlation between maternal employment and overweight children persists regardless of birth order, when we compare the probability to be overweight at each stage of the mother's employment status (e.g. each line in Table 4.3) or for each kind of birth rank (e.g. each column in Table 4.3).

Table 4.4 shows results of a regression using an indicator variable for overweight children as dependent variable and mother's employment status as variable of interest. The six

[53] Mothers who are not working seem to be an exception, as their children tend to be heavier although they stay at home and were able to care for them all day long. Nevertheless, as Table 4.4 reveals, these differences are due to socioeconomic differences between working mothers and mothers staying at home.

columns represent six different specifications where control variables are added stepwise. The first specification shows the raw effect, i.e. no control variables are included. Specifications 2 though 6 increase the set of control variables, first by including personal characteristics of mother and child (age and age squared of mother and child, state, nationality, weight category of mother – underweight, normal weight, overweight or obese, whether the mother is married and the availability of full-day schools in the state). In specification 3, education of the mother is included, while specification 4 additionally controls for the mother's occupational status. Specification 5 adds household income and the last column, specification 6, takes family characteristics (number of siblings and its squares, age difference between siblings and its squares, birth order) into account.

The difference to Table 4.2 is that the variable of interest is maternal employment, measured in different ways. The first panel includes a dummy for fulltime employment (more than 30 hours per week) and one for part-time employment (between 21 and 30 hours per week). The second panel has number of hours worked per week by the mother as regressor, while the third panel contains dummy variables for the number of hours worked per week in sets of ten. It is clearly shown that children are more likely to be overweight if their mother works more hours per week. Point estimates indicate that children of mothers who work fulltime have a 3.2 percentage point higher probability to be overweight compared to children of mothers who are not working or work less than 20 hours per week (first panel, last column of Table 4.4). Nevertheless, there is no difference in childhood overweight for women who work part-time and those who work less. The different specifications reveal that the correlation becomes stronger as more control variables are included. Recalling descriptive statistics in Table 4.3, children of women who do not work have a higher probability to be overweight, contrary to the general trend (that childhood overweight increases with hours worked by the mother). The small raw effect in Table 4.4 (first column) is driven by the fact that there are many overweight children in families where the mother is not working. As soon as I control for education and job position of the mother, the point estimate becomes larger. If I exclude non-working mothers form the regression (not shown in a table), results stay more or less constant throughout all specifications and the effect including all controls remains unchanged (rounded to two decimal digits). This shows that the high probability of having an overweight child for non-working mothers is driven by observed characteristics of the mother (education and job position which can be interpreted as socioeconomic status) and vanishes if all controls are included. Therefore, I keep non-working mothers in the sample and concentrate on the specification with the full set of control variables.

Taking number of hours worked per week as variable of interest, I find a strong relationship between hours worked by the mother and the likelihood that the child is overweight. Point estimates illustrate that 20 more hours of work (e.g. from part-time to fulltime employment) are associated with an increase of the probability to be overweight of 2.4 percentage points. Dummies for working hours (panel 3) also indicate a strong relationship between hours worked and a child’s overweight status. The reference category (working between 21 and 30 hours per week) lies in the middle, while working less is clearly associated with a lower probability to have an overweight child, and working more is significantly positively related to childhood overweight.

Comparing these results to U.S. finding by Anderson et al. (2003, Table 2) indicates that the correlation between maternal employment and overweight children is stronger in Germany.[54] While Anderson et al. (2003) find that mothers who work 10 hours more per week increase a child’s likelihood to be overweight by 0.7 percentage points, results for Germany find an increase by 1.2 percentage points. These differences could partly be explained by differences in the school system. While in Germany only 15.2 percent of all pupils attended a full-day school in 2005 and otherwise return from school in the early

[54] The research design by Anderson et al. (2003) is slightly different. They use ‘average hours per week if working since child’s birth’ as explanatory variable. Moreover, they have some additional control variables and omit some control variable used in my specification. Nevertheless, to get an idea of whether the effect is stronger for the U.S. or Germany, results can be compared.

afternoon without having eaten lunch, nearly all pupils in the U.S. return from school in the late afternoon and generally eat lunch at school. More support for this presumption is given in Appendix 4.1 which shows point estimates for all control variables comparing OLS and probit results.[55] Point estimates for the availability of full-day schools in the state are negatively related to overweight, indicating that full-day schools decrease the likelihood of being overweight. Unfortunately, there is no micro data on school type available in the Micro Census; therefore, I am not able to estimate the correlation between maternal employment and overweight children separately by school type.

4.4.2 The role of a father in the family

The Micro Census dataset includes information on whether children live with both parents or with a single parent, although it cannot be observed whether these are their biological parents.[56] Taking this information, I divide the dataset into married mothers who live with their husband and single mothers. On average, single mothers work more hours per week if employed (29 compared to 24 hours per week), while in both groups about one third is not working. If the mechanism of the effect of maternal employment on overweight children works through the fact that working mothers spend less time cooking, playing and eating with their children and have less time for supervision (as found in Cawley and Liu (2007) and Fertig et al. (2006)), I would expect a stronger correlation between mother's working activities and child's overweight status for single mothers, since they have less time to spend with their children and children do not have a father to spend time with.

Results for married and single mothers are presented in Table 4.5. For married mothers, I distinguish between an estimation using the same set of control variables as above

[55] Since OLS and probit results are very similar, I rely on OLS estimates.
[56] Children living with their father only are excluded from the sample, while single mothers are included to estimate the effect of maternal employment on overweight children.

and a regression including father's controls.[57] As expected, results are much stronger for single mothers, supporting the theory that less time for supervision, cooking and eating might lead to a higher probability that the child is overweight. While for single mothers who work fulltime the probability to have an overweight child increases by 4.4 percentage points, it only increases by 3 percentage points higher for married mothers who live with their husband. When additionally controlling for father's characteristics, the likelihood to have an overweight child is even somewhat lower for fulltime working mothers. This indicates that part of the correlation between maternal employment and overweight children can be explained by the father's characteristics. Nevertheless, controlling for characteristics of the father reduces point estimates somewhat, but results stay significant and the story remains unchanged: there is a negative correlation between maternal employment and the likelihood for children to be overweight. In the following estimations, I will therefore rely on the full sample of married and single mothers and omit controlling for father's characteristics.

4.4.3 The role of birth order and age differences between siblings

In Table 4.6 the dataset of all children is divided by birth order. It shows clearly that the correlation between mother's labor supply and childhood overweight is much stronger for only children or lastborns than it is for firstborns or those in between two siblings. Point estimates indicate that mothers who work fulltime increase the likelihood that their child is overweight by 3.8 and 4.8 percentage points for only children and lastborns, respectively, both being highly significant. Effects for first-born or middle-born children are much smaller, some even insignificant. Explanations for these findings can be found in the literature on the effect of birth order on personality traits by Sulloway (1995, 2001). As stated above, firstborns want to please their parents and meet their expectations, while lastborns can be

[57] The fact that not all fathers answered all questions leads to a slightly smaller sample size for this regression.

described as the rebel of the family. As discussed in Section 4.3, the effect for only children is not clear: while Sulloway (1995, 2001) finds them to be similar to firstborns, Feiner et al. (2003) find significant differences between firstborns and only children, stating that only children have more similarities with lastborns. Results in Table 4.6 support these findings, at least when focusing on the correlation between maternal employment and childhood overweight.

In order to compare effects of birth order without having different family sizes, I also estimate the effect for firstborns and lastborns in 2-child families only. In this case, all mothers have two children, so their working decision should not be influenced by the number of children. As shown in the lower part of Table 4.6, correlations remain basically unchanged. The effect of working fulltime on a child's overweight is three times higher for lastborns and statistically significant at the 1 percent level, while point estimates for firstborns are insignificant. Therefore, family size effects, which might influence a women's working behavior do not seem to play an important role in this case; the effects for lastborns remain much larger, regardless of family size.

But not only birth order is supposed to have an influence on the correlation between maternal employment and childhood overweight. Theoretically, age differences between siblings could also influence a child's behavior if the mother is working in the afternoon. If children are within a similar age range, they always have someone to play with. In contrast, only children and children with much older or much younger siblings might be more on their own, doing less active things like watching TV or playing computer games which might lead to overweight due to a lower activity level or due to increased snacking behavior. Table 4.7 shows the estimates for children with at least one sibling of similar age (age difference less than three years), for children with siblings who are all much older or younger (age difference more than three years) and for only children. As expected, the effect is higher for children not having a sibling of similar age. In fact, this effect is about as high as it is for only children.

While children with siblings of similar age have a 2.4 percentage point increased probability to be overweight if their mother works fulltime, children with much older or younger siblings and only children have a 3.9 and 3.8 percentage points higher probability, respectively. Limiting the sample size to families with two children comes to more or less the same results.

In this context, another interesting aspect is to combine birth order and age difference effects. To this end, I take families with two children and estimate the correlation between maternal employment and overweight status for firstborns and lastborns separately for children with siblings within the same age range and for children no having a sibling of similar age. Results are shown in Table 4.8. Effects of working fulltime on the probability to have an overweight child are large and highly significant for all groups, except for first-born children with siblings of similar age. This group does not seem to react to a working mother, which can be explained by a firstborns' nature to behave in a conscientious and prudential way in combination with the fact that they have a younger sibling as playmate all day long, a sibling to teach or to explain things. Firstborns with siblings of similar age have to grow up faster than other children, take responsibility and are more self-dependent. Because they want (more than children of other birth ranks) to meet their parents' expectations, they behave in a way that leads to a reduced probability to become overweight. In contrast, for all children of other birth ranks or 'birth rank - age difference' combinations, I find a strong and significant correlation between maternal employment and their overweight status.

4.5 Conclusions of Chapter 4

This paper analyses the correlation between maternal employment and overweight children. In contrast to papers trying to estimate a causal effect, this paper focuses on whether the correlation between maternal employment and overweight children is stronger or smaller for a certain group of children, namely children living with a single mother, children of particular birth ranks or children having siblings of similar age. Therefore, the dataset of all children is divided by single or married mothers, by birth rank and by age differences between siblings. Results demonstrate that the correlation is stronger for single mothers, supporting theories by Cawley and Liu (2007) and Fertig et al. (2006), who provide evidence that supervision, time spend with children and nutrition are the main channels to relate mothers' working activities to childhood overweight. Results on the effect of birth order reveal that the correlation between maternal employment and overweight children is much stronger for only children and for lastborns, which can be explained by studies on birth rank and personality, indicating that middleborns and lastborns are more likely to engage in risky behaviors and have a more 'rebellious' personality than firstborns who try to meet everyone's expectations. Lastly, having a sibling of similar age seems to be negatively related to the probability to be overweight if the mother is working. These results indicate that children might spend time differently when having a sibling of similar age. Due to a lack of playmates, only children and children not having a sibling in their age might spend time doing more inactive things such as watching TV or playing computer games. These behaviors lead to a higher probability to become overweight, especially if the mother is working during the afternoon and therefore not able to supervise their children.

In terms of policy implications, I would not conclude that mothers should stay at home and care for their children, but rather interpret results in the following way: If we had full-day schools and full-day kindergartens to supervise children, provide good and healthy food and offer sports activities, the prevalence of overweight children could possibly be reduced.

References for Chapter 4

Anderson, M.P., K.F. Butcher, and P.B. Levine (2003) Maternal Employment and Overweight Children, Journal of Health Economics, Vol. 22(3), 477-504.

Anderson, M.P. and K.F. Butcher (2006) Childhood Obesity: Trends and Potential Causes, The Future of Children, Vol. 16(1), 19-45.

Argys, L.M., D.I. Rees, S. Averett, and B. Witoonchart (2006) Birth Order and Risky Adolescent Behavior, Economic Inquiry, Vol. 44(2), 215-233.

Averett, S. and S. Korenman (1996) The Economic Reality of the Beauty Myth, Journal of Human Resources, Vol. 31(2), 304-330.

Baum, C.L. (2003) Does Early Maternal Employment Harm Child Development? An Analysis of the Potential Benefits of Leave Taking, Journal of Labor Economics, Vol. 21(2), 409-448.

Blake, J. (1981) The Only Child in America: Prejudice versus Performance, Population and Development Review, Vol. 7(1), 43-54.

Booth, A.L. and H.J. Kee (2009) Birth Order Matters: The Effect of Family Size and Birth Order on Educational Attainment, Journal of Population Economics, Vol. 22(2), 367-397.

Case, A., D. Lubotsky, and Ch. Paxton (2002) Economic Status and Health in Childhood: The Origin of the Gradient, American Economic Review, Vol. 92(5), 1308-1334.

Cawley, J. and S. Danziger (2004). 'Morbid Obesity and the Transition From Welfare to Work', Journal of Policy Analysis and Management, Vol. 24(4), 727-743.

Cawley, J. and F. Liu (2007) Maternal Employment and Childhood Obesity: A Search for Mechanisms in Time Use Data, NBER Working Paper No. 13600.

Cawley, J. (2004) The Impact of Obesity on Wages, Journal of Human Resources, Vol. 39(2), 451-474.

Chang, J. and S. Holmberg (2008) The Only Child Myth, ABC News, Retrieved on August 25th, 2008. (http://abcnews.go.com/2020/story?id=3488411&page=1)

Currie, J., M.A. Shields, and S.W. Price (2007) The Child Health/Family Income Gradient: Evidence from England, Journal of Health Economics, Vol. 26(2), 213-232.

Currie, J. and M. Stabile (2003) Socioeconomic Status and Child Health: Why is the Relationship Stronger for older Children?, American Economic Review, Vol. 93(5), 1813-1823.

Feiner, B., Ch. Kolb, S. Peyreder, I. Schaurhofer, R. Weichselbaumer, and E. Zweimüller (2003) Von eigenwilligen Einzelkindern und anpassungsbereiten Erstgeborenen: Vermoderte Klischees oder noch immer aktuell? [Stubborn Only Children and Adaptable Firstborns: Former Prejudices or Still True? – in German], Discussion Paper, Pädagogische Akademie der Diözese Linz.

Fertig, A., G. Glomm, and R. Tchernis (2006) The Connection between Maternal Employment and Childhood Obesity: Inspecting the Mechanisms, CAEPR Working Paper No. 2006-020.

Guo S.S., C. Huang, L.M. Maynaud (2000) Body mass index during childhood, adolescence and young adulthood in relation to adult overweight and adiposity: The Fels Longitudinal Study, International Journal of Obesity and Related Metabolic Disorders, Vol. 24(12), 1628-1635.

James-Burdumy, S. (2005) The Effect of Maternal Labor Force Participation on Child Development, Journal of Labor Economics, Vol. 23(1), 177-211.

Kalies, H., J. Lenz, and R. von Kries (2002) Prevalence of overweight and obesity and trends in body mass index in German pre-school children, 1982 – 1997, International Journal of Obesity, Vol. 26(9), 1211–1217.

Kantarevic, J. and S. Mechoulan (2006) Birth Order, Educational Attainment, and Earnings: An Investigation Using the PSID, Journal of Human Resources, Vol: 41(4), 755-777.

Kromeyer-Hauschild, K., M. Wabitsch, and D. Kunze (2001) Perzentile für den Body Mass Index für Kinder im Alter von 0 bis 18 Jahren [Percentiles of Body Mass Index for Children and Adolescents Aged between 0 and 18 Years], in German, Monatszeitschrift Kinderheilkunde, Vol.149(8), 807-818.

Loke, K.Y. (2002) Consequences of childhood and adolescent obesity, Asia Pacific Journal of Clinical Nutrition, Vol. 11(3), 702-704.

National Center for Health Statistics (2007) http://www.cdc.gov/nchs/products/pubs/pubd/hestats/overweight/overwght_child_03.htm#Table%201.

Register, C.A. and D.R. Williams (1990) Wage Effects of Obesity among Young Workers, Social Science Quarterly, Vol. 71(1), 130-141.

Ruhm, C.J. (2004) Maternal Employment and Adolscent Development, NBER Working Paper No. 10691.

Serdula, M.K., D. Ivery, R.J. Coates, D.S. Freedman, D.F. Williamson, and T. Byers (1993) Do Obese Children Become Obese Adults? A Review of the Literature, Preventive Medicine, Vol. 22(2), 167-177.

Sobal and Stunkard (1989). 'Socio-economic Status and Obesity: A review of the Literature', Psychological Bulletin, Vol. 105(2), 260-275.

Sulloway, F. J. (2001) Birth order, sibling competition, and human behavior, In H. R. Holcomb III (Ed.), Conceptual challenges in evolutionary psychology: Innovative research strategies (pp. 39–83). Dordrecht, the Netherlands: Kluwer Academic.

Sulloway, F. J. (1995) Birth Order and Evolutionary Psychology: A Meta-Analytic Overview, Psychological Inquiry, Vol. 6(1), 75-80.

Zhang, Q. and Y. Wang (2004). 'Trends in the Association between Obesity and Socioeconomic Status in U.S. Adults: 1971 to 2000', Obesity Research, Vol. 12(10), 1622-1632.

Tables and Figures of Chapter 4

Table 4-1: Means for all children and by birth order

	All	Only child	First-born	Middle-born	Last-born
Outcome variable					
Overweight (incl. obesity)	0.17	0.18	0.13	0.16	0.19
Variable of interest					
Mother works full-time	0.22	0.32	0.19	0.14	0.22
Mother works part-time	0.11	0.15	0.09	0.06	0.12
Characteristics of the child					
Female	0.49	0.49	0.49	0.48	0.49
Age	8.82	8.74	9.14	9.38	8.48
Full-days school (in state)	0.11	0.11	0.10	0.10	0.10
Nationality - German	0.96	0.96	0.95	0.95	0.97
Schleswig-Holstein	0.03	0.03	0.03	0.04	0.03
Hamburg	0.01	0.01	0.01	0.01	0.01
Lower Saxony	0.09	0.08	0.09	0.09	0.09
Bremen	0.01	0.01	0.00	0.00	0.00
Hesse	0.08	0.07	0.08	0.07	0.08
Rhineland-Palatinate	0.05	0.05	0.05	0.05	0.05
North Rhine-Westphalia	0.18	0.18	0.19	0.19	0.18
Baden-Wurttemberg	0.15	0.11	0.17	0.19	0.15
Bavaria	0.15	0.13	0.17	0.15	0.16
Saarland	0.01	0.02	0.01	0.01	0.01
Berlin	0.04	0.05	0.03	0.03	0.03
Brandenburg	0.04	0.05	0.03	0.03	0.04
Mecklenburg-West Pomerania	0.02	0.02	0.02	0.02	0.02
Saxony	0.07	0.09	0.06	0.06	0.07
Saxony-Anhalt	0.04	0.05	0.03	0.03	0.04
Thuringia	0.03	0.04	0.02	0.02	0.03
Characteristics of the mother					
Mother is underweight	0.04	0.05	0.04	0.04	0.04
Mother is normal weight	0.67	0.69	0.69	0.62	0.67
Mother is overweight	0.20	0.19	0.20	0.22	0.21
Mother is obese	0.08	0.08	0.07	0.11	0.08
Nationality - German	0.94	0.94	0.93	0.94	0.95
Age	37.3	37.0	35.5	37.7	38.7
Married	0.88	0.76	0.92	0.92	0.91
Education of the mother					
Education - very low	0.27	0.26	0.26	0.31	0.28
Education - low	0.32	0.29	0.35	0.32	0.31
Education - intermediate	0.28	0.31	0.26	0.24	0.28
Education - high	0.13	0.13	0.14	0.13	0.13

Table 4-1: Means for all children and by birth order (continued)

	All	Only child	First born	Middle born	Last born
Job position of the mother					
Not working	0.34	0.25	0.38	0.50	0.32
Blue-collar worker	0.13	0.14	0.11	0.12	0.14
Self-employed	0.05	0.05	0.04	0.05	0.05
White-collar worker	0.45	0.53	0.42	0.31	0.45
Civil servant	0.04	0.03	0.04	0.02	0.04
Household income (in EUR)					
less than 1300	0.10	0.19	0.08	0.07	0.07
1301-1700	0.13	0.16	0.13	0.11	0.11
1701-2300	0.23	0.22	0.25	0.22	0.22
2301-2900	0.25	0.22	0.25	0.25	0.25
2901-4000	0.19	0.13	0.18	0.22	0.21
4001-5000	0.06	0.04	0.06	0.07	0.07
more than 5000	0.05	0.04	0.05	0.07	0.06
Family characteristics					
Number of siblings	1.09	-	1.22	2.53	1.29
Age difference	9.49	-	3.38	2.29	4.01
Only child	0.22	-	-	-	-
First-born children	0.30	-	-	-	-
Middle-born children	0.09	-	-	-	-
Last-born children	0.39	-	-	-	-
n	51,816	11,804	15,347	4,648	20,017

Source: Micro Census 1999, 2003 and 2005, own calculations.

Table 4-2: Correlation between family characteristics and the probability to be overweight: Stepwise including control variables

	1.	2.	3.	4.	5.
Age difference					
Age differences between siblings	0.0110***	0.0110***	0.0095***	0.0096***	0.0096***
	(0.0018)	(0.0018)	(0.0018)	(0.0018)	(0.0018)
Age differences between siblings (squared)	-0.0003***	-0.0003***	-0.0003***	-0.0003***	-0.0003***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Birth order					
Only child	0.045***	0.046***	0.042***	0.043***	0.041***
	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
Middle-born children	0.024***	0.021**	0.017**	0.016*	0.017**
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Last-born children	0.049***	0.052***	0.047***	0.047***	0.047***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
n	51,816	51,816	51,816	51,816	51,816
Child’s characteristics	x	x	x	x	x
Mother’s characteristics		x	x	x	x
Education of mother			x	x	x
Mother’s occupational status				x	x
Household income					x

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. The estimates stepwise increase the set of control variables. In the first column, characteristics of the child (gender, age, age squared, nationality, state of residence, year dummies, and the percentage rate of children visiting full-day schools in the state) are included. The second column controls for age and age squared of the mother and whether the mother is underweight, overweight or obese. The third column adds the educational level of the mother in four levels. In the fourth column, mother’s occupational status is added (white- or blue-color worker, civil servant, in education / apprenticeship / training), while specification 5 additionally controls for household income.
Source: Micro Census 1999, 2003 and 2005, own calculations.

Table 4-3: Percent overweight children by mother's employment status and by birth order

	All	Only child	First-born	Middle-born	Last-born
All	16.5	18.0	13.2	15.5	18.5
Mother is not working	17.7	21.1	14.4	17.7	19.3
Mother works up to 10h per week	14.9	14.7	11.6	12.0	18.4
Mother works 11-20h per week	14.7	17.0	12.4	12.2	15.5
Mother works 21-30h per week	15.6	15.4	12.8	15.9	17.4
Mother works 31-40h per week	17.5	17.8	12.1	14.8	21.4
Mother works more than 40h per week	19.6	23.0	18.8	15.4	18.5
n	51,816	11,804	15,347	4,648	20,017

Source: Micro Census 1999, 2003 and 2005, own calculations.

Table 4-4: Correlation between maternal employment and overweight children: Stepwise including control variables

	1.	2.	3.	4.	5.	6.
Mother works full-time	0.015**	0.018***	0.025***	0.032***	0.034***	0.032***
(>30h per week)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
Mother works part-time	-0.009	0.013	0.007	0.014	0.015*	0.012
(21 – 30h per week)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
No. of hours	0.0000	0.0003	0.0005***	0.0012***	0.0013***	0.0012***
mother is working	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0003)
Mother works 0-10h	0.014*	-0.000	-0.007	-0.022**	-0.023**	-0.019*
	(0.008)	(0.008)	(0.008)	(0.010)	(0.010)	(0.010)
Mother works 11-20h	-0.009	-0.010	-0.013	-0.013	-0.014*	-0.012
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Mother works 31-40h	0.019**	0.010	0.012	0.012	0.013	0.014
	(0.009)	(0.009)	(0.008)	(0.009)	(0.009)	(0.009)
Mother works >40h	0.040**	0.041**	0.045**	0.043**	0.044**	0.043**
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Personal characteristics (child and mother)		x	x	x	x	x
Education of mother			x	x	x	x
Mother's occupational status				x	x	x
Household income					x	x
Family characteristics						x
n	51,816	51,816	51,816	51,816	51,816	51,816

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. The first specification shows the raw effect. Specifications 2 though 6 increase the set of control variables, first by including personal characteristics of mother and child (age and age squared of mother and child, state, nationality, weight category of mother – underweight, normal weight, overweight or obese, whether the mother is married, and the availability of full-day schools in the state). In specification 3, education of the mother is included, while specification 4 additionally controls for mother's occupational (blue- or white-collar employee, civil servant, in education / apprenticeship / training). Specification 5 adds household income and the last column, specification 6, takes family characteristics (number of siblings and its squares, age difference between siblings and its squares, birth order) into account.
Source: Micro Census 1999, 2003 and 2005, own calculations.

Table 4-5: Correlation between maternal employment and overweight children: Dataset divided by mother's marital status

	Mother married		**Mother not married**
	Not controlling for father's characteristics	Controlling for father's characteristics	
All families			
Mother works full-time	0.030***	0.022***	0.044**
(>30h per week)	(0.008)	(0.008)	(0.018)
Mother works part-time	0.015*	0.011	0.009
(21 – 30h per week)	(0.009)	(0.009)	(0.019)
No. of hours	0.0011***	0.0009***	0.0017**
mother is working	(0.0003)	(0.0003)	(0.0007)
Mother works 0-10h	-0.020*	-0.014	-0.025
	(0.011)	(0.011)	(0.031)
Mother works 11-20h	-0.015	-0.012	-0.008
	(0.009)	(0.010)	(0.020)
Mother works 31-40h	0.009	0.004	0.031*
	(0.010)	(0.010)	(0.018)
Mother works >40h	0.041**	0.038**	0.043
	(0.020)	(0.020)	(0.038)
n	42,927	40,894	8,889
2-child families			
Mother works full-time	0.028**	0.021*	0.036
(>30h per week)	(0.011)	(0.011)	(0.028)
Mother works part-time	0.021*	0.019	-0.013
(21 – 30h per week)	(0.012)	(0.012)	(0.029)
No. of hours	0.0012***	0.0008**	0.0011
mother is working	(0.0004)	(0.0004)	(0.0011)
Mother works 0-10h	-0.033**	-0.027*	0.020
	(0.014)	(0.015)	(0.050)
Mother works 11-20h	-0.018	-0.017	0.011
	(0.013)	(0.013)	(0.030)
Mother works 31-40h	0.004	-0.001	0.053*
	(0.014)	(0.014)	(0.031)
Mother works >40h	0.002	-0.001	0.022
	(0.026)	(0.027)	(0.057)
n	24,148	23,121	3,662

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. All results include the full set of control variables.
Source: Micro Census 1999, 2003 and 2005, own calculations.

Table 4-6: Correlation between maternal employment and overweight children: By birth order

	Only child	First-born	Middle-born	Last-born
All families				
Mother works full-time	0.038***	0.011	0.030	0.046***
(>31h)	(0.014)	(0.011)	(0.023)	(0.012)
Mother works part-time	-0.001	0.013	0.037	0.020
(>20h & <31h)	(0.015)	(0.014)	(0.033)	(0.013)
No. of hours	0.0015***	0.0010**	0.0013*	0.0013***
mother is working	(0.0005)	(0.0004)	(0.0008)	(0.0004)
Mother works 0-10h	-0.029	-0.026	-0.051	-0.009
	(0.021)	(0.016)	(0.036)	(0.016)
Mother works 11-20h	0.003	-0.006	-0.041	-0.025*
	(0.015)	(0.015)	(0.034)	(0.013)
Mother works 31-40h	0.026*	-0.008	-0.021	0.029**
	(0.015)	(0.015)	(0.036)	(0.015)
Mother works >40h	0.079**	0.060*	-0.012	0.015
	(0.033)	(0.036)	(0.058)	(0.027)
n	11,804	15,347	4,648	20,017
2-child families				
Mother works full-time		0.013		0.044***
(>31h)		(0.012)		(0.013)
Mother works part-time		0.017		0.018
(>20h & <31h)		(0.015)		(0.014)
No. of hours		0.0011**		0.0012***
mother is working		(0.0005)		(0.0005)
Mother works 0-10h		-0.034*		-0.021
		(0.018)		(0.018)
Mother works 11-20h		-0.009		-0.019
		(0.016)		(0.014)
Mother works 31-40h		-0.011		0.028*
		(0.016)		(0.016)
Mother works >40h		0.052		-0.018
		(0.039)		(0.028)
n		12,519		15,291

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. All results include the full set of control variables, except family characteristics.
Source: Micro Census 1999, 2003 and 2005, own calculations.

Table 4-7: Correlation between maternal employment and overweight children: By age differences between siblings

	Age difference<=3	Age difference>3	Only child
All families			
Mother works full-time	0.024**	0.039***	0.038***
(>30h)	(0.011)	(0.013)	(0.014)
Mother works part-time	0.015	0.024	-0.001
(>20h & <30)	(0.012)	(0.016)	(0.015)
No. of hours	0.0009**	0.0014***	0.0015***
mother is working	(0.0004)	(0.0005)	(0.0005)
Mother works 0-10h	-0.015	-0.026	-0.029
	(0.015)	(0.019)	(0.021)
Mother works 11-20h	-0.016	-0.026	0.003
	(0.013)	(0.016)	(0.015)
Mother works 31-40h	0.005	0.013	0.026*
	(0.014)	(0.017)	(0.015)
Mother works >40h	0.035	0.012	0.079**
	(0.029)	(0.030)	(0.033)
n	24,652	15,360	11,804
2-child families			
Mother works full-time	0.023*	0.037**	
(>31h)	(0.013)	(0.015)	
Mother works part-time	0.006	0.030*	
(>20h & <31)	(0.014)	(0.018)	
No. of hours	0.0009**	0.0013***	
mother is working	(0.0005)	(0.0005)	
Mother works 0-10h	-0.016	-0.040*	
	(0.017)	(0.022)	
Mother works 11-20h	-0.002	-0.031*	
	(0.015)	(0.018)	
Mother works 31-40h	0.016	0.005	
	(0.017)	(0.018)	
Mother works >40h	0.028	-0.025	
	(0.034)	(0.034)	
n	16,097	11,713	

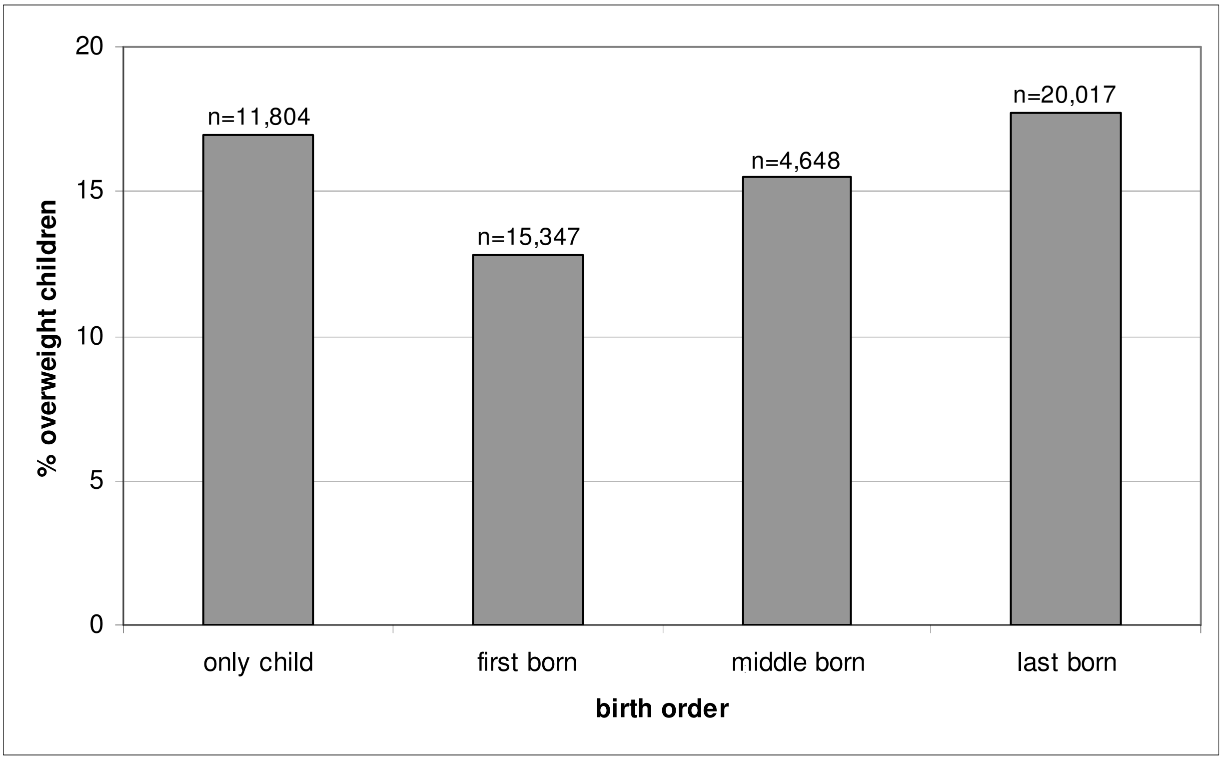
Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. All results include the full set of control variables, except family characteristics.
Source: Micro Census 1999, 2003 and 2005, own calculations.

Table 4-8: Correlation between maternal employment and overweight children: By age differences between siblings and by birth order (families with two children)

	First-born		**Last-born**	
	Age difference<=3	Age difference>3	Age difference<=3	Age difference>3
Mother works full-time	-0.001	0.037*	0.055***	0.039**
(>31h)	(0.016)	(0.021)	(0.018)	(0.019)
Mother works part-time	-0.001	0.049*	0.016	0.022
(>20h & <32)	(0.019)	(0.025)	(0.018)	(0.022)
No. of hours	0.0004	0.0021***	0.0016***	0.0010
mother is working	(0.0006)	(0.0008)	(0.0006)	(0.0007)
Mother works 0-10h	-0.016	-0.066**	-0.020	-0.024
	(0.022)	(0.029)	(0.022)	(0.029)
Mother works 11-20h	0.010	-0.043*	-0.015	-0.027
	(0.020)	(0.026)	(0.018)	(0.023)
Mother works 31-40h	-0.004	-0.024	0.041*	0.021
	(0.020)	(0.027)	(0.022)	(0.023)
Mother works >40h	0.032	0.078	0.029	-0.060
	(0.045)	(0.070)	(0.042)	(0.037)
n	7,774	4,745	8,323	6,968

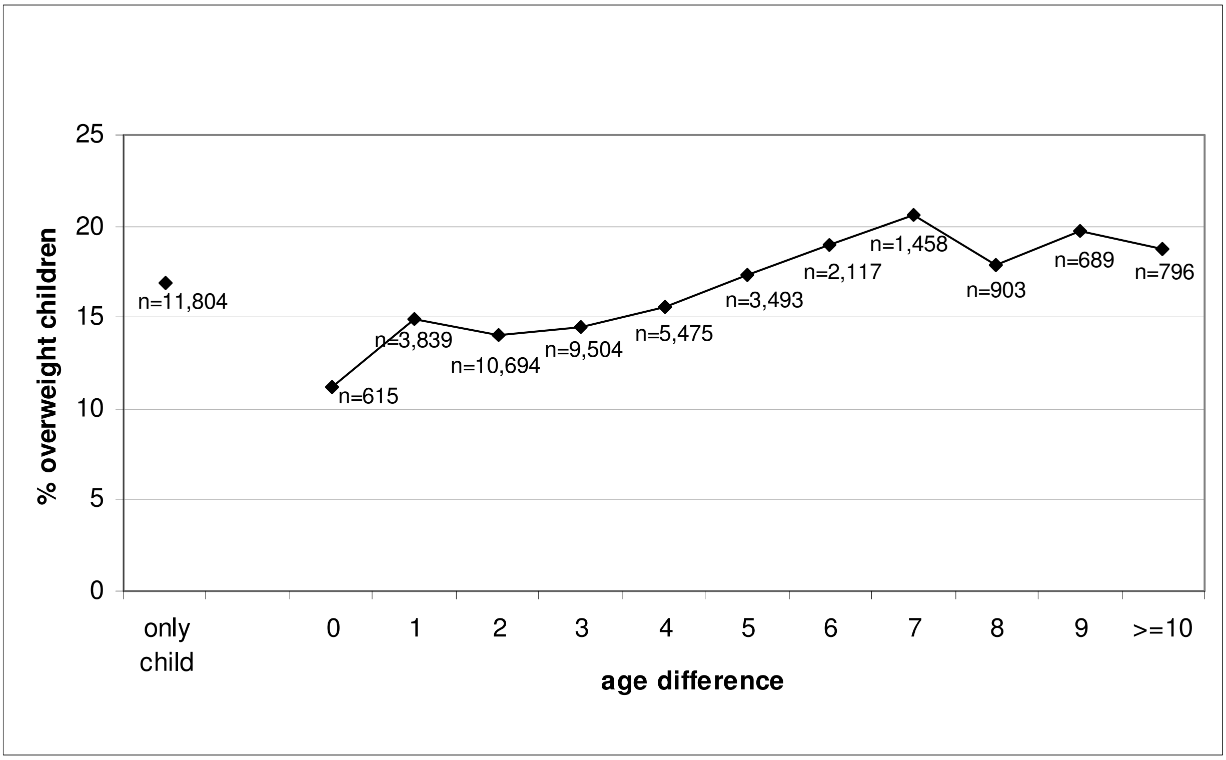
Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively. All results include the full set of control variables, except family characteristics.
Source: Micro Census 1999, 2003 and 2005, own calculations.

Figure 4-1: Percent overweight children by birth order



Source: Micro Census 1999, 2003 and 2005, own calculations.

Figure 4-2: Percent overweight children by age difference between siblings



Source: Micro Census 1999, 2003 and 2005, own calculations.

Appendix to Chapter 4

Appendix 4-1: OLS vs. probit results

	OLS	Probit
Mother works full-time	0.032***	0.034***
	(0.007)	(0.008)
Mother works part-time	0.012	0.014
	(0.008)	(0.009)
Characteristics of the child		
Female	-0.030***	-0.030***
	(0.004)	(0.004)
Age	0.017***	0.018***
	(0.004)	(0.004)
Age (squared)	-0.001***	-0.001***
	(0.000)	(0.000)
Full-day schools (in state)	-0.118*	-0.120*
	(0.066)	(0.069)
Nationality - German	0.001	0.005
	(0.031)	(0.024)
Nationality - EU	0.025	0.022
	(0.050)	(0.045)
Schleswig-Holstein	0.006	0.005
	(0.016)	(0.016)
Hamburg	-0.055**	-0.055***
	(0.022)	(0.020)
Lower Saxony	0.030**	0.030**
	(0.012)	(0.013)
Bremen	-0.026	-0.024
	(0.023)	(0.022)
Hesse	-0.002	-0.003
	(0.010)	(0.010)
Rhineland-Palatinate	-0.011	-0.012
	(0.013)	(0.012)
Baden-Wurttemberg	-0.025**	-0.026**
	(0.010)	(0.010)
Bavaria	-0.029**	-0.029**
	(0.012)	(0.011)
Saarland	0.036	0.033
	(0.023)	(0.023)
Berlin	-0.002	-0.004
	(0.014)	(0.015)
Brandenburg	0.026	0.025
	(0.016)	(0.016)
Mecklenburg-West Pomerania	0.016	0.018
	(0.021)	(0.021)
Saxony	0.016	0.017
	(0.013)	(0.014)
Saxony-Anhalt	0.040***	0.038**
	(0.015)	(0.016)
Thuringia	-0.006	-0.006
	(0.018)	(0.017)

Appendix 4-1: OLS vs. probit results (continued):

	OLS	Probit
Characteristics of the mother		
Mother is underweight	-0.020*	-0.022*
	(0.012)	(0.013)
Mother is overweight	0.065***	0.067***
	(0.007)	(0.007)
Mother is obese	0.109***	0.114***
	(0.010)	(0.011)
Nationality of mother - German	-0.069***	-0.071***
	(0.026)	(0.026)
Nationality of mother - EU	-0.030	-0.021
	(0.036)	(0.028)
Age of mother	0.000	0.000
	(0.004)	(0.004)
Age of mother (squared)	0.000	0.000
	(0.000)	(0.000)
Married	0.001	0.001
	(0.009)	(0.008)
Education of the mother		
Education - very low	0.053***	0.052***
	(0.008)	(0.009)
Education - low	0.011	0.011
	(0.007)	(0.008)
Education - below average	-0.012	-0.014
	(0.008)	(0.009)
Job position of the mother		
Blue-collar worker	-0.003	-0.003
	(0.009)	(0.008)
Self-employed	-0.017	-0.017
	(0.012)	(0.012)
White-collar worker	-0.015**	-0.016**
	(0.007)	(0.007)
Civil servant	-0.004	-0.004
	(0.014)	(0.015)
Household income (in EUR)		
less than 1300	0.024**	0.022**
	(0.011)	(0.011)
1301-1700	0.014	0.013
	(0.009)	(0.009)
1701-2300	-0.001	-0.001
	(0.007)	(0.007)
2901-4000	-0.009	-0.011
	(0.007)	(0.007)
4001-5000	-0.009	-0.011
	(0.011)	(0.012)
more than 5000	0.005	0.002
	(0.012)	(0.013)

Appendix 4-1: OLS vs. probit results (continued):

	OLS	Probit
Family characteristics		
Number of siblings	0.022**	0.020*
	(0.010)	(0.011)
Number of siblings (squared)	-0.003	-0.002
	(0.002)	(0.002)
Age difference	0.013***	0.013***
	(0.003)	(0.003)
Age difference (squared)	-(0.001)***	-(0.001)***
	(0.000)	(0.000)
Only child	0.318**	0.429**
	(0.146)	(0.207)
Middle-born children	0.014	0.016
	(0.009)	(0.010)
Last-born children	0.044***	0.046***
	(0.005)	(0.006)
n	51,816	51,816

Note: *, ** and *** denote significance at the 10%, 5% and 1% level, respectively.
Source: Micro Census 1999, 2003 and 2005, own calculations.