# Economic incentives, employment and health

**Elisabeth Fevang**


Thesis for the Degree of Philosophiae Doctor (PhD)

at the University of Bergen, Norway

Date of defence: 24.05.2019

## Acknowledgements

I am grateful to the Frisch Centre for employing me, and giving me the opportunity to write this thesis.

A number of people has inspired me during the work with this thesis. First, I would like to thank my excellent supervisor and co-author Knut Røed whose guidance and support has been invaluable. I am also thankful to my second supervisor Astrid Grasdal, for constructive criticism and enthusiastic guidance. Besides Knut Røed, I would like to thank Bernt Bratsberg, Simen Markussen and Ines Hardoy for being brilliant co-authors, and for inspiring discussions and encouragements.

Thanks to all colleagues at The Frisch Centre for making my days both pleasant and inspirable. In particular I want to thank Sverre Kittelsen, Oddbjørn Raaum and Ole Røgeberg for their leadership, Simen Gaure for help with estimations, Tao Zhang, Gry Nystrøm and Jørg Gjestvang for their organizational help, Karen Evelyn Hauge and Marte Eline Ulvestad for general support.

I am also thankful to the academic and administrative staff at the Department of Economics at the University of Bergen for their kindness and precious support. In particular, I want to thank Otto Lillebø and Maja Weemes Grøtting for help whenever asked.

I want to thank my family and friends for reminding me of the amazing life outside the office. Finally, a special thank goes to Pål, for your support, understanding and patience, and to our daughters, Julie and Maren, for all your love and inspiration.

## Abstract

The goal of this thesis is to shed new light on the mechanisms behind the high rates of disability-related non-employment in Norway, and to find explanations for the apparent rise of labor market exclusion problems within some demographic groups. The thesis focuses on the role of economic incentives, for individuals as well as firms, and on possible trends in the competitive environment that may (or may not) have changed the health- and productivity requirements in the labor market.

In the first paper we study whether and how recipients of temporary disability insurance (TDI) respond to economic incentives. In order to identify causal effects, we make use of a reform of the TDI system in Norway which was implemented in January 2002. The reform involved a new principle for calculation of the benefits, which changed from being based on the entire income history to being based on income in the last year (or the last three years) prior to disablement. In addition, the minimum level of benefit was raised while the maximum level of child allowance was reduced. This result in changing benefit level, where the benefit level increased for some individuals and reduced for others. We find that the benefit level has a causal impact of the outcome and duration of TDI. According to the point estimates, a 10% cut in the benefit level would induce a 3.3% increase in the transition rate to employment, 2.5% increase in the transition rate to permanent disability, and a 3.9% increase in the transition rate to unemployment. The results are in line with previous findings indicating that there is a significant labor supply potential among temporary disabled people, which can be realized by financial incentives. However, it does not necessarily follow that cutting benefit level is the desired policy from a welfare perspective. Many of the recipients suffer from severe physical and mental illnesses and generous benefits protect them, as well as their dependents, from poverty. Additionally, it provides claimants having the capacity to return to work, with more time to find a suitable and viable job match.

The second paper studies the effect of firm incentives on sickness absence behavior. In most of the industrialized countries (including Norway) the employers are responsible for the costs during an initial period of sickness absence spell, after which the public insurance system covers the costs. Hence, the employers have incentives to prevent short-term absences. But when absence spells stretch beyond the co-payment period, employers may not put much effort in facilitating a quick return to work, since return to work potentially involves new absence spells where the employers are again financially responsible. We examine the impacts of employers' incentives by

exploiting a reform in the Norwegian sick leave insurance scheme. The reform was implemented in 2002 where employers' pay liability was removed for pregnancy-related illnesses. The intention with this reform was to make it more attractive for employers to hire young women. Our findings show that firm incentives actually affect sickness absence behavior by raising short-term absenteeism significantly, while the duration of long-term spells declined. According to the point-estimates the reform increases the probability of starting a period of sickness absence by 10 %, but the probability of ending a period of sickness declines by 12 % for spells exceeding the earlier copayment period. We also find some evidence indicating that the reform actually affect the job-opportunities for young women positively. By following individuals after graduation at school, the reform raised the employment propensity one year after graduation by around 1.5 percentage points for young women in general and by 3.0 percentage points for those who were pregnant at the time of graduation. This implies that there is a trade-off between incentives for sick-leave prevention and incentives for employing workers with high expected absenteeism.

The third paper investigates how exogenous changes in employment opportunities influence take up of disability insurance. Exogenous variations in employment opportunities are measured by variation in firms' economic performance – including profitability, downsizing and firm closure – and fluctuations in local industry-specific labor market conditions. The data we use is Norwegian employer-employee registers together with firms' audited accounts and information from the bankruptcy courts. With data about bankruptcy we are able to distinguish mass layoffs from organizational restructuring, demergers, and takeovers. The estimation results show that job opportunities have significant impact on take up of disability benefits, particularly for men. Job loss, in terms of bankruptcy, more than doubles the risk of entering permanent disability retirement for men while raising entry by approximately 50% for women. Furthermore, it doubles the risk of nonparticipation for both men and women. We also detect that other indicators, as profitability, downsizing and local labor market tightness affect the probability of claiming disability benefits, as well as the probability of being outside the labor force. Putting altogether, the paper shows that there is a considerable element of substitution between unemployment and disability insurance schemes.

In the fourth paper, I explore how employment propensities and earnings of vulnerable groups have developed relative to the population at large. Vulnerable groups are defined as individuals having either poor health, low cognitive ability or coming from low socioeconomic

classes. My main indicator of poor health is low birth weight, which is observed for both men and women. In addition, I use information about height, Body Mass Index (BMI) and cognitive ability measured at age 18-19 for men entering the military service. Socioeconomic class is defined according to parents' earnings rank during their age 50-54. For men, lower birth weight and underweight at age 18-19 has become a stronger predictor of low earnings and non-employment, while there is quite constant effect of height and obesity. For women, where birth weight is the only health-measure I have, I do not find any evidence of changing impact. The influence of cognitive skills on labor market performance has become less important over time, which is due to decreasing returns to high ability. The most striking finding, however, is that poor social background has become a steadily more important determinant of non-employment and low earnings.

# Contents

# Chapter 1:

Introduction

# 1. Introduction

Compared to other OECD-countries, Norway has high take-up rates of health-related benefits (sickness absence and disability insurance) while the rates on unemployment insurance are low (OECD, 2010). The purpose of this thesis is to shed new light on the mechanisms behind the high rates of disability-related non-employment in Norway, and to find explanations for the apparent rise of labor market exclusion problems within some demographic groups. The thesis focuses on the role of economic incentives, for individuals as well as firms, and on possible trends in the competitive environment that may (or may not) have changed the health- and productivity requirements in the labor market. It consists of four papers. Paper I investigates whether economic incentives affects the duration and outcome of temporary disability insurance spells. Paper II is also about the influence of economic incentives, but in this paper the focus is on how employers (not employees) respond to changes in the system of sickness insurance. Paper III studies how changes in employment opportunities influence take up of disability insurance, and paper IV focuses on whether the employment opportunities among vulnerable groups have changed over time.

All the four paper in my thesis seek to find explanations behind the high share of individuals receiving health-related benefits in Norway and why the fraction is higher today than what it was some decades ago. I focus on three possible explanations. The first explanation is that economic incentives for (potential) employees and/or employers are designed in a way that unintentionally promotes this outcome. The second explanation points to a hidden unemployment problem, where entry into disability insurance is triggered by lack of employment opportunities. Finally, the third explanation focuses on changes in the labor market, with increased emphasis on productivity and efficiency, which might have narrowed the job opportunities for individuals with health problems or low ability.

I start in this introductory chapter of my thesis by presenting an overview of the system of social security institutions in Norway, including a description of the reforms used in paper I and paper II, while section 2 discusses the concepts of "health" and "disability" in the context of empirical research. Section 3 provides separate summarizes of the four papers. In light of the findings of the different papers and related literature, I discuss different explanations of the high

rates of health-related benefits in section 4, while section 5 provides some policy implications based on the previous discussion. Finally, section 6 concludes.

### *1.1 Institutional background*

Norway has one of the most generous system of social security benefits in the world, providing universal coverage for income loss due to both unemployment and short- and long-term disability. For individuals with health problems there are three social insurance programs that provide wage replacement: sickness absence benefits (sick pay), temporary and permanent disability insurance. Unemployed individuals may be covered by unemployment insurance. Both for workers and working age individuals without labor market attachment, health problems may qualify for temporary or permanent disability insurance while unemployment insurance is restricted to workers with a minimum of past working experience.

**Unemployment insurance (UI)**

In order to be entitled to unemployment insurance, individuals have to satisfy a minimum earnings requirement based on labor earnings in the calendar year before job loss (or, if higher, the average of the three previous calendar years). In 2017, this threshold was 140,000 NOK Additionally, working hours must be reduced by at least 50%. The replacement rate of UI is 62.4% of previous earnings up to a threshold.[1] As of January 2003 maximum duration was shortened from three to two years.

**Sick pay**

Being employed for at least four weeks, all workers are eligible for sick pay for absence spells lasting up to 1 year. The replacement rate is 100% up to a ceiling of 6 times the base amount in the Norwegian pension system.[2] Sickness is graded from 20 to 100%. The first 16 days of each spell is paid for by the employer, whereas the social insurance system covers the wage costs from the

[1]Maximum threshold was 6 times the base amount corresponding to 560,000 NOK in 2017 value, which is similar to the maximum threshold of disability insurance. Hence, maximum payout of unemployment insurance is 351,000 NOK, while maximum payout of disability insurance is 377,000 NOK.

[2] Employers may offer compensation for workers earning more than the threshold of 560,000 NOK. Dale-Olsen (2018) shows that approximately 60% of workers above the earnings threshold are offered additional compensation for sickness absence.

17th day. There is no experience rating. In order to be entitled to sick pay the absence must be clearly due to own illness or injury. Absence spells lasting more than three (or eight, depending on agreement) calendar days must be certified by a physician. According to the working Environment Act, absent employees cannot be dismissed on grounds that are related to their sickness. When the period of sickness absence has expired, however, the employers can legally lay off absent workers on grounds directly related to their sickness. The firm has no responsibility for subsequent social insurance payments.

**Temporary disability insurance (TDI)**

Individuals who have not recovered after 12 months on sick leave are entitled temporary disability insurance, given that their work capacity is reduced by at least 50% due to illness, injury or defect. TDI is also offered to people who are not employed at the time of disablement. During TDI enrolment, claimants may be offered – and also required to take part in - medical treatment, training, education, courses or other activities intended to help them (back) into employment. The replacement rate is typically around 66% with a minimum and maximum threshold. Maximum duration of TDI is now three years, but it is possible to apply for extension of the period. TDI corresponds to work assessment allowance which was introduced March 1st 2010. Work assessment allowance replaced medical and vocational rehabilitation benefits (1992-2010) and time-limited disability benefits (2004-2010). Before March 1st 2010, there were different rules for maximum length of the different benefits and it was no clear limit on the overall duration of TDI.[3]

**Permanent disability insurance (PDI)**

If the capacity to work is permanently reduced by at least 50%, people may apply for permanent disability insurance, where the compensation rate is roughly the same as temporary disability insurance. Many individuals granted permanent disability insurance have experienced longer periods on temporary benefits and social assistance, particularly among the youngest claimants. For example, Fevang and Røed (2006) show that as long as ten years prior to the first take-up of permanent disability insurance, 45% of the claimants already received some kind of transfers. For

[3] Maximum duration for rehabilitation benefits was 52 weeks, but claimants could apply for extension of the period. Time-limited disability benefits had a maximum duration of four years.

those who were below age 40 at the time of enrolment, this was the case for as much as 67% of claimants. Permanent disability insurance is graded from 50 to 100%.

**Description of reforms utilized in paper I and paper II**

In order to investigate the role of economic incentives, paper I and paper II utilize two different reforms of the benefit system in Norway.

Calculation of both UI-benefits and health related benefits are based on previous earnings. In 2002, the calculation of TDI benefits was subjected to a reform that changed the replacement rate in different directions for different claimants, depending on several observable and predetermined background characteristics. This reform introduced a new principle for calculating benefits, which went from being based on the entire labor-income history of the individual to being based on labor-income in the last year (or the last three years) prior to disablement. Furthermore, maximum child allowance was reduced and the minimum level of benefits was raised. Examples of persons who would receive more with entry after than before the reform are immigrants with few years of residence in Norway and claimants with very low and unstable past earnings. Examples of individuals with higher benefits before the reform are claimants with a declining income path before disablement and claimants with many children earlier entitled to higher (means-tested) child allowance (see Hardoy *et al*. 2004). This reform is utilized in paper I.

Except from the first 16 days of sickness absence, which are generally covered by the employer, all the social insurance programs are paid for by the social security administration. This implies that employers have strong incentives to prevent short-term absenteeism, but limited incentives to prevent long-term absenteeism. In particular, they may exert little effort to facilitate a smooth return to work for employees already being long-term absent, since the return to work is associated with a potential future cost of new absence-spells. In paper II we exploit a reform (introduced in April 2002) of the copayment period for pregnant employees where firms' pay liability were removed from pregnancy-related absences. This reform involves a change of the firms' incentives to influence absence behavior. Short-term absenteeism become less costly, but for pregnant workers already being sick a return to work is not associated with any risk of potential costs with new absence spells. More specifically, if employers respond to economic incentives, we will expect both an increased probability of starting a period of sickness absence and an increased probability of returning back to work for absent employees.

### *1.2 The concepts of health and disability*

Both the concepts of "health" and "disability" are multidimensional and complex phenomena which, to some extent, depend on the social context and available jobs. They are intrinsically unobserved, and must, in empirical analyses, therefore be replaced by proxy variables either capturing subjective assessments of own health or objective indicators that are considered to be correlated with the underlying health status. The concepts of health and disability are important for every articles in the thesis, particularly paper IV.

**Health**

A widely used definition of "health", provided by the World Health Organization (WHO) explains that health is "a state of complete physical, mental and social wellbeing and not merely the absence of disease or infirmity" (WHO, 1946). The definition is holistic since it includes both the physical, mental and social dimensions and does not only concentrate on the absence of disease. However, the formulation has been criticized for being too ambitious and with this definition almost everyone would be considered as unhealthy (Huber *et al.*, 2011).

When describing health, we often use three different concepts; disease, illness and sickness (see, e.g., Boyd, 2000). Disease refers to a pathological process being diagnosed by a physician or another medical expert. Examples of diseases are cancer, diabetes, multiple sclerosis and mental health disorders such as bipolar disorder and schizophrenia. Illness is a subjective feeling of being unwell. Examples are fatigue, fever, pain, weakness, discomfort, distress, confusion. Disease is often the cause of illness, but illness can also occur without an underlying disease or the disease may be undeclared. Being ill is a personal experience and a given disease may trigger symptoms varying tremendously between individuals and also from time to time for each person. In some cases, a person has a disease without feeling ill. Sickness is then the social role a person with disease or illness takes, for example being absent from work.

Using survey data for the period 1988-2001 containing yearly observations of 3,500 employed or self-employed Swedish individuals, Wikman *et al.* (2005) investigate the prevalence of disease, illness and sickness absence and the overlap between them.[4] The authors find that almost

[4] Register data was used in order to measure sickness absence. A person was considered to have an illness if he/she reported symptoms such as pain, sleeping disorders, fatigue and anxiety. A person had a disease if he/she answer yes to the question: "Have you any long lasting disease, trouble from a handicap, or any other weakness" and being

70% of the respondents report to have some kind of illness, while 40% have a long-term disease, and one in seven was observed with an absence spells lasting for more than two weeks during one year. Despite a high prevalence of illness and disease, 80% consider themselves to be in good or very good health. The overlap between the different concepts was quite low; 8% was observed with both illness, disease and sickness absence while 25% had none of these. The rest combined the three concepts in various ways. The authors also shows that the different concepts of morbidity showed different trends over the period 1988-2001; the fraction reporting poor general health was quite stable while there was an increase both in reported illness and disease. Sickness absence increased somewhat, but only during the latest years. This illustrates that health is a complex phenomenon meaning that change in one aspect does not necessarily trigger change in another aspect.

**Disability**

Disability is also a complex and multidimensional phenomenon. The World Health Organization explains it as follows:

> "Disabilities is an umbrella term, covering impairments, activity limitations, and participation restrictions. An impairment is a problem in body function or structure; an activity limitation is a difficulty encountered by an individual in executing a task or action; while a participation restriction is a problem experienced by an individual in involvement in life situations. Disability is thus not just a health problem. It is a complex phenomenon, reflecting the interaction between features of a person's body and features of the society in which he or she lives". (see https://www.who.int/topics/disabilities/en/)

Disability is historically related to the medical model (see WHO, 2011). The medical model focuses on how a disability diagnosis limits functioning, which may be corrected or diminished with medical interventions. The role of medical professionals is important in order to cure or manage disabilities medically. During the last decades there has been a shift from the medical model to the social model. Within the social model people are disabled by barriers in the society, rather than by their bodies. Examples of barriers are people's attitude towards difference, like

diagnosed according to ICD-9. Self-reported health was measured by the question; "In your opinion, how is your state of health? Is it very good, good, fairly, bad, very bad. Note that the indicator for sickness absence is measured on an annual basis, while the other indicators are measured for a more limited time period.

employers assuming that disabled people can't do certain work tasks. WHO's definition of disability represents a compromise of the medical and social model understanding "disability as a dynamic interaction between health conditions and contextual factors (both personal and environmental)" (WHO, 2011. p. 4).

The prevalence of disability is often examined on the basis of surveys with self-reported assessments of own health. The measured prevalence is sensitive to the number and type of questions included, the range of explicit disabilities, inclusion of certain groups, f. ex. institutionalized individuals (WHO, 2011, chapter 2). In Norway, one way of measuring disability rates is to use the Norwegian labor Force Sample Survey. According to this survey, a high fraction of working-age individuals reports having some kind of disability. The survey from 2017 shows that 17% (15% of men and 19% of women) of the population between 16 and 66 years report having a physical or mental disability.[5] The same survey shows that the employment rate among disabled people was 43% in 2017 compared to 73% in the working-age population in general, and this rate has remained quite stable since 2006 when the question of disability was included in the survey. This illustrates that many disabled individuals manage to work.

**Work capacity**

In order to be entitled to health-related benefits, health problems must influence work capacity. Work capacity is related to the concept of disability, but a person reporting some kind of disability does not necessarily have reduced work capacity. It may depend on the type of work being available for the individuals. For example, a given health limitation (f.ex. hearing impairments) will affect work capacity in some type of jobs, but have no or only minor impact on work capacity in other types of jobs.

[5] The question of disability in the Norwegian Labor Force Survey is formulated like this: "By disability, we mean physical or mental health issues of a lasting character that may cause limitations in your daily life. This can, for instance, mean reduced sight of hearing, difficulties with reading and writing, mobility impairment, heart- or lung issues, cognitive difficulties, mental disorders, or other things. Are you, in your opinion, disabled?" Interesting to note is that "only" 30% of people with disabilities report having poor health (compared to 6% of the general population). See https://www.bufdir.no/en/English_start_page/Disabilities_in_Norway/Statistics_on_disabilities_in_Norway/

### *1.3 Summaries of the papers*

**Paper I "Temporary Disability and Economic Incentives" (Joint with Ines Hardoy and Knut Røed)**

During the last decades, the share of working-age individuals receiving disability insurance has increased substantially in many countries (Autor and Duggan, 2003; Burkhauser and Daly, 2011; Bratsberg *et al.*, 2013). A number of studies have investigated how unemployment insurance claimants are affected by economic incentives (Fredriksson and Holmlund, 2006; Card *et al.*, 2007; Røed *et al.*, 2008), but only a few studies have focused on the impacts of economic incentives on the duration and exit from temporary or permanent disability.

In order to identify the role of economic incentives we utilize a reform of the Norwegian temporary disability insurance (TDI) system in 2002, involving a full-scale overhaul of the TDI benefit scheme. Claimants were affected differently by the reform. For some individuals with certain income paths and personal characteristics, the benefit level was reduced, whereas it increased benefits for others. Using this random assignment like variation of changes in benefits, we estimate causal impacts of the benefit level on the duration and outcome of TDI spells.

We use multivariate (mixed) proportional hazard rate model to quantify the effects of changes in the benefit level on the duration and outcome of TDI spells. A crucial element of our identification strategy is that we use both the two hypothetic replacement levels that would have had under the old (pre-reform) and the new (post-reform) regimes ($b_i^o, b_i^n$), and the actual benefit level ($b_i^a$). The actual benefit level will be $b_i^o$ for entrants in the old regime and $b_i^n$ for entrants in the new regime. In addition, we account for calendar time effects by including dummies for calendar time. Since the reform affected the claimants in different ways this will not create a perfect multicollinearity problem. However, the hypothetical benefit level will capture all of the spurious effects arising from the fact that replacement levels is based on the income history and the actual benefit level is quasi-random assigned, i.e. it only depends on the timing of TDI spells.

Our main finding is that the benefit level affects the duration and outcomes of TDI spells. For fully disabled TDI claimants, the estimates imply that a 10% reduction in the replacement level generate a 3.3% increase in the hazard rate to employment, a 2.5% increase in the hazard rate to permanent disability, and a 3.9% increase in the hazard rate to unemployment. Our results support earlier findings that there is a significant labor supply potential among temporary disabled persons,

and that the realization of this potential to some extent can be realized by means of financial incentives.

The reform was intended to be cost-neutral, but this paper shows that the benefit level actually increased by 14%. Results from simulations imply that the reform, by lowering the exit rates out of TDI, may have been responsible for increasing the stock of TDI claimants by approximately 4-5%. By comparison, the number of TDI claimants rose by as much as 35% from 2001 (the last year before the reform) to 2004, and this increase seems to be a part of an upward trend. Hence, in order to understand the overall rise in TDI, other explanations are called for.

**Paper II "The Sick Pay Trap" (Joint with Simen Markussen and Knut Røed)**

A large body of research has documented that workers respond to economic incentives of the sickness insurance scheme, in the sense that if sickness absence becomes more costly for the workers, they will also tend to have less sickness absence (Henreksson and Persson, 2004; Johansson and Palme, 2005; Ziebarth, 2013; Ziebarth and Karlsson, 2014; D'Amuri, 2011).[6] However, there is little evidence regarding the impact of firm incentives. OECD (2010, p. 133) notes that countries where employers are responsible for a large share of their employees' sick pay costs tend to have lower absence than countries where the public insurance system cover (most of) the bill, and also that absenteeism has dropped significantly in the Netherlands and UK after a shift of financial responsibility towards employers. Yet, to our knowledge, no studies have established a causal relationship between firm incentives and worker absenteeism. The design of firm incentives with respect to sick-leave prevention also involves a potential tradeoff between sick-leave and labor market exclusion: While more extensive pay liability improves incentives for absence prevention for already employed individuals, it may at the same time undermine incentives for hiring persons perceived to have a high risk of absence.

In Norway, employers are financially responsible for sick pay during the first 16 days of a sickness absence spell, while the public insurance system covers the whole bill after this period. This division of financial cost with sick pay, where employers are financially responsible for sick pay during an initial period of a workers' sick leave, while the public insurance system covers the bill after some duration threshold are also typically the case in most other industrialized countries

[6] Individuals responsiveness to the system of sickness absence is further confirmed by more recent studies (see Aaviksoo and Kiivet, 2016; Bøckerman *et al*., 2018; Eliason *et al*, 2018)

(OECD, 2010, Table 5.1). This means that firms have strong financial incentives to prevent short-term absences. But when absence spells stretch beyond the co-payment period, employers may not put much effort in facilitating a quick return to work, since return to work potentially involves new absence spells where the employers are again financially responsible.

By using Norwegian administrative register data, we examine the impacts of employers' incentives by exploiting a reform in the Norwegian sick leave insurance scheme. The reform was implemented in 2002 where employers' pay liability was removed for pregnancy-related illnesses. The intention with this reform was to make it more attractive for employers to hire young women. Markussen *et al.* (2011) show that the increased risk of absenteeism associated with pregnancies is substantial; the hazard rate of entering into a sick-leave spell is raised by a factor of five at the onset of a pregnancy and further to a factor of 15 during the last 2-3 months before delivery. Hence, the reform clearly removed a potentially important disincentive with respect to hiring young women; but at the same time it also enhanced employers' incentives to prevent sick leave among pregnant workers.

On the one hand, the reform made short-term absence – absence spells with durations up to 16 days – less costly for the firm. On the other hand it also made it less risky letting long-term absent pregnant employees return to work, since the employers no longer were responsible for the sick pay costs associated with new absence spells. Hence, the reform offers a neat setting for identifying the impacts of firm incentives.

Our findings show that firm incentives actually affect sickness absence behavior considerably, as the reform raised short-term absenteeism, while the duration of long-term spells declined. We estimate that the reform increased the entry-rate into absenteeism by 10%, but also raised the transition rate back to employment by 12% for spells exceeding 16 days. This suggests that policy makers indeed may have good reasons to focus on improving employer incentives in order to reduce absenteeism.

We also find some evidence indicating that the reform actually affected the job-opportunities for young women positively. By following individuals after graduation at school, we show that the reform raised the employment propensity one year after graduation by around 1.5 percentage points for young women in general and by 3.0 percentage points for those who were pregnant at the time of graduation. This implies that there is a trade-off between incentives for sick-leave prevention and incentives for employing workers with high expected absenteeism.

**Paper III "Job Loss and Disability Insurance" (Joint with Bernt Bratsberg and Knut Røed)**

In welfare states, there seem to be a grey area between unemployment and disability insurance. The motivation behind this study is that the recent rise in claimants of disability insurance does not seem to be driven by deterioration of health conditions, and countries with comprehensive disability insurance programs also tend to have low unemployment rates (OECD, 2010; Røed, 2012)

Since the individual risks of disability and unemployment are highly correlated, the causal effect of job opportunities on disability insurance enrollment is difficult to identify based on observational data alone. Our empirical strategy is to exploit exogenous sources of variation in individual employment opportunities, generated by variation in employers' economic performance – including profitability, downsizing, and firm closure – and fluctuations in local industry-specific labor market tightness, to identify causal impacts. The empirical basis is Norwegian administrative employer-employee registers, merged with firms' audited accounts and information collected from bankruptcy courts. The bankruptcy data make it possible to distinguish genuine mass layoffs from organizational restructuring, demergers, and takeovers.

A broad international literature has analyzed adverse consequences of job displacement (see, e.g., Hamermesh, 1987; Ruhm, 1991; Neal, 1995; Kletzer, 1998; Kuhn, 2002; and Hallock, 2009), including two studies based on Norwegian employer-employee data (Rege *et al.*, 2009; Huttunen *et al.*, 2011).[7] The present paper adds to the literature in several directions. It is, to our knowledge, the first study exploiting data on mass layoffs measured by recorded bankruptcies which allow us to distinguish between different kind of closures; closure due to bankruptcy, closure due to voluntary liquidation, and takeovers (with or without bankruptcy). Based on estimates of the overall number of involuntary job losses in the economy – including those from stable and growing firms – it is also the first study to assess the total impact of job loss on the frequency of disability insurance claims. In addition, we add to the literature by examining more specifically the influences of firms' economic performance and of alternative (local) employment opportunities on workers' probability of entering disability insurance programs. And, finally, we examine the *interaction* between these various measures of employment opportunity to test whether the

[7] Since this paper was published, additional studies have found adverse health consequences of job loss (Browning and Heinesen, 2012; Black *et al*., 2015; Schaller and Stevens, 2015; Michaud *et al*., 2016).

probability that job loss leads to a disability insurance claim declines with local labor market tightness.

We find that exogenous changes in employment opportunities, measured by closure or downsizing of a firm, affect the risk of being dependent of disability insurance and non-participation, and the effects are more pronounced for men. In support of the hypothesis that disability and unemployment are substitutable, we also identify significant interaction effects between job loss and local labor market conditions. The more difficult it is to find a new job, the higher is the probability that displacement leads to disability retirement.

Taken together, the results in this paper points to a considerable element of substitutability between unemployment and disability insurance. Our findings suggest that the process of reallocating redundant workers from old to new employers is far from seamless, and that many displaced workers permanently change status from supporting the welfare state to becoming supported by it. Significant human capital resources are squandered in this process. The finding that loss of employment is among the major causes of disability program entry – whether it stems from genuine health effects or from adverse shocks to the expected value of labor market participation for given health levels – suggests that appropriate solutions to the "disability problem" should address strategies for improving the employment opportunities of potential claimants rather than focus exclusively on income insurance. If job loss and unemployment are among the root causes of the rising disability problem, it is probable that provision of employment opportunities is among its remedies.

**Paper IV "Vulnerable groups and labor market performance. Towards more sorting in the labor market?**

In Norway, as well as in other OECD-countries, policy makers seek to promote a high labor market participation rate and an 'inclusive labor market'. According to the European Commision (2017) "Labor markets are inclusive when everyone of working age can participate in paid work, especially vulnerable and disadvantaged people". At the same time many would argue that the labor market has become tougher because of increased focus on reorganizations and productivity. Empirical studies have detected that mass layoffs and organizational changes raise the probability of permanent exclusion from the labor market, (Rege *et al.*, 2009; Bratsberg *et al.*, 2013) and it may also adversely affect the health of the employees (Ferrie, 2001; Kivimäki *et al.*, 2001).

However, there is no clear evidence that these events have become more prevalent over time, or that these are implemented in a less inclusive fashion. Moreover, there have also been changes that have improved working conditions for many employees, for instance access to better physical aids, less manual work and more flexible work schedules.

The aim of this paper is to investigate how employment propensities and earnings of vulnerable groups have developed relative to the population at large. Vulnerable groups are defined as individuals having either poor health, low cognitive ability or a disadvantaged family background. My main indicator of poor health is low birth weight, which is observed for both men and women. In addition I use information about height, Body Mass Index (BMI) and cognitive ability measured at age 18-19 for men entering the military service. Family background is described in terms of socioeconomic class, which is defined by parents' earnings rank during their age 50-54.

A large literature documents that there has been increasing returns to skills as technological changes have increased the demand for high-skilled workers (Bound and Johnson, 1992; Juhn *et al.*, 1993; Acemoglu and Author, 2011). After 2000 there has been little or no growth in cognitive skill-intensive occupations (Acemoglu and Author, 2011; Beaudry *et al.* 2014), but there has been a strong growth in jobs requiring high social skills (Deming, 2017; Edin *et al.*, 2017). The research suggests that the main explanation behind the findings is that technology to an increasing extent substitute for cognitive skills while social skills are more difficult to replace (Deming, 2017).

Empirical evidence on whether the impact of health and social background has changed over time, is limited. Markussen and Røed (2017) show that being born into the poorest families have become a stronger predictor of non-employment and lower earnings. Comparing twin pairs born in different decades, Bharadwaj *et al.* (2018) does not find any evidence of changing impact of birth weight on earnings and high school completion.

A key finding in my study is that the labor market has become gradually more sorted with respect to socioeconomic status (SES); and being born into a low SES family has become an ever stronger predictor for poor labor market performance as young adult, both in terms of non-employment and low earnings. This is the case for both men and women. Although the social gradient in labor market performance becomes slightly weaker when birth weight, height, BMI and cognitive ability are controlled for, such controls do not alter the conclusion with respect to the rising influence of social background. For men, I also find indications of increased sorting with

respect to health. Although this relationship is not as robust as that for social background, most of the results point in the direction that men with poor health have experienced a steady decline in relative employment prospects. In particular, men with low birth weight and underweight at age 18-19 are to an increasing extent non-employed at prime age. For women, where birth weight is the only available health-measure, I do not find any changing impact on employment.

Consistent with earlier research, I find that the impact of cognitive skills on labor market performance has diminished over time. This finding is primarily driven by decreasing returns to very high ability; the influence of having low – relative to medium – ability has been more stable. Given that my indicators for social background, health, and ability are all noisy, and at the same time highly interrelated, it is clear that I cannot provide a complete decomposition of how these different factors affect adult labor market performance. In particular, it is probable that my health indicators are too crude to fully account for the relationship between social background and health. Another challenge for the identification of health effects is that the correlation between the various health indicator and actual health in adulthood may have changed over time. This is particularly the case for low birth weight, as huge improvements in neonatal health care both has raised the probability of surviving with poor health (implying a stronger relationship between low birth weight and poor adult health) and limited the adverse consequences of low birth weight, given survival (implying a weaker relationship). To assess this challenge, I return to an analysis of the relationship between birth weight and alternative adult health measures after having presented the main results. It essentially fails to identify significant changes in the relationship between birth weight and adult health, although it cannot rule out such changes either.

### *1.4 Possible explanations of the high rate on health-related benefits*

The four papers in the thesis are concerned about three different explanations of why there is a high rate of individuals on health-related benefits. These are economic incentives (including substitution between unemployment insurance and health-related benefits), a hidden unemployment problem and changes in the labor market. In light of our results and related literature I will discuss the different explanations more generally. I will also briefly mention other possible explanations such as changes in health, higher share of older individuals in the working- age population and less family support.

**Economic incentives**

Except for some high-earnings workers in private sector, all employees in Norway are fully insured against income loss due to sickness absence. A large body of research have found that employees respond to economic incentives by being more likely to call in sick when replacement rate is higher. However, although workers receive 100% compensation, workers have incentives to be present at work during period of illness since sickness absence may generate lower subsequent earnings and employment (Hansen, 2000; Ichino and Moretti, 2009; Markussen, 2012).

In case of illness, there are also many non-pecuniary incentives influencing the propensity to be absent from work, f. ex. loyalty to the company and how absenteeism will influence colleagues and customers (see Morken *et al.*, 2012). Furthermore, social norms at the workplace (Godøy and Dale-Olsen, 2018) and attitudes towards possible reasons for sickness absence (Hauge and Ulvestad, 2017) may play an essential role.[8]

The importance of economic incentives on sickness absence are not restricted to employees. Based on findings from paper II we suggest that employers also contribute to the high rate of sickness absence we observe in Norway, since employers are not financially responsible for absence lasting more than 16 days. Furthermore, we show that the system of pay liability may create a sick pay trap where employees may put little effort in facilitating return to work for long-term absent employees since a possible new period of absence will involve financial cost for the firms. It may also be that employees take this system into account when (together with the physician) deciding the length of absenteeism. Workers may evaluate how their health problems influence own work capacity which in turn affects his/her colleagues and possible customers/patients/students. If, for instance, the sick worker can be replaced by a (healthy) substitute worker, it might be that the length of absenteeism will be set to ensure that the worker is (again) fully productive and/or a possible risk of new absence spell is minimized.

If workers have not recovered after one year of sickness absence, they may apply for temporary disability insurance, where the replacement rate is typically lower than sick pay (around 67%). We do find that economic incentives affect the duration and outcome for the claimants; higher replacement rate implies longer duration of the benefit and lower transition rate to

[8] Hauge and Ulvestad (2017) find an association between sickness absence and attitudes, but it was mainly due to differences in short-term leave.

employment, other benefits (permanent disability insurance and unemployment insurance) and non-participation without benefits.

Our research suggests that some labor supply among temporary disabled individuals may be realized by financial means. Note, that the temporary disabled individuals in Norway typically do not have a job.[9] In general, a period out of work is typically associated with substantial subsequent earnings losses as this may influence future career opportunities and wage growth, meaning that the long-term economic loss of periods on benefits are often larger than what is reflected in the replacement level. On the other hand, for individuals being out of work, available jobs may involve lower earnings compared to what they received in previous jobs. And in some cases it may be that claimants do not benefit from working.

Based on information about benefits and potential full-time income after tax, Nordberg (2007) constructs total tax rates for benefit receivers in Norway, which are measured as "the fraction of the employers wage cost not gaining the employee, as fraction of total wage costs". He finds that the great majority receives a substantial higher income when working, but approximately 4% of the benefit claimants face total tax rates above 100%, meaning that they are better off if they continue to be recipients instead of working.[10] Another 24% face total tax rates between 80 and 100%. For some (potential) employees it will also be additional costs associated with working, f.ex. travel costs and daycare expenses, indicating that even more recipients are better off when not working.

Changes in the replacement level do not affect transition to employment only. We do observe that it also affects transition to other benefits as well as non-participation without benefits. Whether a transition out of temporary disability insurance is a transition to employment is influenced by local labor market conditions. That employment opportunities play an essential role is further confirmed by findings in paper III, where we find that the probability of entering a disability insurance programme after job loss is affected by available jobs[11] Hence, when evaluating potential success of different reforms of social insurance it is not obvious whether a

[9] Individuals on full temporary disability insurance typically do not have a job. Some claimants are, however, classified as partly disabled and probably have a part-time job. 14% of the claimants in the dataset used in paper I are partially disabled.

[10] Fevang *et al*. (2005) and Hernæs *et al*. (2016), making similar calculations for benefit claimants in Norway, also find that around 4% of the recipients are economically better if they continue to receive claimants instead of working.

[11] Significance of local labor market conditions on take up of disability insurance is also find in Black *et al*. (2002) and Charles *et al*. (2018)

reduction in take-up of social insurance is accompanied by a corresponding rise in employment rates.

As described earlier, there has not been any general increase in the take up of social security benefits in Norway during the last decades. However, a higher replacement level of health-related benefits in combination with a tightening of the criteria for being eligible for unemployment insurance may explain the rising significance of health-related benefits. Some non-employed individuals, such as youths without working experience, are not entitled to unemployment insurance, meaning that temporary disability insurance may be the only alternative for income support for this group.[12]

Does the combination of high rates on health-related benefits and low rates on unemployment insurance affect labor supply? Some kind of activity aimed at returning to work is required both as recipients of unemployment benefits and temporary disability insurance. Recipients of unemployment insurance are obliged to search for jobs, while individuals on temporary disability insurance must participate in agreed activities – f.ex. courses, education, employment scheme or medical treatment. The group of non-employed individuals is a heterogeneous group, with different needs and health status. For some persons it may be beneficial both for health and future earnings to receive temporary disability insurance while taking part in activities (courses, education, etc) while others will not benefit from it. Actually, being diagnosed by a physician, as all individuals receiving disability insurance are, may imply that both case workers and potential employees, as well as the recipients themselves, focus more on their problems than their actual resources.

Utilizing different practices between Norwegian municipalities regarding the leniency of granting temporary disability insurance instead of unemployment insurance to youths, Schreiner (2017) shows that that being granted temporary disability as opposed to unemployment insurance significantly reduces future earnings and increase subsequent welfare dependency. This probably illustrates that too many youths are granted health-related benefits. Whether this result can be transferred to the whole group of non-employed individuals is not clear. As a consequence of the earnings requirement to be entitled for unemployment insurance, the potential for overuse of temporary disability insurance is particularly high in this group.

[12] Non-employed individuals may also be granted social assistance, but only if there is no other possibilities for income/support (including own savings). Social assistance is means-tested, but the payments are typically lower than the other benefits.

Economic incentives may also imply that sickness absence is used as a substitute for unemployment insurance. For example, in case of downsizing, both employers and employees may have mutual interest in using sickness absence instead of unemployment insurance since the compensation level of sickness absence is substantially higher and the employer is not financially responsible for long-term absence.

**Hidden unemployment problem**

As described in paper III there is a grey area between unemployment and disability and in this paper we show that negative shifts in employment opportunities explain part of the disability insurance dependency in Norway.

Negative shift in labor supply following displacement can be described within the context of job search theory (see section 3 in paper III). Job loss can be viewed as a negative shock to the value of continued labor market participation. There may be some costs associated with searching for a new job, job-specific human capital is lost, a new job may be hard to find and is likely to pay less than the current job. This implies that for some groups, inactivity (with or without disability insurance) may be preferred to searching for new jobs.

Being eligible for disability benefits requires minimum 50% reduced work capacity due to illness or injury. As described in section 1.2, some people work despite having some kind of disability, which probably means that within the group of workers there are many potential claimants of disability insurance. A given disability/health problem is likely to reduce the offers of new jobs and job search is considered to be less valuable. And since work capacity is evaluated relative to available job opportunities a job loss may trigger take-up of disability insurance.

In reality, few workers go directly from job displacement to permanent disability. It depends on age, health, labor market opportunities, etc. The replacement rate may also be essential. In addition, the social security administration try to minimize the effect of job loss by offering courses, educations and other measures in order to help non-employed back to work. However, the implication from job search theory is confirmed by the empirical results in paper III in the sense that more limited employment opportunities increase the likelihood of receiving disability insurance.

**Changes in labor demand and in the competition for jobs**

Another possible explanation of higher utilization of health-related benefits is changes in the demand for labor and in the competitive environment. Recent technological change has raised the demand for non-routine and high-skilled labor (see f.ex. Deming, 2017), while an opening up of a common European labor market has put a pressure on employment opportunities, particularly among low-skilled workers. Competition in several sectors has intensified, with increased emphasis on outsourcing, reorganisations and productivity. The pace of structural change has increased also within the public sector, f.ex all government agencies are required to save a given per cent of their budget each year. A more turbulent labor market may increase the risk of developing health problems, and also make it more difficult to be employed with given health problems. In addition, increased job insecurity and job loss may in itself induce health problems. On the other hand, there have also been changes having improved working conditions for many employees, for instance access to better physical aids, less manual work and more flexible work schedules.

In paper III, we have not focused on plant closure only - we also investigate the effects of downsizing and find that the probability of subsequent disability rise with the downsizing level. This effect can, be explained by increased risk of job loss or reduced probability of being rehired by a new employer, especially if indications of health limitations are present, but may also be attributed to health-changes of retained workers. A large epidemiological literature indicate that downsizing affect health of the retained workers by imposing stress due to increased workload and/or increased perception of job insecurity (see e.g. Ferrie 2001). Additionally, Røed and Fevang (2007) show that downsizing process among Norwegian nurses raised the level of non-employment and sickness absence, even when layoffs are not part of the downsizing strategy. Furthermore, Rege *et al*. (2009) show that downsizing increase the risk of take-up of disability insurance. This effect was attributed both to a decrease in subsequent earnings and employment opportunities, and to an adverse effect on health. Among individuals being granted disability insurance, downsizing increases the risk of being diagnosed with psychological conditions which is likely to be related to distress in connection with the downsizing process.

In general, work seems to be good for health, particularly for workers already having a long-standing illnesses or disability (Waddell, 2004; Waddell and Burton, 2006; OECD, 2008). But it may not always be the case. Using longitudinal survey data from Australia, Canada, Korea,

Switzerland and the UK, OECD (2007) confirm that workers mental health is negatively associated with unemployment. When investigating the revers change (from non-employment to employment) the authors find improvements in mental health for those finding a job with standard working-conditions. On the other hand, for non-employed individuals with mental illness who get a job with non-permanent contract, work may not be beneficial for health. Furthermore, Markussen *et al.* (2012) show that partial work during periods of sickness absence decrease the length of absence spells and induce higher subsequent employment rate.

The literature described above, shows that both adverse employment conditions and loss of employment may affect health. In chapter 2 in paper IV, I show that a large body of research has documented that health also is important for earnings and the probability to be employed. The aim of paper IV is to investigate how employment propensities and earnings of vulnerable groups, such as individuals with poor health, have developed relative to the population at large. The findings from paper IV indicate that poor health, to some extent, has become a stronger predictor of non-employment for young adults (age 27-31). However, the results are sensitive to what kind of indicator I use. For men, I find that low birth weight and underweight at age 18-19 affect the probability of being employed, and even to a greater extent over time, while there is quite constant impact of low height and obesity. For women, where birth weight is the only health measure I have, there is no evidence of changing impact.

When it comes to the impact of social background I find clear evidence of steadily increasing negative impact on employment of growing up in the very poorest families, which is in line with findings by Markussen and Røed (2017).[13] All the health indicators I use, as well as the measure for cognitive ability, are correlated with social background. However, the health indicators are imprecise measures of health. , meaning that I am not able to separate out how much of the social gradient is explained by health

The multidimensional aspects of health suggest that my results from paper IV may imply that some dimensions of health has become more important in order to enter and/or remain in the labor market, while other aspects have become less important. The finding that underweight has become a stronger predictor of earnings and non-employment, while there is quite constant effect of height and obesity may indicate that it has become more difficult for young adults to be

[13] Markussen and Røed (2017) also find that poor social background has become more strongly associated with other outcomes, such as educational attainment, and the establishment of a family.

employed when having mental health problems. One possible explanation is that underweight individuals to a greater degree face common mental health problems, such as anxiety and depression.[14] With more emphasize on reorganizations and efficiency it is reasonable to assume that good mental health is more strongly rewarded in the labor market.

One limitations of paper IV is that I focus on labor market outcomes among young adults only, which may imply that I describe more of the process of sorting into jobs than out of jobs. At older ages, individuals tend to experience physical health problems to a greater degree, which may also influence the generalizability of the results.

If the pressure on efficiency and productivity has increased, it is likely that both employer and employees exploit the social insurance system to a greater degree than before. Workers who are, for shorter or longer periods, not able to perform optimally may be encouraged to claim sickness benefits or more long-lasting benefits. Since the social security costs (except the first 16 days of sickness absence) are covered by payroll taxes without experience rating, there are weak incentives for exerting reintegration efforts.

**Other explanations**

There may be other explanations of the high fraction of individuals receiving health-related benefits in Norway. First, the age composition may play an essential role since older individuals are more likely to be on sickness absence and disability insurance. Kann and Sutterud (2017) suggest that this was part of the explanation of increasing fraction of health-related benefits during the period 1992-2016. Second, the rate of family support has decreased (Terum and Hatland, 2014) which may imply that some individuals earlier supported by family members, for shorter or longer periods, are now supported by the welfare state.

Third, there may be changes in health having contributed to the rise in use of health-related benefit. Improvements in medical technology has decreased infant mortality rates and probably

[14] Both underweight, overweight and obesity is associated with increased risk of dying. Whether underweight, in itself, is a risk factor or whether there is a reverse causation in the sense that underweight is a result of preexisting illness and smoking is not clear (see f.ex. Aune *et al*; 2016, Roh *et al*; 2014). Based on data from Switzerland, Roh *et al*. (2016) find that higher mortality risk among underweight individuals are mainly caused by a higher risk of external causes of deaths (e.g. accidents, suicides). Using survey-data from England, Kelly *et al*. (2010) show that underweight individuals are more likely to smoke, to be alcohol obtainers and inactive. In addition, the study find that both underweight and overweight/obesity are associated with a higher prevalence of respiratory disease, less physical activity and poorer mental health.

had a positive effect of most of the survivors (Bharadwaj *et al.*, 2013), but it may also have caused serious health problems for those who were on the margin of survival as babies. Brage and Thune (2015) show that people granted permanent disability insurance before age 25 are dominated by persons with mental retardation and congenital malformations, and part of the recent rice of young benefit receivers is attributed to an increasing fraction of the adult population with severe and chronic health impairments.[15] In general, how better medical treatment affect health for the whole population is not obvious. Some will recover faster and live better with a given diagnosis, but decreased mortality rates may also imply that a larger fraction is living with chronical conditions.

Musculoskeletal pain and mental disorders account for the majority of the disability claimants. Among the youngest, were the disability rates have increased most, more than 50% of the benefit receivers are suffering from mental illness. According to a Norwegian report about mental health in Norway, there is no evidence of increasing rates of mental illness in the population in general. However, they observe an increase in the fraction of young women reporting mental health problems, such as anxiety or depression.[16]

A widening of the conception of disease may have contributed to rising disability rates. One commonly used example of medicalization is Attention Deficit Hyperactivity Disorder (ADHD) were diagnosis and treatment started in the US in the 1960's. Since the early 1990's the disease has become a more global phenomenon and many western countries document an expansion of the diagnosis (Conrad and Bergey, 2014). A growing number of children diagnosed with ADHD is also seen in Norway, and the prevalence of children age 6-17 receiving medical treatment for ADHD is now 3.0% for boys and 1.2% for girls.[17] The diagnosis is set early, normally before age 12, but often continue to persist into adulthood. Thus, we also observe a growing number of adults with ADHD (Conrad and Bergey, 2014).

### *1.5 Policy implications*

The high rates of health-related benefits has been a matter of concern for policy makers during the recent decades. In light of the four papers in this thesis and relevant literature I will discuss different

[15] The fraction of individuals age 18-25 receiving permanent disability insurance increased from 0.8% in 1992 to 1.6% in 2017. Mean age at death for persons with Down syndrome has increased from 16 years in 1969 to 53 years in 2009, which is related to an increasing infant survival rates (Brage and Thune, 2015).
[16] See https://www.fhi.no/publ/2018/psykisk-helse-i-norge/
[17] See https://www.fhi.no/publ/2016/adhd-i-norge/

strategies to reduce the use of health-related benefits; changes in the compensation level, larger copayment for the employers and activity requirements.

**Changes in the compensation level of different benefits**

The compensation rate of sickness benefits has been cut in many western countries and it has been well documented that this is an effective way to reduce absenteeism. In Norway, the replacement rate has been 100% since 1978. The literature has focused on the incentives of the employees when explaining the sensitivity of the compensation rate. But as we have shown in paper II also employers respond to economic incentives. Hence, a reduction in the compensation level of sickness benefits may also affect how employers will prevent absenteeism.

The impact on workers of a reduction in sick pay is well-known; when absence become more costly, the employees will tend to be less absent. For employers, short-term absence (absence less than 16 days) will become *less* costly, meaning that employers will have less incentives to prevent short-term absenteeism. On the other hand, long-term absenteeism will become *more* costly since they will no longer receive full wage-compensation for the absent workers. This may, in turn, reduce transition into disability insurance. However, not every case of sickness absence is undesirable and being present at work when ill may adversely affect others, f.ex. when having a contagious condition or when the illness make the worker more prone to do fatal errors.[18] Sometimes, periods out of work may also be beneficial for health and might prevent future absences.

Although experiences from other countries show that cutting benefits reduce sickness absenteeism, studies have also found that groups of individuals may respond differently. Ziebarth (2013) exploits a reform of the sickness absences system in Germany, were the replacement level of short-term sickness absence (absence shorter than six weeks) was reduced from 100 to 80% while the compensation level of long-term absence spells was cut from 80 to 70%. He finds that benefit cuts reduce short-term absenteeism, but it has no significant effect on long-term absenteeism. There was, however, some exceptions when focusing on subgroups. The lack of an overall effect on long-term absenteeism may be related to the already strict follow-up regime for long-term absent workers in Germany. Using the discontinuity in public sick pay in Norway, Dale-

[18] Pichler and Ziebarth (2017) shows that universal access to paid sick leave decrease the rate of influenza-like disease.

Olsen (2014) investigates how high earnings workers respond to economic incentives, and shows that higher compensation level reduces absenteeism for male workers, but not for female workers.

The findings illustrate that it is not obvious whether, and to what extent, different groups will respond to changes in the replacement level. Dale-Olsen (2018) also points out that a reduction may be offset by supplementary compensation from some employers since part of the firms already provide additional benefits for their high-earnings workers. The replacement level of other benefits is much lower compared to sickness benefits, meaning that the long-term costs of being absent from work is much higher than the short-term cost. However, individuals already receiving long-term benefits do not have a job, and as explained above, around 3-5% is estimated to be better off if they continue to be benefit-receivers instead of working. During recent years, policy makers have tried to overcome this problem, f.ex by making it easier to combine work and benefits.

Despite the relatively generous benefit level of social security in Norway, it does not necessarily imply that cutting benefit level is the desired policy from a welfare perspective. Generous benefits protect those who are unable to work, as well as their dependents, from poverty. In addition, it also provides claimants who ultimately have the capacity returning to work with more time in which to find a suitable and viable job match.

The recent restrictions of the unemployment insurance scheme have excluded a substantial fraction of job seekers from claiming benefits. Both the earnings requirement for being entitled to unemployment insurance have become stricter during the last decades and maximum duration of an unemployment spell has been shortened. Apart from social assistance, temporary disability insurance, which requires that the individuals are being certified by a diagnosis, is the only alternative income support for some non-employed individuals. This has contribute to more substitution away from unemployment insurance, particularly among the youths with no work experience. A way to reduce medicalization and overuse of disability insurance may be to lessen the earnings criteria for being eligible for unemployment insurance (Schreiner, 2017).

**Increased copayment of employers**

A copayment system of sickness absence, where employers cover the first days of absence and the public insurer covers the rest, is the case in many industrialized countries – and the results from paper II show that it may create a sick pay trap. Firms have strong incentives of preventing short-term absence, but when sickness absence last longer than the copayment-period employers have

little incentive of encouraging a quick return to work since the employers will (again) be responsible for the first days of an eventually new period of absence.

The series of reforms in the Netherlands, also involving higher copayments of sickness absence and disability insurance, have reduced the inflow into disability insurance substantially (Koning and Lindebom, 2015). Although the reforms have been successful in order to reduce the fraction of individuals receiving health-related benefits, the side effects may be prominent. The study of Koning and Lindebom (2015) indicates that the reforms have increased the share of unhealthy non-working individuals without benefits. In addition, firms may be less willing to hire workers with health limitations. Policy makers in the Netherlands are now discussing other ways of organizing the system of social security system which provide strong incentives of preventing sickness and disability insurance, but with less financial risks for the firms.

Taken together, the experiences from the Netherlands and results from paper II illustrate that employers substantially influence absence from work. The design of firm incentives with respect to prevention of sickness absence also involves a potential trade-off between sick leave and labor market exclusion: A higher fraction of copayment for employers trigger financial incentives to prevent absenteeism, but it may reduce the incentives of employing persons with high risk of sickness absence. A possible solution to circumvent this problem is to reorganize the system for pay liability in the sense that firms cover a smaller fraction of costs associated with short-term absence and a larger fraction of long-term absence.

**Activity requirements**

In July 2004, the social security administration introduced new rules for sickness absence in Norway. The new rules include stricter activity requirements, f.ex. (more) use of graded absenteeism and compulsory dialogue meeting for long-term absent workers.

Since the introduction of the new rules, employers are obliged to arrange a meeting for workers during the first seven weeks of sickness absence. The aim of the meeting is to agree on a plan to facilitate a quick return back to work, f.ex by graded absenteeism. If the absent worker has not recovered within 26 weeks the local social security administration is obliged to arrange a second meeting with the absent employees, the employer and the physician having certified the absence. How this second meeting is followed-up varies between counties and Markussen *et al.*, (2017) use this as an instrument to evaluate the reform. They find that dialogue meeting reduces the length of

absence spells. Markussen *et al.* (2012) evaluate how more use of graded absenteeism as opposed to full absenteeism by utilizing different leniency among physicians to grade sickness absenteeism. Their estimates suggest that grading an absence spell shortens the length of absence spell and raises subsequent employment propensity.

The results associated with grading and dialogue meetings may justify a more intense usage of those measures, but it is not necessarily the case that more involvement of the physician is the way to go. The physician plays a key role in case of medical interventions, but the physician may have less information about how a given illness will influence work capacity. A more fruitful way to prevent long-term absenteeism may be to arrange for a tighter dialogue between the employer and the absent worker.

One example of decreased involvement by the physician and increased follow-up by the employers is a reform being introduced in a Norwegian municipality in 2003. The employees working in the municipality were given the right to self-certify themselves for sickness absence for the entire benefit period (365 days). The project involves frequent meetings with the employer and the employer's follow-up of the absent employee is more intensive than elsewhere. Torsvik and Vaage (2015) have evaluated this reform and they find that sickness absence is reduced by more than 20%. Whether this reduction in sickness absence is due to reduced control or increased trust is not obvious since intensified follow-up of absent employees (which is a part of the reform) could be interpreted either as increased concern for the worker or increased monitoring (see Torsvik and Vaage, 2015).[19]

The introduction of the Gatekeeper protocol in 2002 has contributed to reduce inflow into disability insurance in the Netherlands (Koning and Lindebom, 2015). This protocol focuses on early interventions, and within the first eight weeks of absence the employers and workers is obliged to make a return-to-work plan including several dates to modify and evaluate the plan. The plan should make a first assessment of functional limitations and possible need for medical interventions as well as a date for work-resumption. If planned work-resumption has not taken place, the employers have to attach a new assessment to explain the non-fulfillment of the original

[19] In lab experiments, Falk and Kosfeld (2006) find that a majority of the agents reduce their performance when they are controlled compared to a situation when the principal trust their agents. The reason is that most of the agents in the experiment are intrinsically motivated to perform well and the agents perceive control as a sign of distrust and a limitation of own choices, which in turn reduce their motivation. Opportunistic agents, however, perform more (i.e. the minimum level) when they are controlled.

plan. The social security administration could then be responsible for sick pay, but only if they approve the return-to-work plan.

In the Netherlands, they have also tried to reduce disability rates by increasing outflow from disability insurance. These reforms have more uncertain effects (Sonsbeek and Gradus, 2012) which may be related to restricted job opportunities for individuals being out of work for a couple of years.

### *1.6 Concluding remarks*

In Norway, the employment rates are high and the unemployment level is low. However, the fraction receiving health-related benefits is high, which to a large extent, hide an unemployment problem. The rules for being eligible to unemployment benefits have become stricter during the last decades, which may have contributed to more substitution away from unemployment insurance to health-related benefits. A way to reduce medicalization and overuse of disability insurance may be to lessen the criteria for being eligible for unemployment insurance.

Experiences from the Netherlands show that employers play a key role in order to reduce the inflow rate to disability insurance, but large financial cost associated with sickness and disability has probably made the employers more reluctant to hire presumably unhealthy workers. Anyway, the copayment system we have in Norway, where the employer is financial responsible for short-term absence and the public insurers covers the rest, undermine the firms' incentives to prevent long-term absenteeism. A better solution may be to let the firms cover a smaller fraction of costs associated with short-term absence and a larger fraction of long-term absence. Some measures, such as grading of absenteeism and dialogue meetings, have already been implemented in Norway and seems to have the intended effect. These activity requirements have primary focused on the role of the physician, it may be that subsequent measures should emphasize increased follow-up from the employers. By extending the financial responsibility of the employers, their incentives of preventing long-term absenteeism will be promoted.

## References


Aaviksoo, E. and Kiivet, R. A. (2016). Influence of the Sickness Benefit Reform on Sickness Absence. *Health Policy*, Vol. 120(9), 1070-1078.

Acemoglu, D. and Author, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. *Handbook of Labor Economics*, 4 (2011), 1043-1171.

Aune, D., Sen, A., Prasad, A., Norat, T., Janszky, I., Tonstad, S., Romundstad, P. and Vatten, L. J. (2016). BMI and all cause Mortality: Systematic Review and Non-Linear Dose-Response Meta-Analysis of 230 Cohort Studies with 3.74 Million Deaths among 30.3 Million Participants. *BMJ* 2016;353:i2156.

Autor, D. and Duggan, M. (2003). The Rise in the Disability Rolls and the Decline in Unemployment. *The Quarterly Journal of Economics*, Vol. 118(1), 157-206.

Beaudry, P., Green, D. A. and Sand, B. M. (2014). The Declining Fortunes of the Young since 2000. *American Economic Review*, *Papers and Proceedings,* Vol. 104(5), 381-386.

Bharadwaj, P., Lundborg, P. and Rooth, D.-O. (2018). Birth Weight in the Long Run. *Journal of Human Resources*, Vol. 53(4), 189-231.

Bharadwaj, P., Løken, K. V. and Neilson, C. (2013). Early Life Health Interventions and Academic Achievement. *American Economic Review*, Vol. 103(5), 1862-1891.

Black, D., Daniel, K., and Sanders, S. (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.

Black, S. E., Devereux, P. J. and Salvanes, K. G. (2015). Losing Heart? The Effect of Job Displacement on Health. *ILR Review*, Vol. 68(4), 833-861.

Bound, J. and Johnson, G. (1992). Changes in the Structure of Wages in the 1980's: An Evaluation of Alternative Explanations? *American Economic Review*, Vol. 82(3), 371-392.

Boyd, K. M. (2000). Disease, Illness, Sickness, Health, Healing and Wholeness: Exploring some Elusive Concepts. *Journal of Medical Ethics: Medical Humanities*, Vol. 26(1), 9-17.

Brage, S. and Thune, O. (2015). Ung Uførhet og Psykisk Sykdom. *Arbeid og Velferd*, No. 1, 2015.

Browning, M. and Heinesen, E. (2012). "Effect of Job Loss due to Plant Closure on Mortality and Hospitalization" *Journal of Health Economics*, Vol. 31(4), 599-616.

Burkhauser, R. V. and Daly, M. C. (2011). The Declining Work and Welfare of People with Disabilities: What Went Wrong and a Strategy for Change. *AEI Press*, Washington D.C.

Bøckerman, P., Kanninen, O. and Suoniemi, I. (2018). A Kink that makes you Sick: The Effect of Sick Pay on Absence. *Journal of Applied Econometrics*, Vol. 33(4), 568-579.

Card, D., Chetty, R. and Weber, A. (2007). The Spike at Benefit Exhaustion: Leaving the Unemployment System or Starting a New Job? *American Economic Review, Papers and Proceedings,* Vol. 97 (2), 113-118.

Charles, K. K, Li, Y. and Stephens, M. (2018). Disability Benefit Take-Up and Local Labor Market Conditions. *Review of Economics and Statistics*, Vol. 100 (3), 416-423.

Conrad, P. and Bergey M. R. (2014). The Impending Globalization of ADHD: Notes on the Expansion and Growth of a Medicalized Disorder. *Social Science and Medicine*, Vol. 122, 31-43.

D'Amuri, F. (2011). Monitoring, Monetary Incentives and Workers' Rents in Determining Absenteeism. Working Paper, Mimeo, http://sites.google.com/site/fradamuri/home/research.

Dale-Olsen, H. (2018). Sykefravær og Inntektskompensasjon ved Sykefravær? *Søkelys på Arbeidslivet*, No. 1-2, 2018, 77-95.

Dale-Olsen, H. (2014). Sickness Absence, Sick Leave Pay, and Pay Schemes. *Labour*, Vol. 28(1), 40-63.

Deming, D. (2017). The Growing Importance of Social Skills in the Labor Market. *Quarterly Journal of Economics*, Vol. 132(4), 1593-164.

Edin, P. E., Fredriksson, P., Nybom, M. and Öckert, B (2017). The Rising Return to Non-Cognitive Skill. *IZA Discussion Paper No. 10914.*

Eliason, M., Johansson, P. and Nilsson, M. (2018). Forward-looking Moral Hazard in Social insurance. *IFAU Working Paper 2018:11*.

European Commision (2017). http://ec.europa.eu/social/main.jsp?catId=1134&langId=en.

Falk, A. and Kosfeld, M (2006). The hidden costs of control. *American Economic Review*, Vol. 96(5), 1611-1630.

Ferrie. J. E. (2001). Is Job Insecurity Harmful to Health? *Journal of the Royal Society of Medicine*, Vol. 94, 71-76.

Fevang, E., Nordberg, M. and Røed, K. (2006). Veien til Uføretrygd i Norge. *Frisch Report 10 (2006) Oslo*. The Frisch Centre.

Fevang, E., Nordberg, M. and Røed, K. (2005). Formelle og Totale Skattesatser på Inntekt for Personer med Svak Tilknytning til Arbeidsmarkedet. *Frisch Report 3 (2005) Oslo*. The Frisch Centre.

Fredriksson, P. and Holmlund, B. (2006). Improving Incentives in Unemployment Insurance: A Review of Recent Research. *Journal of Economic Surveys*, Vol. 20(3), 357–86.

Godøy, A. and Dale-Olsen, H. (2018). Spillovers from Gatekeeping - Peer Effects in Absenteeism. *Journal of Public Economics*, Vol. 167, 190-204.

Grasdal, A. (2016). De Helserelaterte Trygdeytelsene – Betydningen av Økonomiske Incentiver og Samspill mellom Trygdeordninger. *Tidsskrift for Velferdsforskning*, No. 2, 2016, 102-124.

Hallock, K. F. (2009). Job Loss and the Fraying of the Implicit Employment Contract. *Journal of Economic Perspectives*, Vol. 23(4), 69-93.

Hamermesh, D. (1987). The Costs of Worker Displacement. *Quarterly Journal of Economics*, Vol. 102(1), 51-76.

Hansen, J. (2000). The Effect of Work Absence on Wages and Wage Gaps in Sweden. *Journal of Population Economics*, Vol. 13, 45–55.

Hardoy, I., Torp, H. And Storvik, Aa. (2004). Hvem får Mer og Hvem får Mindre? Effekter av Nye Beregningsregler for Stønader til Livsopphold under Attføring og Rehabilitering. Rapport 2004:14, Institutt for samfunnsforskning.

Hauge, K. and Ulvestad, M. E. (2017). Having a Bad Attitude? The Relationship between Attitudes and Sickness Absence. *IZA Journal of Labor Policy*, Vol. 6(11) (https://doi.org/10.1186/s40173-017-0088-y).

Hernæs, Ø., Markussen, S. and Røed, K. (2016). Kompensasjongrader i Inntektssikringssystemet for Personer med Svak Tilknytning til Arbeidsmarkedet. *Frisch Report 1 (2016) Oslo*. The Frisch Centre.

Huber, M., Knottnerus, A., Green, L., van der Horst, H., Jadad, A. R., Kromhout, D., Leonard, B., Lorig, K., Loureiro, M. I., van der Meer, J. W. M. Schnabel, P., Smith, R., van Weel, C. And Smid, H. (2011). How Should We Define Health. *BMJ 2011;343:d4163*.

Huttunen, K., Møen, J., and Salvanes, K.G. (2011). How Destructive is Creative Destruction? Effects of Job Loss on Mobility, Withdrawal and Income. *Journal of the European Economic Association*, Vol. 9(5), 840-870.

Ichino, A. and Moretti, E. (2009). Biological Gender Differences, Absenteeism and the Earnings Gap. *American Economic Journal: Applied Economics*, Vol. 1(1), 183–218.

Johansson, P. and Palme, M. (2005). Moral Hazard and Sickness Insurance. *Journal of Public Economics*, Vol. 89, 1879-1890.

Kann, I. C. and Sutterud, L. (2017). Stadig Færre på Trygd? *Arbeid og Velferd*, No. 3, 2017.

Kelly, S. J, Lilley, J. M and J. Leonardie-Bee (2010). Associations of Morbidity in the Underweight. *European Journal of Clinical Nutrition*, Vol. 64, 475-482.

Kivimäki, M., Vahtera, J., Ferrie, J. E., Hemingway, H., and Pentti, J. (2001). Organisational Downsizing and Musculoskeletal Problems in Employees: A Prospective Study. *Occupational and Environmental Medicine*, Vol. 58 (12), 811-817.

Kletzer, L.G. (1998). Job Displacement. *Journal of Economic Perspectives*, Vol. 12(1), 115-136.

Koning, P. and Lindeboom, M. (2015). The Rise and Fall of Disability Insurance Enrollment in the Netherlands. *Journal of Economic Perspectives*, Vol. 29(2), 151-172.

Kuhn, P.J., ed. (2002). Losing Work, Moving On. International Perspectives on Worker Displacement. W.E. Upjohn Institute for Employment Research, Kalamazoo, Michigan.

Markussen, S. and Røed, K. (2017). Egalitarism under Pressure: Toward Lower Economic Mobility in the Knowledge Economy? *IZA Discussion Paper No. 10664*.

Markussen, S., Røed, K. and Schreiner, R. (2017). Can Compulsory Dialogues Nudge Sick-Listed Workers Back to Work? *Economic Journal*, Vol. 128(610), 1276-1303.

Markussen, S., Mykletun, A and Røed, K. (2012). The Case for Presenteeism – Evidence from Norway's Sickness Insurance Program. *Journal of Public Economics*, Vol. 96(11), 959-972.

Markussen, S. (2012). The Individual Cost of Sick Leave. *Journal of Population Economics, Vol.* 25(4), 1287-1306.

Markussen, S., Røed, K., Røgeberg, O. J. and Gaure, S. (2011). The Anatomy of Absenteeism. *Journal of Health Economics*, Vol. 30(2), 277-292.

Meyer, B. D. and Mok, W.K.C. (2018). Disability, Earnings, Income and Consumption, *Journal of Public Economics*, forthcoming.

Michaud, P.-C., Crimmins, E. and Hurd, M. (2016). The Effect of Job Loss on Health: Evidence from Biomarkers *Labor Economics*, Vol. 41, 194-203.

Morken, T., Haukenes, I. and Magnussen, L. H. (2012). Attending Work or not when Sick – What Makes the Decision? A Qualitative Study among Car Mechanics. *BMC Public Health*, Vol. 12:813.

Neal, D. (1995) Industry-Specific Human Capital: Evidence from Displaced Workers. *Journal of Labor Economics*, Vol. 13(4), 653-677.

Nordberg, M. (2007). The Total Tax on Labour Income. *Memorandum 5 (2007) Department of Economics. University of Oslo*.

OECD (2017). Health at a Glance 2017. https://www.oecd-ilibrary.org/docserver/health_glance-2017en.pdf?expires=1542096983&id=id&accname=guest&checksum=7CA76831A8571905AF66E3995ACFAD21

OECD (2013). Mental Health and Work: Norway, OECD Publishing, Paris. https://www.oecd-ilibrary.org/employment/mental-health-and-work-norway_9789264178984-en

OECD (2010). Sickness, Disability and Work: Breaking the Barriers - A Synthesis of Findings across OECD Countries, OECD, Paris.

OECD (2008). Are All Jobs Good for Your Health? The Impact of Work Status and Working Conditions on Mental Health. Chapter 4 in OECD Employment Outlook 2008, Paris.

Pichler, S. and Ziebarth, N. (2017). The Pros and Cons of Sick Pay Schemes: Testing for Contagious Presenteeism and Noncontagious Absenteeism Behavior. *Journal of Public Economics*, Vol. 156, 14-33.

Rege, M., Telle, K., and Votruba, M. (2009). The Effect of Plant Downsizing on Disability Pension Utilization. *Journal of the European Economic Association*, Vol. 7(4), 754–785.

Roh, L., Braun, J., Chiolero, A., Bopp, M., Rohrmann S. and Faeh, D. (2014). Mortality Risk Associated with Underweight: A Census-Linked Cohort of 31,578 Individuals with up to 32 Years of Follow-up. *BMC Public Health*, 14:371.

Ruhm, C. J. (1991) Are Workers Permanently Scarred by Job Displacements? *American Economic Review*, Vol. 81(1), 319-324.

Røed, K. (2012). Active Social Insurance. *IZA Journal of Labor Policy*, Vol. 1:8 (doi:10.1186/2193-9004-1-8).

Røed, K., Jensen, P., and Thoursie, A. (2008). Unemployment Duration and Unemployment Insurance - A Comparative Analysis Based on Scandinavian Micro Data. *Oxford Economic Papers*, Vol. 60(2), 254-274.

Røed, K., and Fevang, E (2007). Organizational Change, Absenteeism and Welfare Dependency. *Journal of Human Resources*, Vol. 42(1), 156–193.

Røed, K. and Zhang, T. (2005). Unemployment duration and economic incentives - A quasi random-assignment approach. *European Economic Review*, Vol. 49(7), 1799-1825.

Schaller, J. and Stevens, A. H. (2015). Short-run effects of Job Loss on Health Conditions, Health Insurance and Health Care Utilization. *Journal of Health Economics*, Vol. 43, 190-203.

Schreiner, R. C. (2017): Unemployed or Disabled? Disability Screening and Labor Market Outcomes of Youths" in *Social Insurance Programs, Earnings and Labor Force Participation*, Dissertation for the PhD Degree, University of Oslo, 14.10.2016.

Van Sonsbeek, J. M. and Gradus, R. H. J. M. (2013). Estimating the Effects of Recent Disability Reforms in the Netherlands. *Oxford Economic Papers*, Vol. 65(4), 832-855.

Terum, L. I. and Hatland, A. (2014). Sysselsetting og Trygd under Arbeidslinja? *Søkelys på Arbeidslivet*, No. 1-2, 2014, p. 3-22.

Torsvik, G. and Vaage, K. (2015). Gatekeeping Versus Monitoring: Evidence from a Case with Extended Self-Reporting of Sickness Absence. *CES-ifo Working Paper*, no.5113.

Waddell, G. (2004). The Back Pain Revolution (Second Edition). Elsevier Health Sciences.

Waddell, G. and Burton, A.K. (2006). Is Work Good for your Health and Well-being? Report for Department for Work and Pensions. TSO (The Stationery Office), London.

WHO (1946) Preamble to the Constitution of the World Health Organization. WHO, New York, USA

WHO (2011). World Report on Disability 2011. WHO, New York, USA

Wikman, A. Marklund, S. and Alexanderson, K (2005). Illness, Disease and Sickness Absence: An Empirical Test of Differences between Concepts of Ill Health. *Journal of Epidemiology and Community Health*, Vol. 59(6), 450-454.

Ziebarth, N. (2013) Long-term Absenteeism and Moral Hazard – Evidence from a Natural Experiment. *Labor Economics*, Vol. 24, 277-292.

Ziebarth, N. and Karlsson, M. (2014) The Effects of Expanding the Generosity of the Statutory Sickness Insurance System. *Journal of Applied Econometrics*, Vol. 29(2), 208-230.

# Chapter 2:

Temporary disability and economic incentives

# Chapter 3:

The sick pay trap

# The Sick Pay Trap

Elisabeth Fevang, *Ragnar Frisch Centre for Economic Research*

Simen Markussen, *Ragnar Frisch Centre for Economic Research*

Knut Røed, *Ragnar Frisch Centre for Economic Research*

In most countries, employers are financially responsible for sick pay during an initial period of a worker's absence spell, after which the public insurance system covers the bill. Based on an empirical evaluation of a quasi-natural experiment in Norway, where pay liability was removed for pregnancy-related absences, we show that the system of short-term pay liability creates a sick pay trap: firms are discouraged from letting long-term sick workers back into work since they then face the financial risk associated with subsequent relapses. We present evidence indicating that this disincentive effect is both statistically and economically significant.

## I. Introduction

Based on extensive reviews of disability prevention experiences in 13 countries, the Organization for Economic Cooperation and Development (OECD 2010, 125) argues that "employers are key players in preventing health problems at work and facilitating a swift return to work for people absent from work due to sickness." But, while there is ample em-

This article is part of the project "Absenteeism in Norway—Causes, Consequences, and Policy Implications" and is financed by the Norwegian Research Council (grant no. 187924). Data made available by Statistics Norway have been essential for the research project. Thanks to Christian Brinch, Oddbjørn Raaum, and Christopher Taber for useful comments and to Simen Gaure and Tao Zhang for programming assistance. Contact the corresponding author Knut Røed at knut.roed@frisch.uio.no. A note regarding access to data and programs is available as supplementary material online.

[*Journal of Labor Economics*, 2014, vol. 32, no. 2]

0734-306X/2014/3202-0004$10.00

pirical evidence regarding the responsiveness of absenteeism with respect to worker incentives (Henreksson and Persson 2004; Johansson and Palme 2005; Ziebarth 2009; Ziebarth and Karlsson 2009; D'Amuri 2011), there is little evidence regarding the impact of firm incentives. OECD (2010, 133) notes that countries where employers are responsible for a large share of their employees' sick pay costs tend to have much lower absence rates than countries where employers can pass the costs on to the public purse and also that absenteeism has dropped significantly in the Netherlands and the United Kingdom after a shift of financial responsibility toward employers. Yet, to our knowledge, scientific evidence establishing a causal relationship between firm incentives and worker absenteeism is nonexistent. The design of firm incentives with respect to sick leave prevention also involves a potential trade-off between sick leave and labor-market exclusion: while more extensive pay liability (or experience rating) improves incentives for absence prevention, it may at the same time undermine incentives for employing persons perceived to have a high risk of absence in the first place.

Although the extent of employer co-payment differs sharply across different countries, the incentive structure faced by employers in most industrialized economies typically implies that the firm is responsible for sick pay expenditures during an initial stage of a workers' sick leave but that the national insurance scheme (or another insurer) covers the costs accruing after some duration threshold (OECD 2010, table 5.1).[1] This means that firms do have strong financial incentives to prevent short-term absences. However, in cases where absence spells stretch beyond the co-payment period, employers may (rightly) think that it is not in their interest to facilitate a quick return to work, since the return to work also entails the risk of new short-term absences for which the employers are again financially responsible. Hence, current incentive structures may have the unintended side effect of discouraging employers from exerting appropriate effort to curb long-term absenteeism.

In the present article, we examine empirically the impacts of employers' pay liability by exploiting a reform in the Norwegian sick leave insurance scheme, whereby pay liability for pregnant workers' short-term sick leaves was removed for absences caused by pregnancy-related illnesses (in the

[1] While most European countries have public insurance programs for short-term sick leave, typically with replacement rates ranging from 50% to 100% and maximum durations exceeding 1 year (Petterson-Lidbom and Thoursie 2013), there are no federal legal requirements for paid sick leave in the United States. For companies subject to the Family and Medical Leave Act (FMLA), i.e., firms with at least 50 employees, the act does require unpaid sick leave, however. And many US employers offer sick leave pay on a voluntary basis. Also, some states, including California, New Jersey, and New York, have public programs that partially protect workers against the loss of income due to nonoccupational disability.

period prior to transition to maternity leave benefits). The motivation for this reform was that it was feared that the elevated risk of sickness absence associated with pregnancies made employers reluctant toward hiring young women. Markussen et al. (2011) show that the increased risk of absenteeism associated with pregnancies is indeed substantial; the hazard rate of entering into a sick leave spell with a diagnosis predicting full exploitation of the employer's pay liability—often related to back pain, nausea, or anxiety—is raised by a factor of 5 at the onset of a pregnancy and further to a factor of 15 during the last 2–3 months before delivery. The reform thus clearly removed a potentially important disincentive with respect to hiring young female labor, but at the same time it also changed employers' incentives to prevent sick leave among pregnant workers, for example, by reducing productivity expectations/requirements, modifying regular duties, encouraging colleagues to help out with physically demanding tasks, allowing more flexible and less strenuous hours, or investing in technical (strain-reducing) equipment.

On the one hand, the reform made short-term absence—absence spells with durations up to 16 days—less costly for the firms. On the other hand, it also made it less risky to let long-term absent pregnant workers return to work, since the firms no longer were responsible for the sick pay costs associated with subsequent relapses. Hence, the reform offers a neat setting for identifying the impacts of firm incentives. Based on a combination of regression discontinuity (RD) and difference-in-difference (DiD) methodologies, we show that the reform had significant impacts on the affected employees' absence behavior. Using a bivariate hazard rate model (with nonparametric modeling of unobserved heterogeneity), we find that short-term absenteeism originally covered by firms' pay liability rose significantly, while the duration of noncovered long-term spells declined. For example, we estimate that the reform led to a 10% rise in the entry rate into absenteeism but also in a 12% rise in the rate of work resumption at durations exceeding 16 days. Extrapolating our findings to the economy as a whole, we estimate that the 16-day pay liability system reduces the overall number of covered absence days by 11%, while actually raising the number of noncovered days (days exceeding 16 days duration) by 1.6%, compared to a system with no pay liability at all. Our results are exactly as one would expect on the basis of simple economic theory, provided that firms do have some influence on their employees' sick leave behavior. Our findings thus indicate that policy makers indeed may have good reasons to focus on improving employer incentives in their efforts to curb absenteeism.

We also find some evidence indicating that the removal of pay liability for pregnant workers made the school-to-work transition easier for young women. According to our estimates, the reform raised the employment propensity 1 year after graduation by around 1.5 percentage points for

young women in general and by 3.0 percentage points for those who were pregnant at the time of graduation. This suggests that policy makers may have to trade off incentives for sick leave prevention against incentives for employing workers with high expected absenteeism.

## II. Related Literature

When an insurance scheme is troubled by moral hazard problems, efficiency considerations suggest that coverage should decrease with duration. This is studied extensively in relation to the design of optimal unemployment insurance, starting from the seminal paper by Shavell and Weiss (1979). The argument is simple. In the presence of moral hazard, there is an inevitable trade-off between insurance and incentives. By reshuffling the benefit schedule to provide lower payments tomorrow and higher payments today, such that expected utility remains constant, agents are given stronger incentives to search for jobs.

Many countries have adopted declining benefit schedules for the unemployed, most often in the form of a single drop after some time in addition to an overall duration limitation (Cahuc and Zylberberg 2004, 143). Maximum duration limitations are also typically in place for sickness insurance payments. However, there are also historical examples of sickness benefits that increase with duration. Johansson and Palme (2005) study a reform in Sweden, where a time-constant replacement rate of 90% was replaced by a time-increasing payment schedule with 65% replacement rate the first 3 days, 80% the next 77 days, and then 90% from day 80. Johansson and Palme (2005) found that the reform changed workers' behavior exactly as the altered incentives would imply. As the cost of short-term absence increased, short-term absenteeism dropped. For long-term absences, however, the return-to-work hazard declined as the risk of relapse raised the expected cost of returning to work. This example illustrates a potential benefit trap: having reached the highest level of replacement, it is not particularly tempting to risk a return to the bottom of the replacement ladder.

In the present article, we focus on the employers' incentives rather than those of the employees. But, provided that employers influence their employees' sick leave behavior, the story is basically the same. If the employer is financially responsible for short-term absence only, the firm obviously has incentives to prefer a single long absence spell over many short ones. And when employees have been on sick leave long enough to have exhausted the firms' pay liability, the financial risk associated with possible relapses may convince the employer not to accommodate a quick return to work.

While it has long been recognized in the literature that incomplete experience rating in unemployment insurance systems incites firms to lay off too many workers (e.g., Topel 1983), we have not been able to find any

empirical evidence regarding the causal relationship between firms' sick leave insurance costs and their workers' absenteeism. One of the closest pieces of evidence we have found is Burkhauser, Schmeiser, and Weathers (2012), which shows that disabled workers in the United States whose work limitations were caused by an accident on the job, and who were, hence, more likely to be covered by the experience-rated workers' compensation program than workers whose injuries were not work related, also were more likely to be offered accommodation by their employer.

Our article also relates to a literature on the labor-demand effects of mandated employer provision of employee benefits in general (see Burkhauser and Daly [2011] for a recent discussion). A contribution to this literature with particular resemblance to our own is Gruber (1994), which found that mandates that raised the costs of insuring female employees of childbearing age in the United States (by including childbirth in health insurance plans) did not adversely affect young women's employment rates, as wages adjusted to compensate for the value of the mandated benefits. It is conceivable, however, that this finding could be reversed in labor markets with less individual wage flexibility—like the Norwegian one.

## III. Institutions and Mechanisms

All Norwegian workers are fully insured against sickness absence for up to 1 year, with a 100% replacement ratio.[2] Norway also has a high level of absenteeism. On a typical working day, around 7% of all workers are absent due to sickness. This places Norway among the countries with the world's highest sickness absence rates (see, e.g., Bonato and Lusinyan 2007; Edwards and Greasley 2010). Standard regulations imply that absence spells exceeding 3 days (including weekends, holidays, etc.) need to be certified by a physician. Certification is not formally required until the ninth day for employees in firms participating in a tripartite "inclusive workplace agreement" (IWA) between employers, employees, and the state, although it is common to adhere to the 3-day rule even in these firms.[3] Approximately half of the employees are covered by IWA. In total, around 90% of all absence days in Norway are certified by a physician. The sick pay costs are shared between the employer and the tax payers. The general rule is that

[2] There is a ceiling on annual earnings (in 2012) of NOK 500,000 ( ≈US$85,000). Eligibility requires that the employee have been at work for 4 weeks.

[3] Based on an in-depth study of a large Norwegian IWA-company, Bergsvik, Markussen, and Raaum (2010) show that more than 90% of all self-reported absence spells are 1–3 days long (less than 1% are longer than 5 days), despite the possibility of self-reporting up to 8 days. Statistics reported by the Norwegian Welfare Administration also indicate that the level of physician-certified absence is higher in IWA-firms than in non-IWA-firms.

the first 16 days of each absence spell is paid for directly by the employer, whereas the social security administration pays for the remaining days and also for subsequent rehabilitation or disability benefits.[4] If a new absence spell starts within 16 days after a previous spell was completed, it is counted as a continuation of the previous spell. This implies that a new pay liability period for the firm is not triggered until the worker has been present for at least 16 days. The social security costs are covered through uniform payroll taxes; hence, there is no experience rating. On average, the pay liability system implies that firms' cover around 34% of the overall sick pay costs for Norwegian workers (see Bjerkedal and Thune 2003). In addition, they bear the administrative costs associated with finding replacements and/or reorganizing the work. Although Norwegian firms are not allowed to cut wages in direct response to an employee's absence behavior, existing empirical evidence indicates that some of the costs are passed on to the employees through an impact on subsequent wage growth (see Markussen 2012).

In April 2002, a reform was implemented implying that firms were entitled to exemption from the 16-days pay liability for pregnancy-related absences. Common symptoms associated with pregnancy-related diseases are nausea, bleeding, anxiety, infections, reduced functional level (e.g., due to back pain), and various psychological reactions. The term "pregnancy related" obviously entails an element of subjective judgment, since pregnant workers also may develop illnesses that have little or nothing to do with their pregnancies. Based on the observed differences in absenteeism between pregnant workers and a comparison group of female nonpregnant colleagues that we present in the next section of this article, we estimate that approximately 73% of the absence among pregnant workers is pregnancy related. We do not observe whether a particular pregnant worker's absence spell is really pregnancy related or not.

During periods of sickness absence, Norwegian workers enjoy a special protection against dismissals, implying that they cannot be dismissed on grounds that are related to their sickness.[5] After the 1-year absence period, however, the firm is allowed to lay off the absent worker with direct reference to the sickness. Hence, if an employer for some reason wishes to lay off a worker—but is prevented from doing so due to the general employment protection regulations—the incentives for facilitating that worker's return to work from a long-term absence spell are particularly weak.

[4] Workers who have exhausted their sick pay entitlements but who are still not able to take up work due to sickness are eligible for rehabilitation or disability benefits, depending on the prospects for future recovery/rehabilitation. The replacement ratio for these benefits is around 66%.

[5] The burden of proof lies with the firm. In practice, this implies that absent workers can only be laid off as part of a mass displacement.

Although absence decisions are formally taken by workers and/or physicians, the employer can affect absenteeism in several ways, for example, by (i) monitoring employees (to prevent shirking), (ii) investing in healthy work environments and equipment that can prevent afflictions caused by strain, (iii) offering sick workers modified tasks or changes in the requirements of their jobs, and (iv) allowing the workers more flexible hours and less shift work. These activities obviously involve costs; hence, we may expect employer efforts to depend on the extent to which they bear the direct costs of absenteeism in the first place.

Norway has a generous public program for parental leave benefits. In the period covered by our analysis, there was a paid leave period for 42 weeks with 100% replacement ratio (or 52 weeks with 80% replacement), which has later been raised to 47 weeks (57 weeks with 80% replacement).[6] The leave starts no later than 3 weeks before expected delivery, removing the pregnant workers from the risk of sick leave at this point. It is possible to start the leave period as early as 12 weeks before expected delivery, but this is rarely done in practice.

## IV. Data and Empirical Analysis

The data we use in the present article comprise complete longitudinal administrative records on employment and absence for the period 2001–6, merged with information on firms and workers on the basis of encrypted identification numbers.[7] All absence spells are recorded insofar as they are certified by a physician (regardless of recorded symptoms/diagnosis), typically when they exceed 3 days. This implies that the occurrence of very short absence spells is underreported in our data. There is little that we can do about that. To the extent that the removal of pay liability for short-term absenteeism increased the frequency of short absence spells among pregnant workers, we therefore run the risk of underestimating this effect. We return to this issue below. As explained in the previous section, there is also a possibility of self-certifying absence spells as long as 8 days in firms participating in the IWA. Although this option is rarely used in practice (Bergsvik et al. 2010), we deal with the potential difference in certification patterns by always comparing workers who are subject to the same self-certification regulations.

Our empirical analysis consists of two parts. We first examine the extent to which the removal of firms' pay liability for pregnant workers' sick leaves affected these workers absence behavior. We then investigate whether the reform affected young women's employment opportunities. Given that some of the reform effects we seek to evaluate presuppose employers with

[6] These replacement ratios apply for annual earnings up to a ceiling of around 500,000 NOK (in 2012 value, approximately US$85,000).

[7] Individual-level absence data before 2001 are not available.

a forward-looking behavior, an appropriate structural model would contain elements of dynamic programming. Our empirical approach is confined to a more reduced-form setting, however, reflecting the more modest aim of evaluating the empirical relevance of direct (myopic) and indirect (forward-looking) responses to changes in firm incentives.

## A. Absence Behavior during Pregnancies

To examine the impact of employer incentives on absenteeism we construct the following data set. We start out with all employees who became pregnant between May 2001 and May 2005 and did not make a transition to unemployment, rehabilitation benefits, or disability during the pregnancy. We then follow each pregnant employee for 37 weeks through work presence and sickness absence, that is, until she takes maternity leave and thus is no longer at risk of being absent due to sickness.[8] These records constitute our potential treatment group. Given the rather subtle and potentially conflicting ways in which employer incentives were affected by the removal of pay liability, with likely effects on the incidence as well as the duration of absence spells, we set up a bivariate hazard rate model for transitions between the states of presence and absence to identify the reform effects. But before we turn to that model, we take a closer look at what happened at the exact time of the reform implementation by means of a RD analysis. In both these analyses, we also incorporate a control group of nonpregnant workers; that is, for each pregnancy spell, we pick a female nonpregnant worker from exactly the same workplace, at exactly the same point in time, of approximately the same age, and with a similar earnings level (based on a one-to-one nearest neighbor matching procedure).[9] In the bivariate hazard rate model, the control group plays the essential role of representing the counterfactual trend in absence behavior within a DiD modeling framework. In the RD analysis, we primarily use the control group to perform a placebo analysis.

### *1. Descriptive Statistics*

Table 1 and figures 1 and 2 first offer some descriptive statistics. There are 90,898 pregnancy spells included in our data set and an equally large

[8] We assume that the pregnancy started 37 weeks before an observed transition to maternity leave. Transition to maternity leave almost always occurs 3 weeks prior to expected delivery. Only in 1.6% of the cases does the actual delivery take place more than 6 weeks after transition to maternity leave, and these few cases have been deleted from our sample.

[9] We select the coworker with the closest income level, provided that the age difference is less than 3 years. If we cannot find a female coworker with fewer than 3 years of age difference, the pair is not included in our analysis.

**Table 1**
**Descriptive Statistics of Data Used to Analyze Absence Behavior**

| | Treatment Group (Pregnant) | | Control Group (Nonpregnant) | |
|---|---|---|---|---|
| | Before Reform | After Reform | Before Reform | After Reform |
| Number of spells | 20,845 | 70,053 | 20,845 | 70,053 |
| Characteristics (means): | | | | |
| Age | 30.4 | 30.7 | 30.5 | 30.9 |
| Education: | | | | |
| Compulsory or lower secondary | 11.5 | 13.1 | 12.9 | 15.4 |
| Upper secondary | 34.0 | 34.8 | 37.7 | 38.4 |
| College/university | 54.6 | 52.1 | 49.4 | 46.2 |
| Earnings (deflated US$, 2012) | 69,130 | 69,691 | 66,727 | 67,467 |
| Non-European background (%) | 4.8 | 4.6 | 3.8 | 4.4 |

number of controls.[10] While our matching procedure is aimed at making the control group as similar as possible to the treatment group, it is evident that some small compositional differences remain. In particular, the pregnant women tend to earn slightly more than their nonpregnant colleagues. This difference can either be attributed to positive selection into pregnancies or to the strong earnings incentives that pregnant workers face, given that their subsequent maternity leave benefits are calculated on the basis of their earnings in the 10-month period just prior to expected delivery. The "pregnancy wage premium" is almost the same before and after the reform, however, approximately 3.5%.

Figure 1 illustrates that the rate of sickness absence has been fairly stable throughout our observation window for both pregnant and nonpregnant female workers, except for seasonal fluctuations. The timing of the reform is marked in the figure as a vertical line. Based on a visual inspection of the figure, it is not easy to spot any reform effect. A closer comparison of absence rates by pregnancy duration before and after the reform is provided in figure 2. Panel *a* clearly illustrates the sharp rise in absence rates that typically occur as the pregnancy progresses, with absence rates as high as 40%–70% the last few months before transition to maternity leave. Panel *b* reports the changes in absence rates from before to after the reform at different stages of the pregnancy for the members of the treatment group minus the corresponding changes for the members of the control group. According to these descriptive DiD estimators, the reform apparently raised absenteeism somewhat, with a possible exception for the later stages of the pregnancy. The overall absence rate for pregnant women rose

[10] In total, there were 120,089 pregnancies among employed workers in our data period. We lose 29,191 spells (24.3%) due to lack of appropriate controls.

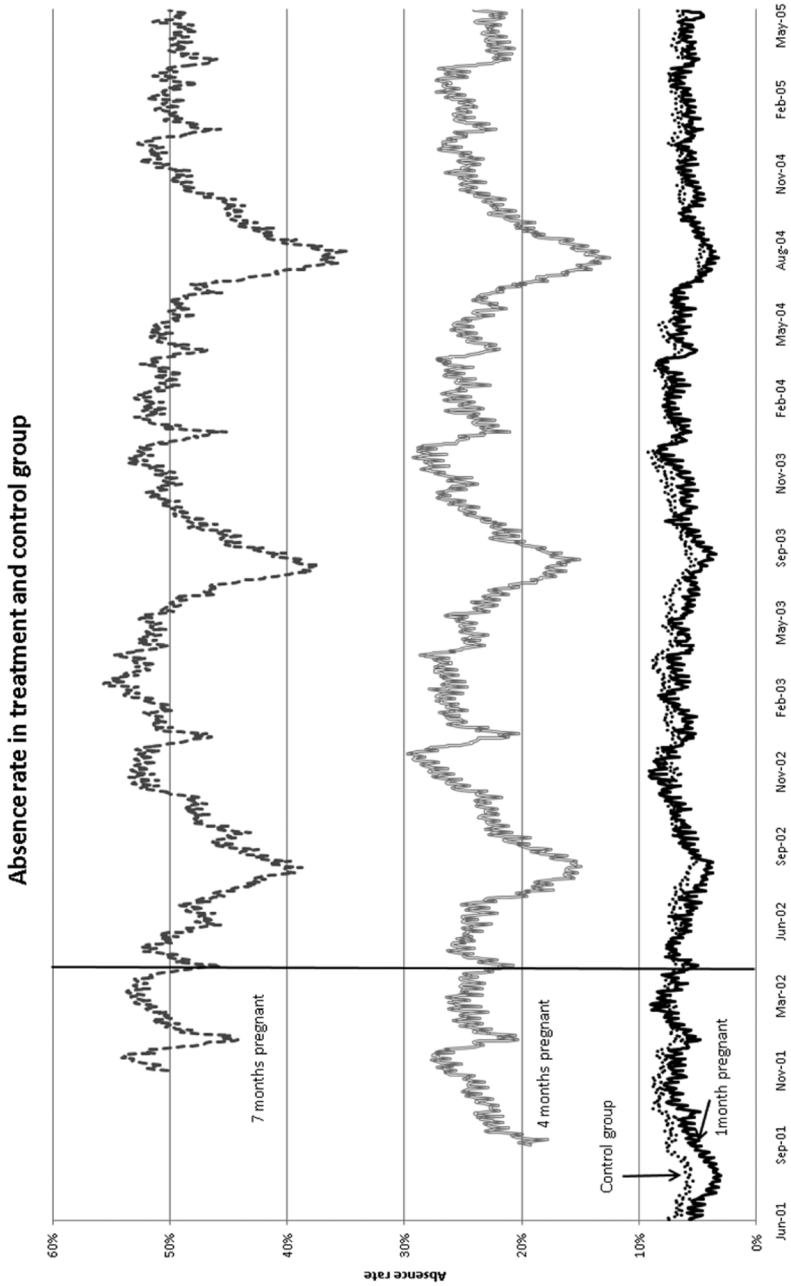



FIG. 1.—Daily absence rates 2001–5 in treatment group (at different stages of the pregnancy) and in control group. Vertical line indicates timing of reform. Daily absence rates are defined as the number of absentees divided by the number of employees at each stage of pregnancy in each of the months displayed.

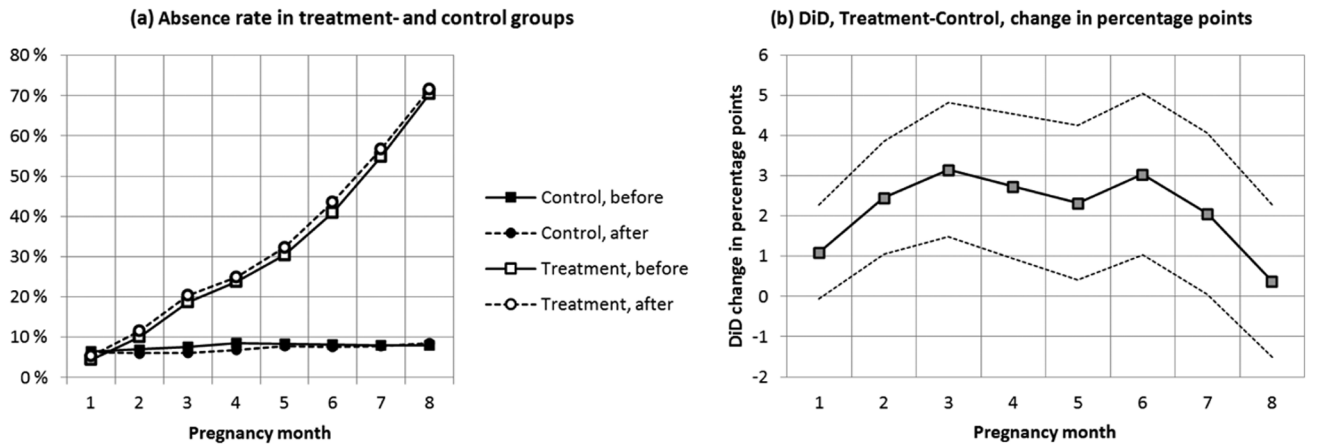



FIG. 2.—Average daily absence rates by number of months into the pregnancy. Treatment and control group before and after the reform (with 95% confidence intervals). For control group members, pregnancy month refers to the number of months since they were picked as controls to a newly pregnant colleague. To prevent seasonal variation from disturbing the before/after reform comparison, the graphs compare workers who became pregnant in May–July 2001 (and their control group members) with those who became pregnant in May–July 2002.

from 32.9% before the reform to 33.8% after the reform, while it dropped from 7.5% to 7.3% for the members of the control group. A simple descriptive DiD estimate for the overall reform effect measured by this outcome is thus a modest increase of 1.1 percentage points, or 3.4% ($100 \times 1.1/32.9$).

### 2. *Regression Discontinuity Analysis*

Figure 3 illustrates the basic idea of a RD analysis. Panel *a* shows how the daily sick leave entry rates among pregnant workers evolved from 90 days before to 90 days after the reform.[11] Due to the seasonal pattern in absence behavior (recall that the reform was implemented on April 1), there is a declining time trend in this period. Imposing a linear calendar time function to control for this development yields a reform effect corresponding to the small upward shift illustrated in the graph. Panel *c* shows how the size of this estimated reform effect varies as we change the bandwidth (the number of days before/after the reform included in the analysis). The shorter the time window, the larger the estimated effects, suggesting (perhaps) that there was a scope for firms/employees to "postpone" the starting date a few days just around the implementation of the reform. The effect is statistically significant, however, regardless of bandwidth. If we rely on the more conservative estimates, that is, those based on the longest time windows, we conclude (from the size of the jump in panel *a*) that the reform raised the daily entry rate to sick leave among pregnant workers by 0.21 percentage points. Since the expected daily entry rate was estimated to be 1.12% just prior to the jump, this corresponds to an increase of 18.8%. As a sort of "placebo test," we examine in panel *e* how the effect estimate changes if we assume alternative false implementation dates. The graph shows that the effect estimates become smaller the further we move away from the true implementation date. Additional placebo tests can be obtained by repeating the whole estimation procedure on the control group. The results are illustrated in panels *b*, *d*, and *f*. The message coming out of these graphs is clear: there is no reform effect whatsoever in the control group.

A natural objection to the regression discontinuity results presented so far is that they may depend heavily on the linearity assumption for calendar time effects. Table 2 shows, however, that the finding of a positive reform effect is highly robust with respect to the specification of the underlying time function. The table presents results for models with from one to four degree polynomials in a time function assumed to be the same before and after the reform and also with from one to four degree polynomials in time functions estimated separately before and after the reform

[11] Daily entry rates are adjusted for weekdays, public holidays, and the first day after holidays.

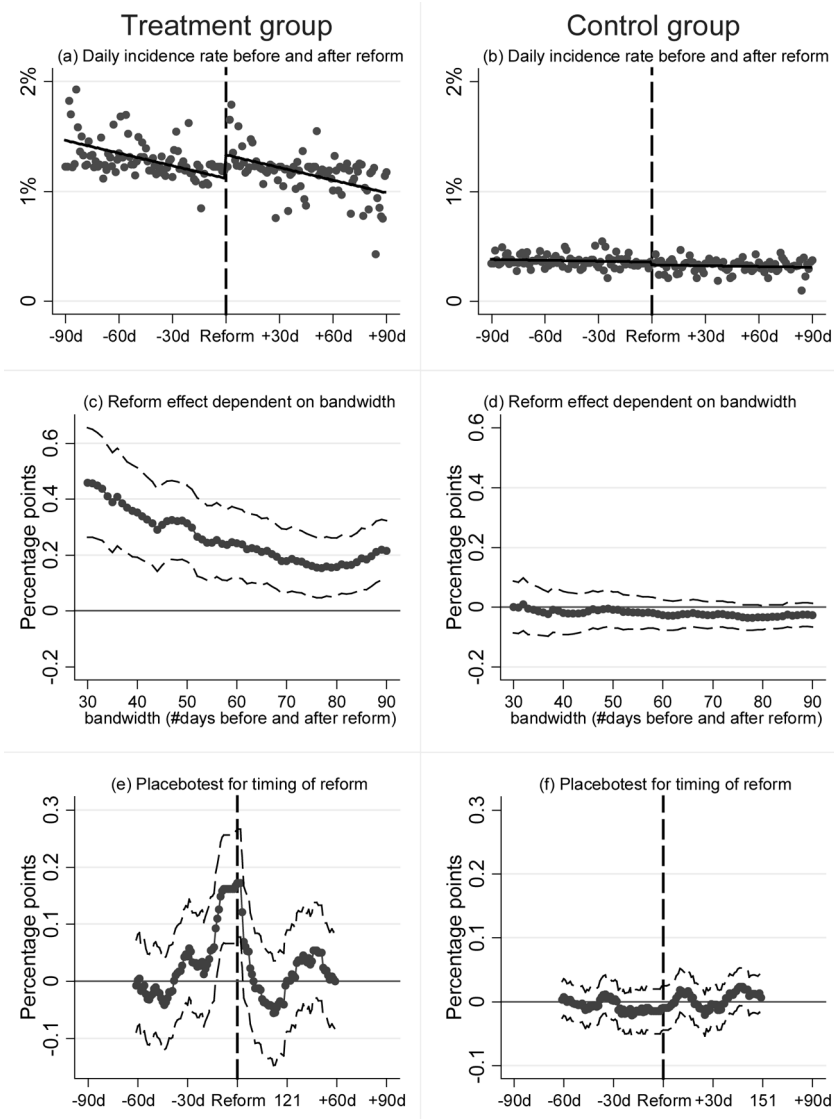



FIG. 3.—Regression discontinuity analysis. Estimated impact on absence incidence of removing firms' pay liability (with 95% confidence intervals). Bandwidth in panels *e* and *f* is 180 days around the true timing of the reform. Reform effects are estimated using a simple linear trend and a reform dummy.

**Table 2**
**Regression Discontinuity Analysis: Estimated Impact on Absence Incidence of Removing Firms' Pay Liability by Bandwidth and Time Function Specification (Percentage Points Change)**

| | Common Time Function | | | | Separate Time Functions Before/After | | | |
|---|---|---|---|---|---|---|---|---|
| Polynomial Degree | 1st | 2nd | 3rd | 4th | 1st | 2nd | 3rd | 4th |
| Treatment group: | | | | | | | | |
| Bandwidth (days): | | | | | | | | |
| [−30, 30] | .46 | .45 | .38 | .37 | .44 | .36 | .57 | .89 |
| | (4.59) | (4.48) | (2.59) | (2.48) | (4.43) | (2.16) | (2.52) | (2.70) |
| [−60, 60] | .24 | .25 | .36 | .35 | .24 | .38 | .54 | .37 |
| | (3.85) | (3.87) | (4.01) | (3.95) | (3.81) | (3.81) | (3.78) | (2.02) |
| [−90, 90] | .21 | .21 | .20 | .20 | .21 | .23 | .56 | .46 |
| | (3.79) | (3.75) | (2.55) | (2.55) | (3.72) | (2.50) | (4.49) | (2.87) |
| Control group: | | | | | | | | |
| Bandwidth (days): | | | | | | | | |
| [−30, 30] | .00 | −.00 | −.02 | −.01 | .00 | −.01 | .11 | .19 |
| | (.01) | (.00) | (−.32) | (−.18) | (.01) | (−.15) | (1.08) | (1.20) |
| [60, 60] | −.03 | −.03 | .00 | .00 | −.03 | .10 | −.00 | .05 |
| | (−1.10) | (-1.08) | (.08) | (.04) | (−1.07) | (.22) | (−.01) | (.62) |
| [−90, 90] | −.03 | −.03 | −.02 | −.02 | −.03 | −.02 | .01 | .01 |
| | (−1.39) | (−1.33) | (−.80) | (−.83) | (−1.31) | (−.56) | (.30) | (.16) |

NOTE.—Values in parentheses are *t*-values.

(in the latter case the reported reform effect is the shift that occurred at the time of reform implementation). All the results are presented for three alternative bandwidths and for both treatments and controls. We interpret the results as convincing evidence that the reform indeed raised the frequency of absence spells among pregnant women. All the 24 alternative specifications produce significant positive reform effects for the treatment group. None of them produce significant results for the control group of nonpregnant colleagues.

Taken at face value, the RD analysis indicates that pregnant workers' entry rate to sick leave was raised 18.8% as a direct result of the reform. If this was the only effect of the reform, the overall level of absenteeism would also increase by 18.8%. This is a huge effect compared to the 3.4% increase suggested by the descriptive DiD calculations referred to above. One possible explanation for this apparent discrepancy is that a higher entry rate was not the only effect but that the reform also triggered significant changes in the duration distribution and recurrence pattern of absence spells and consequently also in the composition of present and absent workers. This is what we examine in the next subsections.

### *3. A Bivariate Hazard Rate Model*

We now turn to a more detailed analysis of the reform effects based on a bivariate hazard rate model where we simultaneously model the transi-

tions to and from sick leave. This simultaneity may be empirically important, since we suspect that the reform affected both transitions, potentially changing the composition of the two risk groups of present and absent workers. Based on economic theory, we hypothesize that four different reform effects may have been at work, all operating through firm incentives:

1. A positive direct effect on absence incidence for workers who had not completed a long-term absence recently (fewer than 16 days ago). This follows directly from the fact that the removal of pay liability made new absence spells less costly for the firm.
2. A negative indirect effect on absence incidence for workers who had completed a long-term absence recently. The reason for this is that the prereform pay liability system entailed a firm incentive to advance any expected (or potential) relapses such that they occurred before a new pay liability period was triggered. The removal of pay liability thus eliminated the incentive to push workers at high risk of relapse into sick leave within 16 days.
3. A negative direct effect on work resumption (positive effect on absenteeism) for absent workers who had been sick-listed fewer than 16 days. This follows directly from the fact that the removal pay liability made absence continuation less costly for the firm within this period.
4. A positive indirect effect on work resumption (negative effect on absenteeism) for absent workers who had been sick-listed more than 16 days. The reason for this is that since the reform removed the pay liability associated with subsequent relapses, it also removed an essential part of the firm's economic risk associated with early work resumption.

An important point to note here is that the indirect effects are not necessarily smaller or less important than the direct effects. We use the term "indirect" to emphasize that the effects in question require an element of forward-looking behavior, not that they are of second order. To the contrary, we expect the indirect work resumption effect (item 4 above) to be particularly significant, as the costs associated with frequent relapses could be very large prior to the reform. Whether the direct or the indirect effects dominate the actual reform responses is an empirical question to which our hazard rate model is designed to provide an answer. Note that the second and third effects listed above are likely to be of minor quantitative importance for overall absenteeism, since they are at work only the first 16 days after or during absence spells, respectively. They are nevertheless of scientific interest, since they convey information about employers' ability to affect the pattern of the employees' absenteeism.

In this exercise, we exploit a much larger time window (2001.5–2005.5) and base our inference on differences in the differences between the treatment and the control groups before and after the reform. Hence, it is the interaction of being pregnant and being in the postreform period that identifies the causal effects of the reform. However, since we are interested in both the direct and indirect effects, the pregnancy-reform interaction needs to be further interacted with variables indicating whether pay liability applied or not in the prereform period, which again depended on the duration of ongoing and past (recent) spells. Since the duration of ongoing and past spells obviously may have autonomous effects on absence behavior in ways that differ between pregnant and nonpregnant workers, we need a full set of controls for these factors to identify the reform effects. Let $P$ be an indicator variable denoting potential treatment (pregnancy), let $R$ be an indicator denoting that the reform has been implemented (after April 1, 2002), let $D$ be an indicator denoting that the current state (presence or absence) has lasted fewer than 16 days, and let $I$ be an indicator denoting that the last absence spell (if any) also lasted fewer than 16 days.[12] Furthermore, let $P_j, j = 9, 8, \ldots, 1$ be indicator variables denoting the number of months until expected delivery for those who are pregnant. We write the hazard rate of moving from presence to absence as

$$\begin{aligned}\theta_{1it} = \exp[&x_{it}\beta_{1k} + v_{1i} + \sigma_{1t} + \delta_{1j}P_j \\ &+ \phi_{11}D + \phi_{12}D \times R + \phi_{13}D \times P \\ &+ \phi_{14}D \times I + \phi_{15}D \times R \times I + \phi_{16}D \times P \times I \\ &+ \alpha_{11}P \times R \times \underline{(1 - D + D \times I)} \\ &+ \alpha_{12}P \times R \times \underline{(D - D \times I)}],\end{aligned} \tag{1}$$

where $\underline{(1 - D + D \times I)}$ is for "pay liability did apply" and $\underline{(D - D \times I)}$ is for "pay liability did not apply." The hazard rate of moving back to presence (work resumption) is

$$\begin{aligned}\theta_{2it} = \exp[&x_{it}\beta_{2k} + v_{2i} + \sigma_{2t} + \delta_{2j}P_j \\ &+ \phi_{21}D + \phi_{22}D \times R + \phi_{23}D \times P \\ &+ \alpha_{21}P \times R \times \underline{D} + \alpha_{22}P \times R \times \underline{(1 - D)}],\end{aligned} \tag{2}$$

where $\underline{D}$ is for "pay liability did apply" and $\underline{(1 - D)}$ is for "pay liability did not apply." In equations (1) and (2), $x_{it}$ is a vector of observed time-

[12] The duration of absence spells is measured inclusive of days belonging to previous absence spells that were terminated fewer than 16 days prior to the start of a current spell.

variant covariates and $(v_{1i}, v_{2i})$ are unobserved person-specific, time-invariant covariates subject to a joint, but unknown, distribution.[13] The coefficients of interest appear in the last rows of these expressions $(\alpha_{11}, \alpha_{12}, \alpha_{22}, \alpha_{21})$. If employer incentives do affect employees' absence behavior, we expect that $\alpha_{11} > 0$ (effect 1 listed above), $\alpha_{12} < 0$ (effect 2), $\alpha_{21} < 0$ (effect 3), and $\alpha_{22} > 0$ (effect 4). The first rows of equations (1) and (2) contain controls for observed and unobserved heterogeneity, for pregnancies (treatment), and for calendar time. Note that calendar time effects $(\sigma_{1t}, \sigma_{2t})$ are estimated separately for each month, implying that any general differences between the prereform and postreform periods are absorbed by these effects. The second and third of rows of equation (1) and the second row of equation (2) contain controls for the factors that determined initial pay liability, and to ensure that it is the interaction of treatment and reform $(P \times R)$ that identifies the causal effects, these determinants are interacted with $P$ and $R$ separately. This implies that we allow the effects of spell duration (fewer than/more than 16 days) and the effects of recent absenteeism (fewer than/more than 16 days) to vary between the treatment and the control groups and between the prereform and postreform periods.

To avoid setting up a model for the initial state, we condition on workers having been present for at least 32 days to start with.[14] To derive the likelihood function for observed data, we split each individual's event history into parts characterized by constant explanatory variables and unchanged state. Let $S_{ki}$, $k = 1, 2$, be the set of observed spell parts under risk of event $k$ (sickness, work-resumption) for individual $i$. Let $l_{kis}$ denote the observed length of each of the spell parts, and let the indicator variables $y_{kis}$ denote whether a spell part at risk of transition $k$ actually ended in such a transition or was right-censored. Conditional on unobserved heterogeneity, the likelihood function for individual $i$ can then be written as

$$\begin{aligned} L(v_{1i}, v_{2i}) = & \prod_{s \in S_{1i}} [\theta_{1it}(v_{1i})]^{y_{1is}} \exp\{-l_{1is}[\theta_{1it}(v_{1i})]\} \\ & \times \prod_{s \in S_{2i}} [\theta_{2it}(v_{2i})]^{y_{2is}} \exp\{-l_{2is}[\theta_{2it}(v_{2i})]\}. \end{aligned} \tag{3}$$

[13] Note that if a person is observed with multiple pregnancies during our observation window, she is treated as multiple persons (with a separate unobserved heterogeneity vector for each pregnancy). The same practice applies if a control person is matched to more than one pregnant colleague. The vector of observed covariates includes age (27 dummy variables, corresponding to ages 19–45), calendar month (56 dummy variables), county (19 dummy variables), income (15 dummy variables), education/industry (15 dummy variables), and the local unemployment rate.

[14] This restriction implies that we drop 2,245 treatment-control pairs (2.5%).

Since the likelihood contribution in equation (3) contains unobserved variables, it cannot be used directly for estimation purposes. This problem may be solved by formulating a model for the distribution of unobserved variables and then replacing equation (3) with its expectation. In the present context, it is probable that the unobserved characteristic affecting absence incidence is correlated to that affecting work resumption, implying that manipulation of the inflow is also likely to change the distribution of unobserved work-resumption propensities in the stock of absentees (and vice versa). We thus model unobserved heterogeneity nonparametrically. Under mild technical assumptions, results in Lindsay (1983, theorem 3.1) and Heckman and Singer (1984, theorem 3.5) ensure that for the purpose of maximizing the likelihood, unobserved heterogeneity may be approximated by a discrete distribution with a finite number of support points. Since only a subset of the analysis population actually experiences an absence spell within our observation window, there is potentially a selection problem involved in estimating the work-resumption hazard. This is similar to the problem arising in the timing-of-events evaluation literature when treatment effects are heterogeneous and only a selected subset of potential participants actually receives treatment. Based on a set of regularity conditions, Abbring and Van den Berg (2003, proposition 4) prove nonparametric identification of the multivariate mixed proportional hazard rate model in this context. To ensure identification in our case, we have performed a Monte Carlo analysis based on artificial data of similar sample sizes and with similar selection challenges that we have in the actual application, that is, data where only a selected portion of the sample comes under the risk of one of the hazard rates during the observation window and with a significant correlation between the two unobserved covariates. The exercise is documented in the appendix, available online. It demonstrates that the sequential hazard rate model we use is indeed identified and that it can be accurately estimated with the algorithm used in this article (with correct standard errors). In our actual application, the foundation for identification is further strengthened by the presence of time-varying covariates as reflected in independent calendar time variation (see McCall 1994; Brinch 2007; Gaure, Røed, and Zhang 2007).

The causal effects of the reform are identified by the shift in pregnant workers' absence behavior from before the reform to after the reform, relative to that of the control group members. An important assumption underlying this identification strategy is that the calendar time effects are the same for the treatment and the control groups. If pregnant workers have been subject to different time trends than nonpregnant workers for reasons that are not related to the reform, the estimated reform effects may be biased. Since we have selected control workers from the pregnant workers' own workplaces, the reform effects may also be biased if colleagues affect each other's absenteeism. Existing evidence indicates that peer ef-

fects in absenteeism are empirically relevant (see, e.g., Ichino and Maggi 2000; Bradley, Green, and Leeves 2007; Hesselius, Nilsson, and Johansson 2009).[15] We return to a number of robustness exercises below, with respect to the composition of the control group, with respect to the assumption of a common trend, and with respect to the modeling of unobserved heterogeneity.

Let $Q$ be the (a priori unknown) number of support points in the unobserved heterogeneity distribution and let $\{(v_{1l}, v_{2l}), p_l\}$, $l = 1, 2, \ldots Q$, be the associated location vectors and probabilities. In terms of observed variables, we write the likelihood function as

$$L = \prod_{i=1}^{N} E_{v_i}[L_i(v_{1i}, v_{2i})] = \prod_{i=1}^{N} \sum_{l=1}^{Q} p_l L_i(v_{1l}, v_{2l}), \quad \sum_{l=1}^{Q} p_l = 1, \quad (4)$$

where $L_i(v_{1i}, v_{2i})$ is given in equation (3). Our estimation procedure is to maximize the likelihood function (4) with respect to all model and heterogeneity parameters repeatedly for alternative values of $Q$. The nonparametric maximum likelihood estimators (NPMLE) are obtained by starting out with $Q = 1$ and then expanding the model with new support points until the model is "saturated" in the sense that we are no longer able to increase the likelihood function by adding more points. The preferred model is then selected on the basis of the Akaike Information Criterion (AIC). Monte Carlo evidence presented in Gaure et al. (2007) indicates that parameter estimates obtained this way are consistent and approximately normally distributed. They also indicate that the standard errors conditional on the optimal number of support points are valid for the unconditional model as well and hence can be used for standard inference purposes.[16]

Table 3 presents the estimated reform effects and the effects of the key control variables.[17] The two direct effects of the reform contributed to higher absenteeism: the rate of entry into sick leave spells that used to be subject to pay liability rose by around 10% ($\alpha_{11}$), while the work-resumption rate during the first 16 days of the spells dropped by around 6% ($\alpha_{21}$). The two indirect effects pulled in the other direction, however: the rate of entry into sick leave spells that were exempted from pay liability already before the reform dropped by as much as 19% ($\alpha_{12}$), whereas

[15] It is also possible that there are peer effects in fertility decisions (see Hensvik and Nilsson 2010). Note, however, that we have removed from the control group all women who themselves become pregnant during the treatment group's pregnancies.

[16] Note that the standard errors do not take any within-group correlation in absenteeism into account, which may be present due to our sampling of control persons within firms. We perform a robustness exercise below, where we have sampled control persons from other firms instead.

[17] A complete listing of estimated coefficients is available from the authors.

**Table 3**
**Selected Estimation Results from Hazard Rate Model**

| Reform Effect | Coefficient |
|---|---|
| Reform effect on incidence with initial pay liability ($\alpha_{11}$) | .100*** |
| | (.016) |
| Reform effect on incidence without initial pay liability ($\alpha_{12}$) | −.215** |
| | (.119) |
| Reform effect on work resumption with initial pay liability ($\alpha_{21}$) | −.057** |
| | (.024) |
| Reform effect on work-resumption without initial pay liability ($\alpha_{22}$) | .123*** |
| | (.034) |
| Effects of duration and past absence: | |
| Effect on incidence of current presence spell lasted fewer than 16 days: | |
| For all($\phi_{11}$) | 2.096*** |
| | (.035) |
| Interacted with reform ($\phi_{12}$) | .017 |
| | (.030) |
| Interacted with pregnancy ($\phi_{13}$) | −1.259*** |
| | (.026) |
| Additional effect when last absence spell lasted fewer than 16 days: | |
| For all ($\phi_{14}$) | .420*** |
| | (.104) |
| Interacted with reform ($\phi_{15}$) | −.147 |
| | (.111) |
| Interacted with pregnancy ($\phi_{16}$) | .470*** |
| | (.113) |
| Effect on work resumption of current absence spell lasted fewer than 16 days: | |
| For all ($\phi_{21}$) | .978*** |
| | (.031) |
| Interacted with reform ($\phi_{22}$) | .108*** |
| | (.032) |
| Interacted with pregnancy ($\phi_{23}$) | .253*** |
| | (.037) |

NOTE.—Number of observations = 177,306. Standard errors are in parentheses. Additional controls include age (34 dummy variables), calendar time (64 dummy variables), county (19 dummy variables), income (15 dummy variables), education/industry (15 dummy variables), and local unemployment (in the municipality). The model also includes six support points for the two-dimensional unobserved heterogeneity distribution and the estimated correlation coefficient between $\exp(v_{1i})$ and $\exp(v_{2i})$ is .30. In total, the model contains 348 estimated parameters.

** Significant at the 5% level.

*** Significant at the 1% level.

the work-resumption hazard from long-term absences rose by 13% ($\alpha_{22}$).[18] All these results are exactly as one would expect on the basis of economic theory. Firms responded to the new incentives by reducing efforts to prevent short-term absence, for which they no longer faced any direct costs. At the same time, they apparently became less skeptical toward allowing

[18] The estimated percentage shift in the hazard rates caused by a given variable is given as $100 \times (\exp(\text{estimate}) - 1)$, which for small estimates is approximately equal to the estimate multiplied by 100.

long-term absentees back into work and became more willing to let "risky" presence spells exceed the limit of 16 days, knowing that they no longer were responsible for the sick pay costs associated with relapses.

In order to examine the quantitative importance of short-term pay liability, we have performed a simulation exercise based on the estimated model; that is, we have simulated a large number of spells for typical pregnant workers with and without the direct and indirect effects reported in table 3 "turned on."[19] When we only allow the two direct effects ($\alpha_{11}, \alpha_{21}$) to enter the model (i.e., disregard the indirect effects), it predicts that pay liability reduces the overall expected number of absence days during pregnancies from 85.5 to 79.3, or by 7.3%. However, when we also allow the two indirect effects ($\alpha_{12}, \alpha_{22}$) to enter the model, the predicted number of absence days again rise to 82.8, reducing the overall effect of pay liability to a 3.2% reduction. It is instructive to decompose the predicted effects of short-term pay liability into absence days subject to and absence days exempted from pay liability. Based on our simulations, we find that while pay liability for short-term absences reduces the number of absence days directly covered by pay liability by around 6.8%, it reduces the number of noncovered days by 3.5% only. If we extrapolate our findings to nonpregnant workers, we find that pay liability on short-term absenteeism reduces short-term absenteeism by as much as 11.1%, while it actually increases long-term absenteeism by around 1.6%. The reason why we predict that the effects of pay liability are relatively more important for nonpregnant than for pregnant workers is that a disproportionally large fraction of pregnant workers are "stuck" in long-term absenteeism anyway, with return-to-work rates return so low that the negative proportional shift caused by pay liability ($\alpha_{22}$) is of moderate quantitative importance. Consequently, they also become removed from the risk of incidence.

It is important to bear in mind that all the effect estimates reported here are likely to underestimate the true causal effect of pay liability for the reason that the reform did not really affect all absences among pregnant workers—only those that were "pregnancy related." In Section IV.A.3 above, we estimated that only 73% of pregnant workers' absence spells are pregnancy related, implying that our coefficients are estimated with a significant attenuation bias.

### *4. Robustness*

As discussed above, the results in table 3 are based on the assumption that young, nonpregnant, female colleagues constitute an appropriate con-

[19] Simulations were made for an agent starting a pregnancy with "representative" entry and exit rates, based on the point estimates reported in table 3. We made 177,176 simulations for each set of assumptions regarding the reform effects (corresponding to the actual number of observations).

trol group for pregnant workers, in the sense that they were subject to the same time variation in absence behavior apart from the reform and also were not affected by the reform through peer effects. Yet, while it is conceivable that violations of these assumptions could disturb attempts to identify overall reform effects, it is hard to see why they should yield the particular twist in the work-resumption profiles captured by our estimates. We interpret the finding that the work-resumption rate declined during the first 16 days of absence (for which there used to be a pay liability period) while at the same time it rose at longer durations as convincing evidence of a reform effect.

Table 4 reports the results from a number of robustness exercises. We first examine sensitivity with respect to the common trend assumption. Column 2 reports the estimated reform effects when the time window is reduced to include 22 months only (dropping the last 42 months), while columns 3 and 4 report the results when instead a separate linear or quadratic time trend for pregnant workers is included in the model, respectively. While these modifications do change the estimated effects somewhat—in the direction of weaker direct effects and (in some cases) stronger indirect effects—they do not alter any of the main conclusions. The sensitivity of the point estimates with respect to the common trend assumption may suggest, however, that there was an increasing trend in pregnancy-related absences for reasons not related to the reform. Alternatively, it may have been the case that the full reform effects materialized gradually, as it took some time before employers learned about the new rules. Second, we assess the potential problem that the absence behavior in the control group is causally affected by the absence behavior in the treatment group. Since our control persons are chosen from the same workplaces as the treatment population (to ensure maximum similarity in terms of shocks to the work environment), such a causal relationship could arise due to peer effects or changes in work pressure resulting from colleagues' absence. Column 5 reports the results from a model where we have drawn the control group members from other workplaces instead; that is, for each treatment, we have selected a control person from a different firm within the same industry and with the same IWA status but have otherwise followed the same one-to-one nearest neighbor matching strategy as described above.[20] It is clear that the estimated direct incidence effect as well as the indirect effect on work resumption then become larger. This pattern is consistent with the existence of a significant peer effect, in line with existing empirical evidence. Again, the main conclusions go through, however.

In our baseline model, we have imposed rather restrictive assumptions regarding duration dependence in the return-to-work hazard by allowing

[20] Industries are categorized on the basis of the International Standard Industrial Classification (ISIC) at a three-digit level.

**Table 4**
**Robustness: Estimated Reform Effects under Alternative Modeling Assumptions**

| | Baseline Model (Repeated from Table 3) (1) | Reduced Time Window (2001.5–2003.3) (2) | Separate Linear Time Trend for Pregnant (3) | Separate Quadratic Time Trend for Pregnant (4) | Control Group Selected from Other Workplaces (5) | With Flexible Duration Baseline (Nine Steps) (6) | Without Unobserved Heterogeneity (7) | Absence-Specific Unobserved Heterogeneity (8) |
|---|---|---|---|---|---|---|---|---|
| Reform effect on incidence with initial pay liability ($\alpha_{11}$) | .100*** | .082*** | .042** | .066** | .141*** | .100*** | .096*** | .105*** |
| | (.016) | (.019) | (.020) | (.029) | (.015) | (.015) | (.014) | (.018) |
| Reform effect on incidence without initial pay liability ($\alpha_{12}$) | −.215** | −.226 | −.276*** | −.256** | −.094 | −.224** | −.221** | −.268** |
| | (.119) | (.143) | (.117) | (.121) | (.112) | (.117) | (.102) | (.103) |
| Reform effect on work-resumption with initial pay liability ($\alpha_{21}$) | −.057** | −.031 | −.049* | .034 | −.020 | −.076*** | −.058** | −.048** |
| | (.024) | (.028) | (.029) | (.039) | (.022) | (.024) | (.023) | (.023) |
| Reform effect on work-resumption without initial pay liability ($\alpha_{22}$) | .123*** | .107** | .133*** | .208*** | .174*** | .109*** | .124*** | .106*** |
| | (.034) | (.041) | (.038) | (.045) | (.031) | (.034) | (.033) | (.033) |
| Number of support points in unobserved heterogeneity distribution | 6 | 4 | 5 | 3 | 2 | 5 | 1 | 10 |

NOTE.—Standard errors are in parentheses.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

only a two-step function in the form of a shift after 16 days of absence. To assess whether this may have influenced our results, we report in column 6 estimates based on a much more flexible nine-step function (allowing a shift every thirtieth day). Again, we find that the results are robust, although we estimate a somewhat stronger reform response during the first 16 days and correspondingly weaker response afterward. Finally, we examine robustness with respect to the modeling of unobserved heterogeneity. Column 7 reports the results from a model without unobserved heterogeneity, while column 8 reports the results from a model where the unobserved heterogeneity vector is allowed to change each time a person has returned to work from a completed absence spell (and been at work for at least 16 days). The latter specification is included to allow distinct work-absence cycles to be associated with different unobserved transition propensities, generated, for example, by different diseases. The results change only marginally, indicating that the exact way in which unobserved heterogeneity is modeled is empirically unimportant in this case.

### *5. Comparison with RD Findings and Descriptive Statistics*

While the purely descriptive DiD estimate for the overall reform effect was a 3.4% rise in absenteeism (see Sec. IV.A.1), the simulation exercise based on the hazard rate model indicates a rise around 3.2%. These two very different empirical approaches thus yield more or less the same conclusion regarding the overall effects of the reform. The RD analysis, on the other hand, indicated a shift of the incidence rate alone of 18.8%. This is not necessarily in conflict with the results from the DiD and the hazard rate approaches, since the latter also incorporate effects on spell duration. But even for the incidence effect, the hazard rate model and the RD analysis seem to come up with different results. While the RD analysis indicated a shift in the incidence rate 18.8%, the hazard rate model indicates a shift around 10%. This difference can be explained, however, by endogenous sorting into the risk set of present workers, which is appropriately accounted for in the bivariate hazard rate model but not in the RD analysis. To show this, we reestimated the hazard rate model without including controls for pregnancy-month (and other individual characteristics) in the model, making it more directly comparable to the RD analysis (not shown in tables). We then obtained an entry effect ($\alpha_{11}$) of 0.18 (SE = .01), which is basically in line with the RD estimates with the longest bandwidths. Hence, what apparently happened was that the reform caused entry rates to rise for two very different reasons: (i) because transition to absenteeism for each present worker became less costly for the firm and (ii) because the composition of the population of present workers changed toward individuals with higher absence risks (as long-term absentees to a larger extent returned to

work). While the bivariate hazard rate model is designed to disentangle these two effects, the RD analysis lumps them together. This illustrates a potential limitation associated with using RD in settings were the source of the discontinuity affects behavior such that the composition of the agents under study also changes; the method may be a reliable and robust strategy for ascertaining that behavior really has changed yet be a poor strategy for quantifying these effects at the level of agents.

### B. Employment Opportunities for Young Females

The firms' pay liability for pregnant workers' absences was removed for the purpose of making individuals conceived to have a high risk of becoming pregnant more employable. In this subsection, we evaluate whether the reform had this intended effect or not. Employers' scope for discriminating against pregnant or pregnancy-prone workers is limited insofar as workers already are employed. Hence, to the extent that the higher expected absence costs associated with pregnancies affect employment opportunities at all, they are likely to do so primarily through the hiring decision. We therefore start out with the population of labor-market entrants, that is, persons who completed their educational career (upper secondary or higher) in the period from 2000 to 2004, and we investigate transitions to a first job.[21] It is not obvious how a treatment group should be defined in this case, and it is even less obvious how an appropriate control group can be established. Loosely speaking, the treatment group consists of women at risk of becoming pregnant, which from the employers' point of view basically includes all young women. In a more narrow sense, it consists of already-pregnant job seekers. As our primary strategy, we thus define all young, female labor-market entrants (aged 19–34) to be in the treatment group, whereas we use men in the same age group as controls. In addition, we examine in particular the employment prospects of pregnant graduates.

We first show some descriptive statistics illustrating the school-to-work transitions for males and females prior to and after the reform. We then set up a multivariate regression model accounting for the transition to the first job and investigate whether the reform coincided with a relative improvement in employment chances for female labor-market entrants in general and for pregnant entrants in particular.

#### *1. Descriptive Statistics*

Descriptive statistics for the treatment and control groups are provided in figure 4 and in table 5. We focus on groups who either completed their

[21] A person is considered to have completed education in a particular quarter if he or she studied in that quarter but was not registered in any formal education the subsequent two quarters.

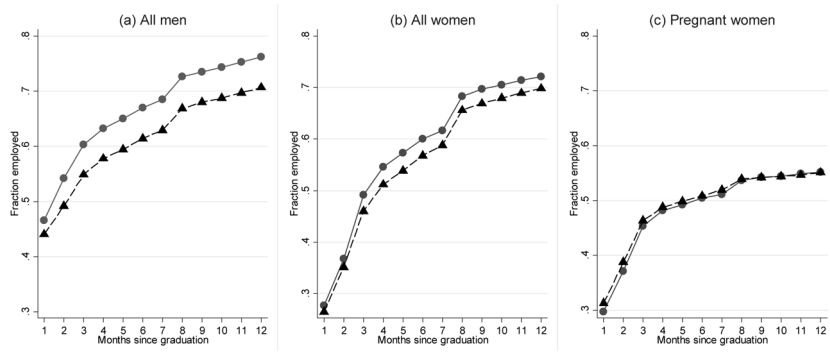



FIG. 4.—Cumulative employment fractions by months since graduation. Before reform, solid gray line with circles; after reform, dashed black line with triangles.

education well before (June 2000–March 2001) or well after (June 2002–March 2003) the reform to avoid cohorts that were only partly affected by the reform. Looking first at the cumulative employment propensities by time since graduation in figure 4, we note that both women and men experienced a significant drop in employment from the former to the latter period but that the drop was largest for men, implying a relative improvement for women (see panels *a* and *b*).[22] Moreover, women who were pregnant at the time of graduation did not experience a drop in employment propensity at all, thus experiencing a quite substantial relative improvement. The descriptive DiD estimator for the reform's effect on young women's overall probability of having found some form of employment within 6 or 12 months after graduation is 2.4 and 3.2 percentage points, respectively (see the numbers reported at the bottom of table 5).

Table 5 also reveals that men and women tend to chose very different educations, as reflected in the much higher level of college and university education for women than for men. In addition, it reveals that a much larger fraction of men than of women has obtained employment well before the time of graduation. An important reason for this is that many secondary educations involve the possibility of a final apprentice year, which implies employment.

### 2. *A Probability Model for Subsequent Employment*

Given the large differences in men's and women's educational choices, it is also probable that their school-to-work transitions have been subject to correspondingly different cyclical and seasonal fluctuations around the time of the reform. To avoid this from biasing our results, we set up a

[22] Employment is defined on the basis of the Norwegian Employer-Employee register. A person is defined to have been employed in a particular month if he or she had a registered job paying a monthly salary of at least NOK 20,500 (2012 value).

**Table 5**
**Descriptive Statistics of Data Used to Analyze Transitions to Employment**

| | Treatment: All Young Women | | Control: All Young Men | |
|---|---|---|---|---|
| | Before June 2000–March 2001 | After June 2002–March 2003 | Before June 20000–March 200101 | After June 2002–March 2003 |
| Number of graduates | 19,589 | 18,843 | 19,674 | 18,462 |
| Age | 25.5 | 25.3 | 25.1 | 24.7 |
| Education | | | | |
| Upper secondary | 44.7 | 46.8 | 64.0 | 65.7 |
| College (low level) | 43.6 | 41.0 | 22.2 | 20.7 |
| University (high level) | 11.7 | 12.2 | 13.9 | 13.6 |
| Non-European background (%) | 5.2 | 6.1 | 5.4 | 5.7 |
| Cumulative fraction employed (%): | | | | |
| 6 months before graduation | 14.7 | 13.9 | 35.9 | 33.2 |
| At graduation | 19.6 | 19.4 | 39.9 | 38.3 |
| 6 months after | 60.0 | 56.8 | 67.0 | 61.4 |
| 12 months after | 72.1 | 69.8 | 76.2 | 70.7 |

multivariate regression model for the school-to-work transition, where we control for all possible combinations of education types and graduation times; that is, we define 49 different education categories and interact them with dummy variables for the before/after reform period.[23]

We focus on the two discrete outcomes of having obtained some form of employment within 6 or 12 months after graduation. For ease of interpretation, we specify the school-to-work transition with a linear probability model, that is,

$$E[y_i] = x_i\beta_3 + \phi_{31}\text{EDUC}_i + \phi_{32}\text{EDUC}_i \times R + \phi_{33}W + \phi_{34}P + \alpha_{31}W \times R + \alpha_{32}P \times R, \quad (5)$$

where $y_i$ is an indicator for employment within 6 or 12 months, $W$ is a dummy for women, and $\text{EDUC}_i$ is a vector with 49 education dummy variables. The pregnancy dummy $P$ is equal to one if a woman was pregnant at the time of graduation or within the first 3 months after graduation. The covariate vector $x_i$ includes the month of graduation (17 dummy variables), age (16 dummy variables), and the local (municipality) rate of youth unemployment (aged 19–34) at the time of graduation. The coefficients of main interest are $\alpha_{31}$ and $\alpha_{32}$.

[23] Educations are classified on the basis of the International Standard Classification of Education (ISCED97).

The main regression results are presented in columns 1 and 2 of table 6. For the 12-month outcome, we find a small positive reform effect on young women's employment propensity equal to 1.6 percentage points. The effect is statistically different from zero at the 5% level. For women who were pregnant at the time of graduation, we find an additional effect of 2.9 percentage points, significant at the 10% level. The effects are imprecisely estimated, however. When we focus on a 6-month period rather than a 12-month period, the estimated effect for women as a whole declines (and become statistically insignificant), whereas the additional effect for pregnant women becomes stronger.

As shown in table 5, many students—particularly men—have obtained some employment long before graduation. Since it is difficult to interpret the school-to-work transition for these graduates, we have also estimated the model without including those who had a job as long as 6 months before graduation. The results from this exercise are displayed in columns 3 and 4 of table 6. While the reform effects for women as a whole are virtually unaffected by this sample restriction, the estimated effects for pregnant women become somewhat smaller. To sum up, it seems that the removal of pay liability for pregnancy-related absences had the intended

**Table 6**
**Estimated Reform Effects on Employment Propensity**

| | Whole Sample | | Reduced Sample: No Employment 6 Months before Graduation | |
|---|---|---|---|---|
| | Employment within 12 Months (1) | Employment within 6 Months (2) | Employment within 12 Months (3) | Employment within 6 Months (4) |
| Effect for all women ($\alpha_{31}$) | 1.59** | 1.00 | 1.62* | 1.08 |
| | (.70) | (.76) | (.84) | (.89) |
| Effect for pregnant ($\alpha_{32}$) | 2.84* | 4.13** | 2.03 | 3.50* |
| | (1.51) | (1.65) | (1.80) | (1.91) |
| Women ($\phi_{33}$) | −.35 | −3.00*** | 2.29*** | .02 |
| | (.49) | (.53) | (.59) | (.63) |
| Pregnant ($\phi_{34}$) | −19.87*** | −12.09*** | −22.48*** | −13.98*** |
| | (1.04) | (1.14) | (1.24) | (1.32) |
| Local youth unemployment rate (ages 19–34) | −6.89*** | −7.65*** | −6.78*** | −7.24*** |
| | (.05) | (.06) | (.06) | (.07) |
| Observations | 76,568 | 76,568 | 57,864 | 57,864 |

NOTE.—Additional controls include age (16 dummy variables), graduation month/year (17 dummy variables), and education type/level before/after reform (97 dummy variables).
* Significant at the the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

effect of making young women more employable. Our best guess is that it raised the probability of having experienced some form of employment 1 year after graduation by 1–2 percentage points for women in general and by 4–6 percentage points for pregnant graduates. To put these estimates into perspective, it may be of interest to examine the sizes of the expected sick pay costs that were effectively removed by the reform. Based on the simulation exercise described in the previous subsection, we compute that being pregnant implies approximately 12 extra sick pay days directly covered by pay liability. The reform's removal of this cost corresponds to a 5% reduction of the total wage bill during a typical pregnancy. On average, female graduates aged 23–27 experienced 0.4 pregnancies during the first 3-year period after graduation. Hence, the reform did have a small, but noticeable, effect on expected wage costs for new female employees.

## V. Conclusion

The findings reported in this article show that employees' sick leave behavior is responsive toward the employer's wage costs during their absence. If a significant part of the costs can be passed on to a public insurer, employees tend to be absent more than if the costs are paid for by their own employer. In most countries, the employer is responsible for the costs associated with short-term absence, while the public insurer covers the direct sick pay costs arising from long-term absence. We have shown that an insurance system with these properties reduces the incidence of absence spells and also raises the probability of quick work resumption. At the same time, however, it also reduces the work-resumption rate at durations exceeding the pay liability period. We conclude that responsibility for short-term sick pay only undermines the firms' incentives to prevent long-term absenteeism, not only because they have too little pecuniary incentive to avoid long-term absence per se but also because a long-term absent worker's return to work entails the risk of costly short-term relapses (for which the employer is again financially responsible). As a result, employers may exert too little effort to prevent long-term absence. Although we must be careful generalizing the quantitative results obtained for pregnant workers to workers at large, we see no reason why the phenomenon discovered in this article should be restricted to pregnant employees. The evidence presented in this article then indicates that the unintended side effect of restricting pay liability to short-term absence spells only—in terms of raising longer-term absenteeism—may be sufficiently large to almost nullify the favorable impacts of a 16-day pay liability period for short-term absenteeism. Extrapolating our findings to the work force at large, we estimate that by making the employers responsible for around one-third of overall sick pay costs in Norway, policy makers achieve a mere 2.7% reduction in overall absence and even a slight increase in long-term absence.

Given that long-term absence is the typical gateway to permanent disability benefits (which are fully covered by taxpayers), insufficient employer incentives to prevent long-term absenteeism may be very costly from a social point of view.

Why, then, have so many countries designed their sickness insurance systems such that employers do not face the direct sick pay costs associated with long-term absenteeism? We see two possible explanations. The first is that absence behavior has typically been considered to be determined by the worker—not the firm. Hence, the focus has been placed on worker incentives rather than firm incentives. The results presented in this article show that this argument is not valid; firms do influence their employees' absence behavior, and they respond to economic incentives. Second, there may be significant administrative costs associated with reimbursing firms for the large number of short-term absences, implying that even a modest pay liability for longer absence spells is difficult to implement without at the same time raising the overall sick leave costs for firms. While this may be a desirable outcome from the perspective of absence prevention, it has been argued that it also undermines employment prospects for job seekers with high expected absence propensity. Our findings indicate that, to some extent, this argument is valid; firms do respond to sick pay liability by being less willing to employ workers expected to be absent a lot. Hence, by raising pay liability for long-term absences without at the same time reducing it for shorter spells, policy makers may undermine incentives to employ marginal workers. A possible solution to this dilemma may be to restructure firms' pay liability so that firms cover a smaller fraction of the short-term absence costs and a larger fraction of the long-term absence costs. This solution requires, however, that the administrative challenges associated with high numbers of reimbursement transactions can be overcome.

We conclude this article with a caveat. Since our data include physician-certified absences only, we obviously cannot evaluate the impacts of pay liability on self-reported absenteeism. While we see no particular reason to expect significant differences in the responses of self-certified and physician-certified sick leave to changes in firm incentives, we cannot rule out that such differences exist.

## References

Abbring, Jaap H., and Gerard J. Van den Berg. 2003. The nonparametric identification of treatment effects in duration models. *Econometrica* 71, no. 5:1491–1517.

Bergsvik, Daniel, Simen Markussen, and Oddbjørn Raaum. 2010. Tok en tredagers? Mønstre i egenmeldt sykefravær (Three days off? Patterns of self-reported absenteeism). *Søkelys på Arbeidslivet* 27, no. 4:379–95.

Bjerkedal, Tor, and Ola Thune. 2003. Hva koster sykelønnsordningen? (Total cost of sick leave benefits: The case of Norway). *Tidsskrift for den Norske Lægeforening* 123 (May): 662–63.

Bonato, Lusine, and Leo Lusinyan. 2007. Work absence in Europe. *IMF Staff Papers* 54, no. 3:475–538.

Bradley, Steve, Colin Green, and Gareth Leeves. 2007. Worker absence and shirking: Evidence from matched teacher-school data. *Labour Economics* 14, no. 3:319–34.

Brinch, Christian. 2007. Nonparametric identification of the mixed hazards model with time varying covariates. *Econometric Theory* 23, no. 2: 349–54.

Burkhauser, Richard V., and Mary C. Daly. 2011. *The declining work and welfare of people with disabilities: What went wrong and a strategy for change*. Washington, DC: AEI.

Burkhauser, Richard V., Maximilian D. Schmeiser, and Robert R. Weathers. 2012. The importance of anti-discrimination and workers' compensation laws on the provision of workplace accommodations following the onset of a disability. *Industrial and Labor Relations Review* 65, no. 1: 350–76.

Cahuc, Pierre, and André Zylberberg. 2004. *Labor economics*. Cambridge, MA: MIT Press.

D'Amuri, Francesco. 2011. Monitoring, monetary incentives, and workers' rents in determining absenteeism. Working paper, Research Department, Italian Central Bank. http://sites.google.com/site/fradamuri/home/research.

Edwards, Paul, and Kay Greasley. 2010. *Absence from work*. Report from the European Foundation for the Improvement of Living and Working Conditions. Dublin: European Working Conditions Observatory.

Gaure, Simen, Knut Røed, and Tao Zhang. 2007. Time and causality: A Monte Carlo evaluation of the timing-of-events approach. *Journal of Econometrics* 141, no. 2:1159–95.

Gruber, Jonathan. 1994. The incidence of mandated maternity benefits. *American Economic Review* 84, no. 3:622–41.

Heckman, James, and Burton Singer. 1984. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica* 52, no. 2:271–320.

Henrekson, Magnus, and Mats Persson. 2004. The effects on sick leave of changes in the sickness insurance system. *Journal of Labor Economics* 22, no. 1:87–113.

Hensvik, Lena, and Peter Nilsson. 2010. Businesses, buddies and babies: Social ties and fertility at work. IFAU Working Paper 2010:9, Institute for Evaluation of Labour Market and Education Policy (IFAU), Uppsala, Sweden.

Hesselius, Patrik, J. Peter Nilsson, and Per Johansson. 2009. Sick of your colleagues' absence? *Journal of the European Economic Association* 7, nos. 2–3:583–94.

Ichino, Andrea, and Giovanni Maggi. 2000. Work environment and individual background: Explaining regional shirking differentials in a large Italian firm. *Quarterly Journal of Economics* 115, no. 3:1057–90.

Johansson, Per, and Mårten Palme. 2005. Moral hazard and sickness insurance. *Journal of Public Economics* 89, nos. 9–10:1879–90.

Lindsay, Bruce G. 1983. The geometry of mixture likelihoods: A general theory. *Annals of Statistics* 11:86–94.

Markussen, Simen. 2012. The individual cost of sick leave. *Journal of Population Economics* 25, no. 4:1287–1306.

Markussen, Simen, Knut Røed, Ole J. Røgeberg, and Simen Gaure. 2011. The anatomy of absenteeism. *Journal of Health Economics* 30, no. 2: 277–92.

McCall, Brian P. 1994. Identifying state dependence in duration models. In *Proceedings of the Business and Economics Section, American Statistical Association, 1994*, 14–17. Washington, DC: American Statistical Association.

OECD (Organization for Economic Cooperation and Development). 2010. *Sickness, disability and work: Breaking the barriers*. Paris: OECD.

Petterson-Lidbom, Per, and Peter Skogman Thoursie. 2013. Temporary disability insurance and labor supply: Evidence from a natural experiment. *Scandinavian Journal of Economics* 115, no. 2:485–507.

Shavell, Steven, and Laurence Weiss. 1979. The optimal payment of unemployment benefits over time. *Journal of Political Economy* 87, no. 6: 1347–62.

Topel, Robert H. 1983. On layoffs and unemployment insurance. *American Economic Review* 73, no. 4:541–59.

Ziebarth, Nicolas. 2009. Long-term absenteeism and moral hazard: Evidence from a natural experiment. Working Paper 888, DIW Berlin.

Ziebarth, Nicolas, and Martin Karlsson. 2009. The effects of expanding the generosity of the statutory sickness insurance system. Health, Econometrics, and Data Group (HEDG) Working Papers 09/35, HEDG, University of York.

# Chapter 4:

Job loss and disability insurance

Labour Economics 24 (2013) 137–150

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

# Job loss and disability insurance

Bernt Bratsberg, Elisabeth Fevang, Knut Røed *

*Ragnar Frisch Centre for Economic Research, Norway*

## HIGHLIGHTS

- We show that job loss is a major cause of disability program entry in Norway.
- The impact of job loss on disability is much larger than previously acknowledged.
- The harder it is to find a new job, the more likely that job loss causes disability.
- The "disability problem" is to a large extent an unemployment problem in disguise.

## ARTICLE INFO

*Article history:*
Received 14 August 2012
Received in revised form 26 April 2013
Accepted 6 August 2013
Available online 29 August 2013

*JEL classification:*
H55
I12
I38
J63
J65

*Keywords:*
Disability
Displacement
Social insurance
Employment opportunities

## ABSTRACT

Based on administrative register data matched with firms' financial statements and closure data collected from bankruptcy proceedings, we show that a large fraction of Norwegian disability insurance claims can be directly attributed to job displacement and other adverse shocks to employment opportunities. For men, we estimate that job loss more than doubles the risk of permanent disability retirement and accounts for one quarter of new disability insurance claims. Firm profitability and tightness of the local labor market also significantly affect employees' likelihood of disability program entry, and the adverse effects of displacement grow stronger when local labor market conditions deteriorate.



## 1. Introduction

In welfare states, the lines between unemployment and disability insurance are blurred. In this paper, we provide new insights on the causal relationships between individual employment opportunities and disability program enrollment. The study is motivated by the observations that the recent rise in disability benefit recipiency has not been paralleled by any deterioration of health conditions, and that countries with comprehensive disability insurance programs also tend to have very low unemployment rates (OECD, 2010; Røed, 2012). Building on job search theory and existing empirical evidence (Autor and Duggan, 2003; Black et al., 2002), we frame our empirical analyses on the notion that there is a gray area between unemployment and disability insurance, and that shocks to individual employment opportunities may trigger disability insurance claims even when health status remains unchanged.

Because the risks of disability and unemployment will be highly correlated at the individual level, the causal effect of employment opportunities on disability program enrollment will be difficult to identify on the basis of observational data alone. Our empirical strategy is to exploit exogenous sources of variation in individual employment opportunities, generated by variation in employers' economic performance – including profitability, downsizing, and firm closure – and idiosyncratic fluctuations in local industry-specific labor market tightness, to identify causal impacts. The empirical basis is Norwegian administrative employer–employee registers, augmented with firms' audited accounts and information collected from bankruptcy courts. The bankruptcy data make it possible to distinguish genuine mass layoffs from organizational restructuring, demergers, and takeovers.

The adverse consequences of job displacement is the focus of a broad international literature (see, e.g., Hamermesh, 1987; Ruhm, 1991; Neal, 1995; Kletzer, 1998; Kuhn, 2002; Hallock, 2009), including two recent

* Corresponding author at: Gaustadalleen 21, 0329 Oslo, Norway. Tel.: +47 2295 8813.
*E-mail addresses:* bernt.bratsberg@frisch.uio.no (B. Bratsberg), elisabeth.fevang@frisch.uio.no (E. Fevang), knut.roed@frisch.uio.no (K. Røed).

0927-5371/$ – see front matter 
http://dx.doi.org/10.1016/j.labeco.2013.08.004

studies relying on Norwegian employer–employee data (Rege et al., 2009; Huttunen et al., 2011).[1] The present paper extends this literature in several directions. It is, to our knowledge, the first study to exploit data on mass layoffs resulting from recorded bankruptcies in order to identify the impacts of exogenous displacement on the subsequent disability program and non-participation propensities of affected workers. Based on estimates of the overall number of involuntary job loss in the economy – including those from stable and growing firms – it is also the first study to assess the total impact of job loss on the frequency of disability insurance claims. We further add to the literature by examining more specifically the influences of firms' economic performance and of alternative (local) employment opportunities on employees' likelihood of entering disability insurance programs. And, finally, we examine the *interaction* between these various measures of employment opportunity to test whether the probability that job loss leads to a disability insurance claim declines with local labor market tightness.

In contrast to the existing literature, the paper also explicitly addresses the problem that the root cause of disability program enrollment may be hidden in events that took place many years prior to actual entry into permanent disability insurance. We show that social security careers ending in permanent disability retirement are often extremely long and intricate. Identification of the triggering causes therefore requires long and detailed labor market histories for the population at risk. In order to assess the impact of, e.g., job loss on the subsequent probability of becoming a disability pensioner, we either have to take into account that the outcome may materialize long after its cause, or we have to examine outcomes that materialize closer in time to their cause, but are highly correlated with the subsequent risk of receiving a permanent disability benefit. In this paper we pursue both these strategies; the former by examining entry into permanent disability insurance up to six years after displacement, and the latter by examining entry into temporary disability programs and withdrawal from the labor market.

Our results show that disability insurance and non-participation risks are indeed significantly affected by exogenous change in employment opportunities. Some of the estimated effects are large from an economic viewpoint, particularly for men. Our most reliable indicator for individual displacement is full-time employment in a firm which will go bankrupt within four years. Holding such a job raises, on average, the risk of entering permanent disability retirement during the upcoming six-year period by 2.0 percentage points for male employees and 1.2 percentage points for female employees, when compared to holding a job in a stable firm. Taking into account that the risk of job loss is present even in stable firms, we estimate that displacement raises the risk of permanent disability retirement by as much as 2.6 percentage points (121%) for men and 1.6 percentage points (48%) for women, ceteris paribus. Extrapolating these effects to all job losses in Norway, we infer that job loss accounts for around 28% of all new disability benefit claims among males and for 13% among females in our data. Not surprisingly, we also find strong impacts on the propensity for non-participation. For men, the probability of being outside the labor force after four years increases by 9.0 percentage points (123%) as a result of exogenous job loss. For women, the probability rises by 12.1 percentage points (98%). Disability insurance and non-participation propensities are also affected by more moderate downsizing processes and even by reductions in firm profitability without any observed downsizing. In addition, employment opportunities outside the current workplace play a significant role. A one standard deviation deterioration in local education/industry-specific labor market tightness (conditional on aggregate labor market tightness) raises the probability of permanent disability retirement by around 0.4 percentage points (14%) for men and 0.5 percentage points (also 14%) for women. In support of the hypothesis that disability and unemployment statuses are substitutable, we also identify significant interaction effects between job loss and local labor market conditions. The more difficult it is to find a new job, the higher is the probability that displacement leads to disability retirement.

The causal relationship between employment opportunity and disability insurance propensity will of course also reflect that job loss and unemployment entail adverse health consequences; see Kasl and Jones (2002) for a survey. In particular, our results show that, for *male* employees, job loss raises the mortality rate over a six-year period by 34 percent. For men, our data therefore support recent evidence from Sweden and the United States showing adverse effects of displacement on mortality risk (Eliason and Storrie, 2009b; Sullivan and von Wachter, 2009). However, we fail to find evidence that displacement has adverse health effects for female workers.

The estimates of causal effects of displacement on the propensities for disability insurance and non-participation presented in this paper are an order of magnitude larger than comparable estimates reported in prior studies, such as Rege et al. (2009) and Huttunen et al. (2011). We find that this disparity largely stems from differences in the operational definition of "displacement." While the findings of the prior studies are based on mass layoffs identified from employment registers alone (with, as noted by the authors, the risk of misclassification in cases of reorganizations, demergers, and takeovers), the mass layoffs exploited in this paper are identified on the basis of auxiliary information taken from bankruptcy proceedings. We demonstrate that this approach reduces attenuation bias otherwise associated with the purely register-based method. The revised effect estimates show that job loss is a major factor behind disability program participation in Norway.

## 2. Institutional background

Workers in Norway are insured against loss of work capacity from health impairment. Social insurance is compulsory and comprises sickness absence benefits, rehabilitation benefits, and disability pension. During sickness absences, the benefit replacement rate is 100%. Sickness absence benefits cannot be paid out for more than 12 months, however. Beyond 12 months, workers are eligible for rehabilitation or disability benefits provided that their work capacity is reduced by at least 50% due to sickness or injury. The replacement ratio associated with rehabilitation benefits or disability pension is typically around 66%. Rehabilitation benefits are temporary (normally 1–3 years), and are paid out during medical and/or vocational rehabilitation attempts. Disability pension is in practice a permanent benefit (lasting until the normal retirement age of 67), as the outflow from disability pension to self-supporting employment is negligible. Except for very short sickness absence spells (three days or less), all social insurance payments require that a physician certifies the health impairment. In more serious cases, the application may also be assessed by independent physicians appointed by the social security administration. It must be certified that health impairment is *the main cause* for the loss of work capacity. If this requirement is met, the law text explicitly states that the social security administration may consider the employment opportunities of the applicant when ruling whether or not the loss of work capacity is sufficiently large to qualify for benefits.

The economic incentives embedded in the social insurance replacement ratios were stable during the time period covered by this paper (1993–2006), although the period covers some attempts at tightening gate-keeping, particularly for disability pensions. For example, the requirement that the certified health impairment must be *the main cause* of the claimant's inability to work was introduced in 1995. Prior to 1995, it was sufficient that health impairment was *among the causes*. In 2000, the rehabilitation requirement was tightened such that disability benefit applicants were required to go through a vocational rehabilitation attempt, unless deemed *obviously futile*.[2] In 2004, the rules regulating the

[1] For previous Norwegian evidence that unemployment is among the key drivers of labor market detachment processes leading to permanent disability retirement, see also Bratberg (1999), Dahl et al. (2000), and Bratsberg et al. (2010).

[2] Apparently, vocational rehabilitation is deemed "obviously futile" quite often. According to our data, as many as 62% of the 2005 disability entrants had never been referred to vocational rehabilitation.

maximum duration of rehabilitation benefit payments were also tightened, leaving less room for extensions beyond one year. The same year saw the introduction of a *time-limited* disability benefit (with a maximum duration of four years). This new benefit effectively substituted for permanent disability pension for younger claimants. However, experiences so far indicate that return to employment from the time-limited disability benefit is modest, and that the arrangement essentially only has postponed entry into the permanent disability program.[3]

The employer is responsible for covering sickness insurance payments during the first 16 days of the sickness absence spell. For longer spells and for permanent disability insurance claims, the costs are covered in full by the public purse. There is no experience rating; hence there are limited pecuniary costs for firms associated with their employees utilizing long-term sickness or disability programs. In fact, when a firm has redundant labor, but finds it difficult to lay off workers due to employment protection regulations, an employee's transition to long-term sickness absence or disability insurance may be profitable for the firm.

Identifying and quantifying the roles of job loss and disemployment in explaining disability insurance claims is especially pertinent to recent developments in Norway. Over the past decades, Norway experienced a staggering rise in temporary and permanent disability program participation. Based on the data used in the present paper, we find that, over the 1993–2006 period, dependency on broadly defined health benefits increased by 34%, from 15.2 to 20.4% of the working-age population, with the ratio of those claiming permanent disability insurance to the number of unemployed rising from 1.2 to 4.0. The growth in disability rolls occurred without any corresponding deterioration in health conditions. To the contrary, subjective health indicators improved, with the proportion of the adult population reporting good or very good health rising from 79% in 1995 to 81% in 2005, and the share reporting bad or very bad health declining from 8 to 6%.[4]

## 3. Theoretical considerations

Although disability insurance eligibility requires at least 50% reduced work capacity due to sickness or injury, it is plausible that individual preferences and labor market opportunities affect application and approval decisions. Job search theory provides a useful framework for thinking about the process of entry into the disability insurance program in this context; see, e.g., Diamond and Sheshinski (1995), Autor and Duggan (2003), and Rege et al. (2009). Individuals are assumed to have preferences over the alternative labor market states of employment, job search, and inactivity (with or without disability benefits); and job displacement can be viewed as a negative shock to the value of continued labor market participation. It follows directly that there potentially is a group of individuals who prefer employment over inactivity, but nonetheless prefer disability benefit application over search for new employment. Autor and Duggan (2003) label this group "conditional disability insurance applicants," as they will apply for disability benefits only in the event of job loss. The intuition behind the conditional application strategy is that job loss shifts the discounted value of labor market participation below that of inactivity. This may happen both because obtaining a new job will incur search costs and because a new job is hard to find and likely to pay less than the prior job. Barth (1997) shows that there is a significant tenure component in Norwegian wage setting partly generated by a delayed compensation strategy (Lazear, 1981). And, as stressed by Bound and Burkhauser (1999), displacement nullifies the value of job-specific human capital and thus reduces the value of continued labor market participation. Recent empirical evidence from Norway also confirms that displacement leads to significant earnings losses (Huttunen et al., 2011).

Given the relatively strong protection against selective dismissals in Norway, it is probable that many existing employment relationships will continue despite loss of productivity caused by reduced health. In the event of job loss triggered by downsizing or closure, however, the same health problem is likely to reduce the arrival rate of new job offers and shift the distribution of wage offers downwards, and hence make job search less attractive. At the same time, the likelihood of being considered eligible for disability benefits may increase following displacement, since work capacity is assessed relative to realistic employment opportunities. This obviously entails elements of discretionary judgment by the social security administration. Røed and Westlie (2012) present empirical evidence showing that the probability of making a direct transition from unemployed job search to temporary or permanent disability enrollment rises significantly with past unemployment experience, indicating that a long and unsuccessful job search is interpreted as evidence of reduced work capacity.

Employment protection legislation does of course not provide full insurance against selective dismissals. Individual workers may legally be laid off in continuing firms if there is a factual foundation for downsizing or reorganization based on the firm's economic performance. Management may further encourage employees to quit the job, perhaps with some severance payment as a carrot, in order to achieve a desired reorganization without triggering labor conflicts. If the probability of disability program entry rises upon job loss, we would expect the future risk of disability retirement to relate negatively to firm profitability, as high profits reduce the likelihood of dismissals and employer-initiated quits.

Extending the job search model with the option of applying for disability benefits further yields the prediction that the probability of being a conditional disability insurance applicant declines with labor market tightness, as the value of unemployment rises, while the value of inactivity declines, with improved employment opportunities. In particular, an important implication of such a model is that the impact of job loss on the rate of disability program entry is larger the more difficult it is to find a new job. We therefore expect to find a negative interaction effect between job loss and labor market tightness in empirical models designed to explain disability program entry.

## 4. Data and identification challenges

The data we use in this paper consist of three parts. The first part covers a detailed account of individual labor market and social security event histories from 1992 to 2007, linked with comprehensive information about each individual. The second part includes a description of firms in terms of their employee composition and economic performance. Indicators for economic performance are constructed from annual audited accounting data, which all limited liability firms in Norway are required to make public. The third part contains information about the nature of firm closures. These data are collected from the Norwegian bankruptcy court system. A generic problem facing research based on administrative employer–employee data is to distinguish genuine mass layoffs from "spurious" layoffs, whereby a firm appears to downsize or close down while in reality it splits into smaller entities, merges with another company, or reorganizes in other ways, perhaps without laying off workers at all. A strategy pursued in the existing literature (Fevang and Røed, 2006; Henningsen and Hægeland, 2008; Rege et al., 2009) is to interpret a mass layoff as spurious when a relatively large fraction of the workers make a transition to the same new firm. But this strategy obviously fails to identify a spurious layoff that splits the workforce, e.g., when a large firm is reorganized into several smaller entities. Defining thresholds for the fraction of workers moving together may also be awkward and result in measurement error for small firms. In the present paper, we exploit additional information that we collect from bankruptcy court proceedings and that allows us to distinguish explicitly between closures due to bankruptcy, closures due to voluntary liquidation, and takeovers (with or without a bankruptcy).

[3] Our data show that, by the end of 2004, 8412 persons received a time-limited disability pension. Three years later only 2% had returned to work. As many as 65% remained on time-limited disability and 29% had entered permanent disability.

[4] These numbers are collected from Statistics Norway's level of living sample surveys, and can be downloaded from www.norgeshelsa.no.

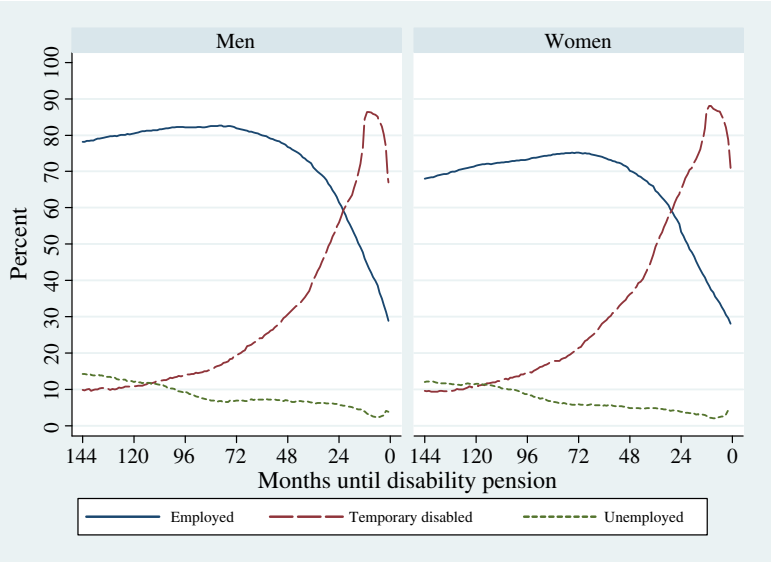



**Fig. 1.** Past labor market states of 2005 permanent disability program entrants. Note: States are not mutually exclusive, as disability and unemployment may be partial and combined with some employment. Populations consist of 13,194 men and 15,993 women of age 30 or above who entered the permanent disability program in 2005.

A particular problem arising in attempts to identify the causal effect of employment opportunities on subsequent disability insurance claims is the long and variable time lags between the presumed *cause* and its observed *effect*. When granting a new disability pension, the social security administration also sets a "disablement date." This date is meant to reflect the occurrence of the health impairment behind the loss of at least 50% of work capacity. Because benefits are based on earnings up to the time of disablement, the date becomes important for the level of benefits; hence its determination is likely to involve some considerate judgment by the case worker. On the basis of disablement dates recorded in our data, we find that the disablement on average occurs three years before entry into the permanent disability program. The variation across individuals is large, however, and for almost 20% of claimants the duration from disablement until disability retirement is more than five years. The typical duration from disablement to disability pension uptake also varies over time, primarily reflecting the various attempts (referred to above) at curbing the inflow to the permanent disability rolls. To illustrate, in our data the average "waiting time" fell from 38 months for 1997 entrants to 32 months for 2000 entrants, after which it rose to 36 months for 2003 entrants to the permanent disability program (we do not have comparable disablement date statistics for later entrants).

Many disability program entrants have long histories of labor market difficulties, often with combinations of unemployment and health problems. In these cases, it is difficult to identify a particular triggering event. Fig. 1 displays the employment and social security histories – month by month – during the 12-year period prior to permanent disability enrollment for men and women age 30 or older who entered the program in 2005. Almost one quarter of this group received social security transfers such as unemployment benefits as long as 12 years prior to obtaining the permanent disability status. Visible signs of health problems in the group as a whole, in the form of declining employment rates and corresponding increases in the proportion claiming temporary health benefits (rehabilitation or long-term sickness benefits), appeared around six years before disability program entry. Three years before entry into permanent disability, around 40% of the men and 50% of the women claimed temporary disability benefits. These patterns show that the road to permanent disability retirement can be long and winding – often involving unemployment spells as well as periods on temporary health benefits – and that very few cases are straightforward in that there is a once-and-for-all health shock leading quickly and directly to disability retirement.[5]

## 5. The effect of employment opportunities on disability program entry

### 5.1. Methodology

To allow for long time lags between employment opportunity shocks (the presumed cause) and entry into the permanent disability program (the possible effect), we have structured our dataset into three four-year time periods, starting at the end of the *base years* of 1993, 1997, and 2001, respectively. We condition the analysis on workers holding a full-time job on January 1st following the base year. In addition, we exclude workers with recent social insurance spells and drop from the samples those who received social security benefits for more than six months during the prior two years. We then examine the probability of permanent disability retirement as well as of transitions to states that involve a high risk of subsequent entry into the permanent disability program, as functions of, inter alia, exogenous change in employment opportunities. We limit the analysis to employees in private sector single-plant firms with more than 10 employees and for which we have access to audited accounting data (which includes all limited liability firms).[6] We also limit attention to individuals who were between 20 and 63 years of age in the base year and who resided in Norway throughout the analysis period. All analyses are conducted separately for men and women.

[5] The apparent decline in temporary disability just before entry into permanent disability displayed in Fig. 1 mirrors the occurrence of a "benefit vacuum" period after temporary disability insurance options are exhausted, but before the application for permanent disability benefits has been approved.

[6] A key to interpretation of our results is that workplace events can be considered *exogenous* with respect to the behavior of the individual employee. Since this assumption may be questionable for small workplaces, below we also present results based on samples of workers in large firms (more than 50 employees) to examine the robustness of our findings. The reason why we restrict attention to single-plant firms is that accounting and closure/takeover data are available at the company level. Hence, the accounting and closure data can be directly matched to workplace data for single-plant firms only. Finally, by focusing on single-plant firms we avoid complications caused by within-firm job transfers following plant closures (Huttunen et al., 2011).

**Table 1**
Analysis populations and the distribution of outcomes.

| | Men | | | Women | | |
|---|---|---|---|---|---|---|
| Base year: | 1993 | 1997 | 2001 | 1993 | 1997 | 2001 |
| Observations | 130,786 | 189,703 | 203,781 | 44,549 | 59,272 | 70,373 |
| Disability insurance (temporary or permanent) during next 4 years (%) | 9.2 | 12.6 | 13.9 | 14.4 | 19.2 | 21.3 |
| Out of labor force 4 years later (%) | 7.8 | 10.6 | 11.1 | 14.2 | 16.8 | 17.1 |
| Permanent disability program within 6 years (%) | 3.0 | 3.1 | 2.8 | 3.9 | 4.1 | 3.4 |

We focus on three alternative outcome measures for the individual:

1. Whether claiming *disability insurance* – temporary or permanent – during the four-year period following the base year.[7]
2. Whether *outside the labor force* four years after the base year.[8]
3. Whether entered the *permanent disability* program within six years of the base year.[9]

Table 1 lists the sizes of the analysis populations and the distribution of outcomes. Comparing the three periods, we note that the incidence of permanent disability program participation rose somewhat between the first and the second period, after which it declined to a level below that of the first period. The incidence of temporary (and permanent) disability program participation rose sharply throughout the three periods; for women it increased quite dramatically, from 14.4% in the 1994–97 period to 21.3% in the 2002–5 period. We interpret the shift from permanent to temporary disability program participation in the third period as reflecting attempts by the social security administration of curbing inflows into permanent disability retirement through more ambitious rehabilitation attempts; see Section 2.

Empirical analysis of the causal impact of employment opportunities on the likelihood of claiming disability benefits requires observed variation in employment opportunities that is *exogenous* to each individual's disability program propensity. Our data give three potential sources of such variation. Two of these operate at the workplace level and consist of mass layoffs and variation in firm profitability, respectively. The third operates primarily at the region-by-occupation level and consists of fluctuations in demand for the type of labor that the worker has to offer outside the present employer. While a mass layoff will have a very direct effect on the displaced workers' employment opportunities, a prediction from the theoretical framework of Section 3 is that poor (or deteriorating) firm performance may involve small-scale layoffs that place pressure on employees to quit "voluntarily" and/or to claim disability benefits of some kind. Fluctuations in local labor demand impinge on the employment opportunities for anyone searching for a new job.

In this setting, true exogeneity of workplace-specific employment opportunities might be questioned as the quality of a firm's workforce also will affect its economic performance and, hence, the likelihood of laying off workers. Moreover, firm-specific employment opportunities may correlate with other disability risk factors related to, e.g., occupation and work practices. We address these possible problems by applying extensive controls for potentially confounding factors, by examining differences in employee composition between different types of firms, and through extensive robustness checks of our findings with respect to the composition of the analysis population. These checks include analyses where we focus on large firms only, as reverse causality is more likely to be a concern for small firms.

For mass layoffs, we have chosen a forward-looking setup and assess the impacts of closure and downsizing events over a four-year period after the base year. This is motivated by the idea that "early leavers" may have started the search for a new job in response to information about an impending mass layoff, leaving remaining workers at the time of mass displacement a selected subset of the original workforce; see Kuhn (2002) for a discussion. The downsizing indicators are computed in a similar fashion as in Rege et al. (2009, p. 764), i.e., as the percent change in the number of full-time equivalent workers between the start of each period and the date exactly four years later.[10] When a workplace is downsized by 100%, we have – in contrast to prior studies – collected direct information on the reason behind the closure, i.e., whether it resulted from a bankruptcy, a voluntary liquidation, or a takeover. Firms' profitability is measured by the annual rate of return on invested capital. We include both initial profitability (in the base year) and the change in profitability over the next four years as explanatory variables in our models.

In order to extract and isolate exogenous variation in local labor market tightness, we start out by constructing two individual and time-specific tightness indices; one reflecting the probability of *becoming* unemployed, the other reflecting the probability of *finding a new job* given unemployment. Gaure and Røed (2007) show that the transition rates between unemployment and employment capture the cyclical fluctuations in labor demand better than the corresponding rates of unemployment. Both indices are computed on the basis of auxiliary (logit) regression models. To be precise, let $u_{it} = 1$ if person $i$ becomes unemployed in period $t$ and let $e_{it} = 1$ if the unemployed person finds new work within one year. We then set up the following models:

$$\begin{aligned} &\Pr(u_{it} = 1) = l(\phi_t + x_{it}\varphi_t), \\ &\Pr(e_{it} = 1 | u_{it} = 1) = l(\psi_t + x_{it}\pi_t), \\ &t = 1994{-}1996, 1998{-}2000, 2002{-}2004, \end{aligned} \tag{1}$$

where $x_{it}$ includes a large set of individual characteristics (to be explained below) including type of work (educational attainment and industry) and region (travel-to-work area) of residence, and $l(.)$ denotes the logit function, $l(a) = \exp(a)[1 + \exp(a)]^{-1}$. Based on these regressions we compute for all individuals and each of the three periods the predicted linear unemployment and reemployment propensity indices, $\widehat{\phi}_t + x_{it}\widehat{\varphi}_t$ and $\widehat{\psi}_t + x_{it}\widehat{\pi}_t$.[11] The two indices are by construction functions of individual covariates and will, at face value, *not* be independent of the error term in statistical models of individual disability program or labor market withdrawal propensities. As we explain below,

[7] Temporary disability is measured as having spells of medical or vocational rehabilitation or at least six months of long-term sickness leaves during the four-year interval.

[8] Being outside the labor force after four years is defined on the basis of social security and annual earnings data as either 1) having annual earnings or self-employment income below 144,000 NOK (2009 currency; approx 18,000 €) during the last calendar year, 2) receiving permanent disability or rehabilitation benefits in the month of December that year, or 3) receiving long-term sickness benefits in December and for at least six months out of four-year period ending that month. This definition ensures that individuals who either have earnings that are incompatible with self-sufficiency or are observed to rely on long-term social security transfers are classified as being outside the labor force.

[9] Our measure of permanent disability also includes the formally time-limited disability benefit introduced in 2004.

[10] Note that we do not exploit information on individual layoffs in order to avoid complications from selection bias in cases where some workers are retained by the firm (Henningsen and Hægeland, 2008).

[11] The two indices are designed to measure labor market tightness in the first three years of each four-year period. We do not include the fourth year for the reason that labor market tightness is likely to affect the three outcome measures with some time lag.

**Table 2**
Employment opportunities — descriptive statistics.

| | Men | | | | Women | | | |
|---|---|---|---|---|---|---|---|---|
| Base year: | All | 1993 | 1997 | 2001 | All | 1993 | 1997 | 2001 |
| Observations | 524,270 | 130,786 | 189,703 | 203,781 | 174,194 | 44,549 | 59,272 | 70,373 |
| Age | 39.4 | 39.3 | 39.1 | 39.9 | 38.7 | 37.9 | 38.6 | 39.3 |
| Education | | | | | | | | |
| Compulsory | 24.3 | 26.8 | 24.8 | 22.2 | 24.4 | 27.9 | 24.9 | 21.7 |
| Secondary | 56.1 | 55.2 | 56.6 | 56.3 | 54.1 | 56.7 | 55.4 | 51.4 |
| College/University | 19.1 | 17.6 | 18.2 | 20.9 | 21.0 | 14.9 | 19.3 | 26.3 |
| Earnings in base year (1000 NOK, 2009-value) | 402 | 373 | 389 | 434 | 297 | 262 | 287 | 327 |
| Percent subject to | | | | | | | | |
| Closure w/ bankruptcy | 2.6 | 1.4 | 2.6 | 3.2 | 1.8 | 1.0 | 1.8 | 2.3 |
| 10–20% downsizing | 9.3 | 6.9 | 10.4 | 9.8 | 9.9 | 8.6 | 10.2 | 10.5 |
| 20–35% downsizing | 8.9 | 5.3 | 10.5 | 9.6 | 10.1 | 7.4 | 11.5 | 10.6 |
| 35–99% downsizing | 14.0 | 9.8 | 15.5 | 15.4 | 15.6 | 12.7 | 18.3 | 15.0 |
| Liquidation | 5.0 | 4.7 | 5.7 | 4.5 | 5.6 | 5.1 | 6.6 | 5.2 |
| Takeover | 10.0 | 9.3 | 12.7 | 7.9 | 10.9 | 10.6 | 13.5 | 9.0 |
| Return on capital | 0.072 | 0.079 | 0.091 | 0.055 | 0.072 | 0.087 | 0.086 | 0.050 |
| Change return on capital | −0.008 | −0.001 | −0.047 | 0.023 | −0.007 | −0.012 | −0.042 | 0.026 |
| Risk of unemployment | 14.9 | 14.9 | 13.2 | 16.4 | 15.9 | 16.7 | 14.1 | 17.0 |
| Prob. of reemployment | 68.8 | 73.8 | 70.8 | 64.5 | 58.1 | 57.0 | 61.8 | 56.2 |

Note: Individual characteristics (age, education, earnings) are measured in base year, while firm downsizing and closure indicators refer to four-year period following the base year.

we deal with this endogeneity problem by controlling for $x_{it}$ in all analyses where the indices appear as explanatory variables, in essence isolating the variation in labor market opportunities that arise from time-varying *effects* of individual characteristics $(\widehat{\varphi}_t, \widehat{\pi}_t)$, in particular those driven by differences in cyclical conditions related to education, industry, and region.

Table 2 provides a descriptive overview of our analysis populations and the variables designed to represent change in individual employment opportunities. Males are strongly overrepresented in the dataset, reflecting our focus on full-time employees in the private sector. Workplace turbulence (in the form of downsizing, closure, or takeover) generally increased from the first to the second period, and declined slightly in the third period. An important exception to this pattern is the bankruptcy rate, which rose significantly over the full data period. Another important pattern to emerge from Table 2 is that takeovers make up a majority of the firm closures in the data. Around 18% of male and female full-time employees in our dataset work in a firm that "disappears" over the next four years,[12] but almost 60% of these jobs are subject to a firm takeover or acquisition and are therefore less likely to entail displacement than jobs in firms that go bankrupt. Note that while we, in cases of firm closure, can use the bankruptcy data to distinguish genuine mass layoffs from, e.g., takeovers and demergers, we are not able to make this distinction for more moderate downsizings. Hence, our downsizing indicators are likely to be "inflated" by organizational changes that do not really involve collective layoffs.

For our three ultimate outcome measures, we estimate the following models:

$$\begin{aligned}\Pr\left(y_{ijt}=1\right) &= l\left(\alpha_{jt}+z_{it}\delta_j+x_{it}\beta_j+\gamma_j\left(\widehat{\phi}_t+x_{it}\widehat{\varphi}_t\right)+\lambda_j\left(\widehat{\psi}_t+x_{it}\widehat{\pi}_t\right)\right)\\ &= l\left(\alpha^*_{jt}+z_{it}\delta_j+x_{it}\beta_j+\gamma_j x_{it}\widehat{\varphi}_t+\lambda_j x_{it}\widehat{\pi}_t\right),\\ \alpha^*_{jt} &= \alpha_{jt}+\gamma_j\widehat{\phi}_t+\lambda_j\widehat{\psi}_t,\end{aligned} \tag{2}$$

where $y_{ijt}(j = 1,2,3)$ denote the three dichotomous outcome indicators described in Table 1, observed for individual $i$ in time period $t$. The vector $z_{it}$ contains all workplace-specific covariates such as initial firm size, downsizing, closure, turnover, and profitability.[13] As explained above, the vector of individual characteristics ($x_{it}$) contains information about the (initial) type of work and region of residence. Since we do not have direct information about occupations, the type of work is proxied by a combination of educational attainment and industry (resulting in 21 different job type categories). In addition, we include information about age (i.e., 44 age dummies), nationality (eight classes), actual work experience (six classes), base year log earnings and the change in log earnings from the year prior to the base year, initial family situation (i.e., marital status, number of children and labor market status of the spouse; 10 categories), travel-to-work area (90 categories), and, for older workers, entitlement to early retirement. A complete listing of the explanatory variables ($x_{it}$, $z_{it}$) is provided in Appendix A.

A key point to note is that the coefficient vector $\{\beta_j,\gamma_j,\lambda_j\}$ in Eq. (2) can be separately identified only because there is time variation in the parameter estimates $\widehat{\varphi}_t$ and $\widehat{\pi}_t$. Without the $t$-subscript on these parameters, the regressors $x_{it}, x_{it}\widehat{\varphi}_t$, and $x_{it}\widehat{\pi}_t$ would be perfectly collinear. We have deliberately constructed the model this way in order to ensure that it is *only* the idiosyncratic changes in labor market tightness over time that identify the effects of employment opportunities on the risk of disability program entry and non-participation. In practice, the key source of identification is that different industries and economic regions were subject to different cyclical fluctuations during the three observation periods. For example, while employment opportunities in the manufacturing industries and in agriculture declined over time, particularly for workers with low educational attainment, the employment opportunities in retail, restaurants, and tourism improved.

Since an important aim of this paper is to assess the extent to which individual displacement affects the risk of subsequent disability insurance uptake, we place considerable emphasis on the effects of working in a firm that is going to close down due to bankruptcy over the upcoming four-year period. As Table 2 showed, in any of the three four-year intervals only between 1.0 and 3.2% of workers in our data actually experienced a bankruptcy. This does not imply, however, that displacements are rare. According to Salvanes (1997), as many as 10% of Norwegian jobs are eliminated in a typical year. We therefore expect displacement to be relatively common even

[12] In addition, there are some jobs in our dataset that seemingly disappear because of mismatches between firm identifiers in the two main data sources. Specifically, 1.16% of males and 1.25% of females work in firms that disappear from the employer–employee data during the upcoming four years but do not close down according to the accounting data; and 0.98 and 1.36% work in firms that vanish from the accounting data but not from the employer–employee data. We include these jobs in our analyses, but mark the observations as firm-identifier mismatches.

[13] For firms that close down during the period, we set the change in profitability equal to the sample mean in order to keep the observation in the analysis. Since we have separate dummy variables for firms that close down, this does not affect the estimated effects of the change in profitability, but it does imply that closure effects are measured relative to firms with mean change in profitability.

**Table 3**
Incidence of registered unemployment during four-year period and mean disability and participation outcomes by downsizing and closure status. Average over three sample periods.

| | Men | | | | Women | | | |
|---|---|---|---|---|---|---|---|---|
| | Registered unemployed, 4 yrs (%) | Temp or permanent disability, 4 yrs (%) | Out of labor force after 4 yrs (%) | Permanent disability, 6 yrs (%) | Registered unemployed, 4 yrs (%) | Temp or permanent disability, 4 yrs (%) | Out of labor force after 4 yrs (%) | Permanent disability, 6 yrs (%) |
| Closure w/ bankruptcy | 56.5 | 18.8 | 18.8 | 4.9 | 62.2 | 24.7 | 27.9 | 4.3 |
| No downsizing (<10%) | 12.4 | 11.3 | 8.5 | 2.6 | 13.1 | 17.6 | 14.0 | 3.4 |
| 10–20% downsizing | 17.9 | 13.1 | 11.3 | 3.4 | 19.1 | 20.1 | 16.9 | 4.0 |
| 20–35% downsizing | 21.9 | 13.7 | 11.6 | 3.3 | 23.7 | 20.5 | 18.1 | 4.3 |
| 35–99% downsizing | 26.5 | 14.0 | 13.0 | 3.7 | 29.5 | 19.9 | 19.8 | 4.3 |
| Liquidation | 19.6 | 10.5 | 10.4 | 2.6 | 25.3 | 18.5 | 17.6 | 3.5 |
| Takeover | 20.0 | 11.6 | 10.9 | 2.7 | 21.6 | 19.8 | 16.5 | 4.4 |

**Table 4**
Descriptive statistics by firm closure and downsizing status.

| | Closure w/ bankruptcy | Liquidation or takeover | Downsizing | No downsizing(<10%) |
|---|---|---|---|---|
| Outcome (%) | | | | |
| Temporary or permanent disability (4 yrs) | 19.9 | 14.1 | 15.3 | 12.8 |
| Out of labor force (4 yrs) | 20.5 | 12.7 | 13.7 | 9.9 |
| Permanent disability (6 yrs) | 4.8 | 3.3 | 3.7 | 2.8 |
| Sickness absence in base year (%) | 11.9 | 10.7 | 11.5 | 10.4 |
| Sickness absence yr before base yr (%) | 9.6 | 9.1 | 9.7 | 9.0 |
| Female (%) | 18.9 | 26.8 | 26.5 | 23.9 |
| Age | 38.1 | 38.8 | 39.6 | 39.3 |
| Education | | | | |
| Compulsory | 28.4 | 24.2 | 25.1 | 23.8 |
| Secondary | 56.7 | 54.5 | 55.1 | 56.1 |
| College/University | 14.1 | 20.9 | 19.3 | 19.6 |
| Earnings in base yr (1000 NOK, 2009) | 346 | 379 | 374 | 378 |
| Plant size | 61.6 | 109.4 | 146.5 | 110.9 |
| Number of workers (all three periods) | 16,462 | 107,409 | 195,047 | 379,546 |

Note: Sickness absence is recorded in a certain year if the person had at least one absence spell exceeding 16 days.

in stable or growing firms. Table 3 shows how the downsizing and closure indicators correlate with subsequent incidences of registered unemployment (within the corresponding four-year downsizing/closure period) in our data. With unemployment incidence rates of 57% for men and 62% for women, entry into registered unemployment is indeed *much* higher among workers exposed to a bankruptcy-driven closure than among other workers.[14] It is nonetheless clear from the table that unemployment is relatively frequent regardless of the type of downsizing event. The table also reveals that the prevalence of our disability and non-participation outcome measures are higher for workers that faced workplace restructuring than for workers in stable or growing firms, and that, at least for men, the bankruptcy category stands out with high future incidence rates of disability program entry and labor force withdrawal.

To obtain a rough estimate of the overall level of displacements in our own data, we use the unemployment frequencies reported in Table 3 as a starting point. If we assume that all employees in the "closure with bankruptcy" category are actually displaced, we can infer that 56.5% of displaced male workers and 62.2% of displaced female workers register as unemployed during the four-year period in question. If we assume that these same propensities to register for unemployment also apply to workers who lose their job in other (non-bankruptcy) firms, we can use the numbers listed in Table 3 to back out the total number of job losses in our data. Doing this exercise separately for men and women, we estimate that around 31% of both male and female employees in our dataset lose their job over a four-year period.[15] Even in the no-downsizing bracket (<10%), we find that the four-year job-loss rate is 22% for men and 21% for women. To the extent that we interpret the effects of working in a bankruptcy-exposed firm – as opposed to working in a firm with no downsizing – as representing the causal effect of displacement, our estimates will thus clearly be subject to *contamination bias* (Heckman and Robb, 1985). We return to the issue of contamination bias in Section 5.2 below.

As stressed by Rege et al. (2009), the estimated impact of firm closure may be affected by *selection bias* if workers in closing firms differ systematically from workers in continuing firms. Table 4 provides descriptive statistics for the workforces of firms in the various downsizing categories. These statistics show that there are in fact large differences in worker composition across categories. In particular, bankruptcy

[14] It is of interest to note that liquidations seem to involve unemployment entries at the same level as relatively small downsizings. This suggests that liquidations lead to fewer displacements than bankruptcies, although both events involve firm closure. Probable reasons for this pattern is that the classification "liquidated firms" contains some false closures and that an organized liquidation gives more room for maintaining viable economic activities within new firm structures compared to an outright bankruptcy.

[15] The assumption that the propensity for unemployment registration is the same for all types of job loss is of course questionable. On the one hand, one could argue that the marginal employee in a stable firm has weaker labor market prospects than the average employee displaced from a bankrupt firm. Moreover, selective layoffs may carry a stigma and serve as an adverse signal about an employee's productivity; see Gibbons and Katz (1991). These factors imply higher unemployment registration propensities for job losses in stable firms, and thus fewer actual job losses behind a given number of registered unemployed. On the other hand, job losses in continuing firms are typically announced well in advance of the event, leaving displaced workers with more time to search for new jobs and hence avoid being registered as unemployed. And congestion effects in local labor markets may imply that mass layoffs have larger adverse consequences than other layoffs. Such factors suggest higher registration frequencies for job losses in closing firms. It is also worth noting that our 31% estimate is only slightly below what would be expected on the basis of the 10% annual job elimination rate reported by Salvanes (1997), which – provided that the risk is independently distributed across individuals over time – yields a 35% cumulative displacement rate over a four-year period ($1 - 0.9_4$).

**Table 5**
Estimated percentage point impacts of employment opportunities on disability program entry and non-participation.
Average marginal effects (robust standard errors in parentheses).

| | Men | | | Women | | |
|---|---|---|---|---|---|---|
| | Temp. or permanent disability 4 yrs | Out of labor force 4 yrs | Permanent disability 6 yrs | Temp. or permanent disability 4 yrs | Out of labor force 4 yrs | Permanent disability 6 yrs |
| Closure with bankruptcy | 4.72 | 6.99 | 2.02 | 4.30 | 9.57 | 1.23 |
| | (0.53) | (0.46) | (0.23) | (0.79) | (0.86) | (0.40) |
| No downsizing (<10%) | Ref. | Ref. | Ref. | Ref. | Ref. | Ref. |
| 10–20% downsizing | 0.46 | 1.48 | 0.37 | 1.14 | 1.52 | 0.18 |
| | (0.22) | (0.35) | (0.09) | (0.38) | (0.36) | (0.16) |
| 20–35% downsizing | 1.05 | 1.65 | 0.45 | 1.09 | 2.07 | 0.41 |
| | (0.20) | (0.21) | (0.10) | (0.37) | (0.36) | (0.18) |
| 35–99% downsizing | 1.68 | 2.89 | 0.86 | 0.75 | 3.96 | 0.72 |
| | (0.21) | (0.26) | (0.10) | (0.32) | (0.35) | (0.16) |
| Liquidation | 0.78 | 3.04 | 0.68 | 1.63 | 4.52 | 0.77 |
| | (0.29) | (0.32) | (0.15) | (0.49) | (0.53) | (0.24) |
| Takeover | 0.01 | 0.39 | 0.07 | 1.30 | 0.90 | 0.21 |
| | (0.20) | (0.25) | (0.08) | (0.34) | (0.33) | (0.15) |
| Initial rate of return on capital[a] | −0.12 | −0.58 | −0.09 | −0.60 | −0.46 | −0.11 |
| | (0.08) | (0.19) | (0.03) | (0.13) | (0.14) | (0.06) |
| Change in return on capital[a] | −0.21 | −0.41 | −0.06 | −0.30 | −0.39 | −0.09 |
| | (0.08) | (0.15) | (0.03) | (0.13) | (0.14) | (0.06) |
| Risk of unemployment[a] | 1.67 | −0.07 | 0.47 | 1.57 | −0.30 | 0.23 |
| | (0.29) | (0.28) | (0.13) | (0.46) | (0.42) | (0.20) |
| Probability of reemployment[a] | −0.72 | −1.52 | 0.10 | −0.62 | −2.23 | −0.29 |
| | (0.23) | (0.20) | (0.09) | (0.46) | (0.40) | (0.17) |
| Percent with outcome = 1 | 12.22 | 10.10 | 2.96 | 18.81 | 16.23 | 3.77 |

Number of observations: 524,270 (men) and 174,194 (women). Standard errors are clustered within 34,620 (men) and 29,700 (women) firm-by-period cells. The following controls are included in the regressions (number of categories for categorical variables in parentheses): Education/industry (21), age (44), nationality (8), actual work experience (6), initial level and change in log earnings, family situation (10), region of residence (90), size of municipality (5), firm size (4), employee turnover in base year (5), time period (3), firm-identifier mismatch (3), and, for old workers, entitlement to early retirement programs (2).

[a] The variables are standardized, such that they are centered on zero and have a unit standard deviation. Marginal effects are calculated as the effect of a one standard deviation change in the explanatory variable.

firms have fewer female employees, lower fractions of highly educated workers, and lower average earnings than stable firms. Bankruptcy firms also tend to be smaller than other firms. Given the sample sizes reported at the bottom of the table, these differences cannot be attributed to randomness alone; hence they must be accounted for in the empirical analysis. For the analysis, it would be of concern if workers' reliance on health-related benefits in bankruptcy firms deviated from that in other firms even prior to the start of the analysis period. As our analysis samples are conditioned on not having received any *long-term* health benefits prior to the outcome period, such sorting problems should primarily show up in observed short-term benefits, i.e., sick pay. The numbers in Table 4 indicate that the rate of sickness absence during the base year indeed is somewhat higher in bankruptcy firms than in other firms. The year before the base year, however, there are only minor differences between the different firm types. A possible interpretation of these patterns is that the higher absence rate in soon-to-go-bankrupt firms reflects that the downsizing process has already started in some of these firms.

To formally test for whether employees in closing firms, conditional on our explanatory variables, have higher initial absence rates than employees in stable or growing firms, we estimate separate models with indicators for sickness absence in the base year and in the year before the base year, respectively, as the dependent variable. The models are formulated exactly as the models we use for other outcome variables and include the same control variables (see Eq. (2)). Results (not reported in tables) show that the estimated average marginal effect of working in a closing (bankruptcy) firm on absenteeism in the base year is equal to 0.86 percentage points for men (t-value = 2.69) and −0.19 percentage points for women (t-value = −0.80). For the year before the base year, however, we fail to uncover significant differences across firm types; 0.29 percentage points (t-value = 1.28) for men and 0.26 percentage points (t-value = 0.33) for women. We interpret these findings as supporting evidence for the hypothesis that the higher absence rate in the base year in soon-to-go-bankrupt firms captures an early causal effect of the turbulence and stress associated with the forthcoming closure; see Røed and Fevang (2007). The failure to identify significant differences in the year prior to the base year indicates that compositional differences by closure status is not driven by sorting of employees across firms. We nevertheless return to the issue of sorting in terms of past sickness absence in the robustness exercises below.

### 5.2. Results from the baseline model

Table 5 presents our key results regarding the impacts of employment opportunity on subsequent disability program entry and non-participation for men and women, respectively. For ease of interpretation, we report average marginal effects (multiplied by 100); i.e., the mean percentage point impact of the explanatory variable on each of the three outcome probabilities. Average marginal effects are computed on the basis of relevant comparisons only; for dummy variable sets with more than two categories, each category's average marginal effect is calculated for observations belonging to the category in question and the reference category only (see Bartus, 2005). A complete listing of estimated coefficients is available from the authors.[16]

As Table 5 shows, employment opportunities have large and statistically significant effects on disability program entry and non-employment propensity. For both men and women, the probability of claiming permanent disability benefits after six years, and the

[16] In order to account for any covariance between employees working at the same establishment (and to correct for Moulton (1986) bias), we cluster standard errors within firm-by-period cells. Were we instead to cluster at the establishment level (to also account for any serial correlation across periods), standard errors would be slightly larger than those reported in the tables. To illustrate, the standard error of the coefficient of the bankruptcy variable in the male permanent disability logit equation becomes 0.06118 (21,332 cluster units) as opposed to 0.06082 (34,620 clusters). Note also that the three periods will contain multiple observations of some of the workers in our sample (the baseline samples consist of 524,270 observations of 347,748 males and 174,194 observations of 128,391 females). Using clustering to account for serially correlated errors among individuals with multiple observations raises standard errors by an even smaller amount than clustering within firms.

likelihood of being out of the labor force after four years, rise monotonically with the level of workplace downsizing, ceteris paribus. All three outcome propensities decline with the employer's economic performance and, at least for men, with improvements in local labor market tightness as captured by the risk of unemployment and re-employment variables.

As explained above, our most reliable indicator of *exogenous displacement* is the "closure with bankruptcy during the next four years" variable. As shown in Table 5, such an event raises a male worker's probability of claiming *permanent* disability benefits after six years by 2.0 percentage points when compared to working in a stable or growing firm with average profitability. Given the large and variable time lags in entry into permanent disability status described in Section 2, and because virtually all permanent disability benefit claims are preceded by extended periods on temporary disability benefits and/or by self-supported periods outside the labor force, it is of interest to examine the impacts on these outcomes as well. According to the estimates in Table 5, a bankruptcy raises a male full-time worker's probability of claiming either temporary *or* permanent disability benefits by 4.7 percentage points and the probability of labor force withdrawal (measured four years after the base year) by 7.0 percentage points. These large additional flows into temporary disability and non-participation show that the 2.0 percentage point rise in the permanent disability program participation rate identified after six years does not capture the full effect of displacement.

The effects of job loss on disability insurance claims and non-participation are large for women as well, though generally smaller than those for men when measured relative to the average outcome within gender. For a female full-time worker, bankruptcy raises the risk of permanent disability program entry by around 1.2 percentage points. The risk of temporary or permanent disability rises by around 4.3 percentage points. One reason why the effects tend to be smaller for women than for men, may relate to gender differences in mental distress associated with unemployment – and perhaps not being able to fulfill the traditional breadwinner role – a point to which we return in Section 5.5 below. It is worth noting that the overall impact of bankruptcy on the probability of non-participation is larger for women than for men; the likelihood of non-participation following bankruptcy goes up by 9.6 percentage points for women (compared to 7.0 for men). But, because our analysis covers private sector employees only – leading to a huge overrepresentation of men – some caution is warranted when interpreting gender differences in effect estimates.

The estimates listed in Table 5 show the effect of working in a bankruptcy firm as opposed to a stable or growing firm, and not the effect of displacement per se. We can nevertheless use the estimated effects to evaluate the underlying causal impacts of displacement. As we argued in Section 5.1, displacement is relatively common even in stable and moderately downsizing firms. This implies that the estimated effects of closure with bankruptcy reported in Table 5 in fact understate the causal effects of displacement. Adjusting the point estimates for contamination bias caused by inclusion of treated (i.e., displaced) employees in the non-treatment (no downsizing) group, we find that displacement on average raises the permanent disability program propensity for men by 2.6 percentage points (121%) and by 1.6 percentage points (48%) for women.[17] Likewise, the risk of temporary or permanent disability following job loss rises by 6.0 percentage points (60%) for men and by 5.5 percentage points (32%) for women. Finally, accounting for contamination bias, displacement raises the non-participation propensity by 9.0 percentage points (123%) for men and by 12.1 percentage points (98%) for women. Based on the (admittedly questionable) assumptions that these effects are representative for all displaced workers in our dataset and that our estimate of the overall number of job losses is correct (see Section 5.1), we estimate that displacements account for fully 28% of all new permanent disability benefit claims among males and for 13% among females (see footnote 17 for the exact calculations). Similarly, we find that for men (women), displacements account for 28 (23) percent of transitions to non-employment and for 16 (9) percent of transitions to temporary or permanent disability programs.

The economic performance of surviving firms – as measured by the annual return on their capital base – also has statistically significant effects on transitions into disability programs and non-participation (conditional on the observed level of downsizing). Although the effects on disability benefit claims are moderate in size, they are far from negligible. For example, a one-standard-deviation deterioration in initial profitability and its four-year change will raise the female entry rate into temporary or permanent disability by 0.9 percentage point (0.6+0.3). Our interpretation of this finding is that poor economic performance of the employer does entail small-scale displacement and places pressures on employees with poor health.

Local industry-specific labor market conditions significantly affect transitions into disability programs and non-participation. For example, a one standard deviation increase in the unemployment incidence index raises the likelihood of entering a temporary or permanent disability program by 1.7 percentage points for both men and women (around 14% for men and 9% for women). A negative shock to the local labor market resulting in higher unemployment risk and reduced likelihood of reemployment (both of a magnitude of one standard deviation) is predicted to raise the inflow rate to permanent disability by 0.4 percentage point (i.e., by 14%) for men and 0.5 percentage point (also 14 percent) for women.

Our estimated displacement effects are considerably larger than those reported in two recent studies also based on Norwegian register data. Rege et al. (2009) find that workers originally employed in plants that downsized by more than 60% between 1995 and 2000, were 24% more likely to utilize disability pensions in 2001 than comparable workers in non-downsizing plants. And Huttunen et al. (2011), who define displaced individuals as workers who separate from plants that reduce employment by 30% or more, report that the probability of being outside the labor force is 3.4 percentage points higher seven years after displacement than for otherwise similar, but non-displaced, workers. When we replicate the definition of downsizing used by Rege et al, we also replicate their main result.[18] The implication is that the conventional definition of downsizing and closure based on employer–employee data imparts attenuation bias in estimates. Although both studies take steps to eliminate false downsizings and/or focus on high-seniority workers, register-based downsizing indicators will invariably capture some false downsizings and closures related to outsourcing, demergers, and other forms of organizational change. Moreover, some separations are voluntary, even when they occur in downsizing firms. In fact, the authors point out themselves that their strategies for identifying displacement will involve some misclassifications. Our results, showing much larger effects of displacement on disability benefit uptake and labor market withdrawal, suggest that this indeed is the case.

[17] We adjust for contamination bias by dividing the estimated average marginal effect of "closure with bankruptcy" by the estimated fraction of non-displaced workers in non-downsizing firms. To illustrate, for men the adjusted effect is calculated as $2.02/(1-0.22)=2.59$, where 0.22 is the estimated fraction of displacement over the four-year interval among males in non-downsizing firms; see Section 5.1. We compute the counterfactual disability entry rate – the rate that would have prevailed in the absence of any displacements – as the actual entry rate minus the product of the estimated average effect of displacement and the computed overall rate of displacements. In the example given for men, this yields a counterfactual non-displacement disability rate of 2.14. As the observed rate in the data is 2.96 (see the bottom row of Table 5), we estimate the fraction of overall disability entries that can be attributed to displacements to be $(2.96-2.14)/2.96=0.28$.

[18] Rege et al. (2009) report an estimated odds-ratio associated with 60–100% downsizing of 1.30. Our own corresponding estimate is 1.31.

**Table 6**
Heterogeneous effects of bankruptcy.
Average marginal effects (robust standard errors in parentheses).

| | Men | | | Women | | |
|---|---|---|---|---|---|---|
| | Temp. or permanent disability 4 yrs | Out of labor force 4 yrs | Permanent disability 6 yrs | Temp. or permanent disability 4 yrs | Out of labor force 4 yrs | Permanent disability 6 yrs |
| Bankruptcy | 4.33 | 9.69 | 0.49 | 7.35 | 12.44 | −0.13 |
| | (1.98) | (2.25) | (0.36) | (4.97) | (5.71) | (0.85) |
| Reemployment index[a] | −0.90 | −0.97 | 0.00 | −0.56 | −2.07 | −0.11 |
| | (0.21) | (0.17) | (0.03) | (0.45) | (0.39) | (0.07) |
| Bankruptcy * reemployment[a] | −1.19 | −0.75 | −0.13 | −3.13 | −3.77 | −0.23 |
| | (0.52) | (0.38) | (0.06) | (1.43) | (1.35) | (0.18) |
| Bankruptcy * (age > 50) | 0.62 | 1.85 | 0.05 | −2.27 | 0.61 | 0.01 |
| | (0.80) | (0.74) | (1.17) | (2.14) | (2.24) | (0.37) |
| Early retirement elig. | −3.12 | 5.92 | −0.42 | −3.23 | 9.26 | −0.57 |
| | (0.25) | (0.48) | (0.02) | (0.71) | (1.01) | (0.06) |
| Bankruptcy * early retire't elig | −2.62 | −1.02 | −0.42 | −5.59 | 1.36 | −0.56 |
| | (1.25) | (0.96) | (0.10) | (4.56) | (6.39) | (0.41) |
| Log earnings base yr | −3.68 | −4.53 | −0.62 | −1.39 | −7.20 | −0.52 |
| | (0.20) | (0.16) | (0.04) | (0.39) | (0.34) | (0.07) |
| Bankruptcy * log earn base yr | 4.09 | 2.44 | 0.50 | 5.55 | 5.29 | 0.71 |
| | (0.67) | (0.47) | (0.14) | (1.74) | (1.39) | (0.38) |

Control variables include the downsizing, closure, and firm characteristics listed in Table 5 as well as all controls listed in note to Table 5. In addition, the regressions control for interactions between bankruptcy and education/industry, nationality, work region, municipality, firm size and turnover, and time period. The baseline bankruptcy effect is evaluated for a native-born, low-educated manufacturing worker in Oslo and employed in a small firm with low turnover during the first observation period of the study. See also notes to Table 5.

[a] Marginal effects are calculated as the effect of a one standard deviation change in the explanatory variable.

### 5.3. Heterogeneous effects

According to the theory outlined in Section 3, substitutability between unemployment and disability insurance schemes implies that there is an interaction effect between displacement and local industry-specific labor market tightness. In particular, a prediction from the framework is that the risk of disability benefit uptake following displacement will be higher when it is difficult to find a new job. To examine this possibility – and also investigate the existence of other potential heterogeneous effects – we have estimated models that allow for interactions between the bankruptcy variable and labor market and individual characteristics. Table 6 presents some key results (the full set of results is available from the authors). As predicted, disability benefit uptake depends on local labor market conditions, and poor employment prospects aggravate the adverse effect of displacement. In fact, the coefficient of the interaction term between closure with bankruptcy and the reemployment index is negative for all outcome measures and for both genders. To illustrate, a one standard deviation increase in the reemployment index reduces the probability that a bankruptcy-affected male worker receives temporary or permanent disability benefits by a statistically significant 1.2 percentage point and that of a female worker by as much as 3.1 percentage points. The evidence is thus consistent with the conclusion of Couch and Placzek (2010) that the adverse consequences of job loss are greater during economic downturns.

The table further shows that transition rates to disability programs and out of the labor force following displacement are slightly higher for older workers. This conclusion is turned upside down, however, for workers eligible for early retirement. For the latter group of workers, there does not seem to be any effect of displacement on disability program entry at all, indicating a strong element of yet another social program substitutability, this time between early (state subsidized) retirement and disability pensions. This interpretation is reinforced by the coefficient estimates showing that, among displaced workers, those eligible for early retirement are less likely to enter disability programs, but much more likely to leave the labor force than workers not eligible for early retirement. Another point to note from Table 6 is that there is a tendency for "the social gradient" in disability program entry to be weaker for the flows generated by mass layoffs. This is illustrated by the impacts of prior earnings. In general, there is a strong negative relation between prior earnings and the likelihood of disability benefit uptake, particularly for men. The relation likely reflects heterogeneity in health – in that poor health causes both low earnings and disability – and that the opportunity costs of disability program enrollment are larger for workers with high earnings. Interestingly, this relationship vanishes in bankruptcy firms. Upon job loss, the local labor market opportunities apparently become more important relative to individual background characteristics, again supporting the notion of unemployment-disability substitution.[19]

### 5.4. Robustness analyses

Even though the results presented in Table 5 account for a rich set of control variables, we cannot a priori rule out that employees in downsizing and closing firms differ systematically from employees in stable or growing firms. For example, the layoff process in closing firms may have started during or before the base year, leaving a selected group of employees in terms of unobserved disability risk. Moreover, there is the concern of reverse causality: If many workers in a small firm become disabled, this may have detrimental effect on the firm's economic performance, and can – at least for small firms – even cause bankruptcy.

Tables 7 and 8 report the estimated average marginal effects of our key explanatory variables from a number of robustness exercises for men and women, respectively. To ease comparisons, in column I we first list the estimates from the baseline model. In column II, we examine whether the estimated effects of bankruptcy are impacted by inclusion of the firm profitability and local labor market tightness measures in the empirical model. The results show that this is not the case — if anything, dropping these measures raises the estimated impact of bankruptcy. Column III lists the estimated effects based on employees in the restricted sample of firms that did not downsize at all during the two years prior to the outcome period. If our results were driven by early sorting caused by an ongoing downsizing process, we would expect estimates to be sensitive to this sample condition. As it turns out, they are

[19] The attempts at tightening gate-keeping referred to in Section 2 might be expected to have affected caseworkers' scopes for considering applicants' employment prospects and thus reduced the effect of job loss over time; see Gruber and Kubik (1997), Campolieti (2004), and de Jong et al. (2011). Although not statistically significant, results indicate somewhat lower bankruptcy effects for men towards the end of our sample period. For women, we do not uncover any systematic differences in estimated bankruptcy effects across the three periods.

**Table 7**
Robustness analysis for men. Estimated percentage point impacts of employment opportunities on temporary or permanent disability program entry and non-participation. Average marginal effects (AME).

| | I Baseline model | II Omit profits and labor demand indices | III Firm size stable last two years | IV More than 50 employees | V No welfare benefits prior two years | VI With controls for past absence | VII Include multi-plant firms | VIII Include region-specific time dummies | IX Employed in the same firm prior five years |
|---|---|---|---|---|---|---|---|---|---|
| Observations | 524 270 | 524 270 | 489 368 | 232 684 | 388 592 | 524 270 | 1 137 749 | 524 270 | 208 311 |
| *A) Temporary or permanent disability program after 4 years* | | | | | | | | | |
| Closure w/ bankruptcy | 4.72 | 4.90 | 4.55 | 4.30 | 4.45 | 4.44 | 4.35 | 4.68 | 6.68 |
| Return on capital | −0.12 | | −0.14 | 0.00 | −0.07 | −0.11 | −0.19 | −0.12 | −0.21 |
| Change in ret. capital | −0.21 | | −0.24 | −0.06 | −0.15 | −0.22 | −0.19 | −0.18 | −0.13 |
| Risk of unempl. index | 1.67 | | 1.65 | 2.09 | 1.42 | 1.62 | 1.15 | 1.97 | 1.69 |
| Prob. of reempl. index | −0.72 | | −0.77 | −0.28 | −0.59 | −0.55 | −0.74 | −0.81 | −0.71 |
| Percent w/ outcome = 1 | 12.22 | 12.22 | 12.18 | 12.49 | 8.44 | 12.22 | 11.46 | 12.22 | 12.87 |
| *B) Out of labor force after 4 years* | | | | | | | | | |
| Closure w/ bankruptcy | 6.99 | 7.82 | 6.95 | 7.26 | 6.28 | 6.80 | 6.75 | 7.01 | 9.17 |
| Return on capital | −0.58 | | −0.42 | −1.08 | −0.59 | −0.58 | −0.34 | −0.52 | −0.97 |
| Change in ret. capital | −0.41 | | −0.33 | −0.69 | −0.40 | −0.42 | −0.22 | −0.35 | −0.59 |
| Risk of unempl. index | −0.07 | | 0.08 | −0.60 | −0.16 | −0.10 | −0.43 | −0.18 | −1.41 |
| Prob. of reempl. index | −1.52 | | −1.58 | −1.61 | −1.36 | −1.42 | −1.66 | −1.60 | −1.56 |
| Percent w/ outcome = 1 | 10.10 | 10.10 | 9.98 | 10.71 | 7.63 | 10.10 | 9.82 | 10.10 | 10.15 |
| *C) Permanent disability program after 6 years* | | | | | | | | | |
| Closure w/ bankruptcy | 2.02 | 2.15 | 1.83 | 2.50 | 1.76 | 1.95 | 1.86 | 2.01 | 3.55 |
| Return on capital | −0.09 | | −0.08 | −0.11 | −0.06 | −0.09 | −0.08 | −0.08 | −0.09 |
| Change in ret. capital | −0.06 | | −0.07 | −0.03 | −0.05 | −0.07 | −0.10 | −0.05 | −0.06 |
| Risk of unempl. index | 0.47 | | 0.44 | 0.65 | 0.31 | 0.45 | 0.30 | 0.53 | 0.75 |
| Prob. of reempl. index | 0.10 | | 0.10 | 0.22 | −0.02 | 0.13 | −0.03 | 0.12 | 0.26 |
| Percent w/ outcome = 1 | 2.96 | 2.96 | 2.90 | 3.22 | 2.03 | 2.96 | 2.94 | 2.96 | 4.23 |

not. Column IV presents estimates for employees in large firms only (more than 50 employees). If our results reflected reverse causality, the estimated impacts should drop significantly when we restrict the sample to employees in large firms. They do not.

Columns V and VI report estimates based on the sample limited to workers without welfare benefits at all during the past two years, and estimates based on the full sample, but with additional controls included for past absences (in the form of dummy variables indicating incidences of long-term absence in the base year and in the year before the base year), respectively. If our results were driven by systematic sorting of employees with poor health into bankruptcy firms, the estimated impacts of bankruptcy should drop in these exercises. Once again, they do not.

Column VII presents estimates based on the extended sample of workers employed in multi-plant as well as single-plant firms. If workers in single-plant firms differ systematically from those in multi-plant firms, our results might not generalize to workers at large. Effect estimates based on the extended sample change only marginally relative to the baseline, though, and the slight decline in the estimated effect of bankruptcy is consistent with our presumption that bankruptcies in large (multi-plant) companies often entail the continuation of some of the plants' economic activities, and hence that bankruptcy is a less precise indicator of job loss in multi-plant than in single-plant firms.

Column VIII lists estimates from a model where we have allowed the time dummy variables to vary by region (with the country divided into five regions). If there were regional trends in disability uptake not caused by business cycle developments, our baseline model could confound such trends with business cycle effects. As it turns out, when we allow for region-specific trends, the within-region estimates of labor market tightness effects are, if anything, larger than the estimates of the baseline model. Again, deterioration of local re-employment opportunities raise the probability of disability program entry.

Finally, column IX presents estimates based on reduced samples conditioned on stable employment in the same firm for at least five years. In the literature, restricting the sample to high-seniority workers is a common practice, typically for reasons of eliminating voluntary quits and firings for cause from the group of displaced workers.[20] It is also probable that job loss is a more severe shock for high-seniority workers with more job-specific human capital and a stronger expectation of remaining in their current job than for recent hires. As the column shows, the estimated impacts of bankruptcy rise significantly when we impose the seniority restriction. While the pattern to some extent is explained by much lower contamination of displacements in the reference group of stable firms (not shown), the substantial difference from the baseline nonetheless indicates that the adverse effects of job loss increase with seniority. An implication for the empirical job-loss literature is that studies that focus on high-seniority workers may exaggerate the average impact of worker displacement.

The main message coming out of the robustness exercises is that the estimated marginal effects from our baseline model are highly robust with respect to data delimitation and model specification. If anything, the estimated bankruptcy effects from our baseline model turn out to be on the conservative side; most of the robustness exercises yield stronger effects. For the other parameters of interest (i.e., the coefficients of the profitability and labor market tightness variables), there are only minor variations across the different model specifications and samples.

### 5.5. Effects on health

Our finding that employment opportunities have a strong impact on subsequent disability benefit claims does not necessarily imply that the disability status results *directly* from unemployment only. Previous evidence from Norway suggests that job loss adversely affects employees' physical and mental health conditions (Rege et al., 2009), and

[20] A number of studies adopt the U.S. Bureau of Labor Statistics definition of displacement and limit samples to workers with at least three years of seniority (Fallick, 1996). See also the discussions of high vs. low tenure workers and the implications for measurement of displacement effects in Jacobson et al. (1993) and von Wachter et al. (2009).

**Table 8**
Robustness analysis for women. Estimated percentage point impacts of employment opportunities on temporary or permanent disability program entry and non-participation. Average marginal effects (AME).

| | I Baseline model | II Omit profits and labor demand indices | III Firm size stable last two years | IV More than 50 employees | V No welfare benefits prior two years | VI With controls for past absence | VII Include multi-plant firms | VIII Include region-specific time dummies | IX Employed in the same firm prior five years |
|---|---|---|---|---|---|---|---|---|---|
| Observations | 174 194 | 174 194 | 135 216 | 73 063 | 119 858 | 174 194 | 401 060 | 174 194 | 61 320 |
| *A) Temporary or permanent disability program after 4 years* | | | | | | | | | |
| Closure w/ bankruptcy | 4.30 | 4.98 | 4.10 | 2.47 | 4.75 | 4.21 | 3.59 | 4.34 | 6.09 |
| Return on capital | −0.60 | | −0.60 | −0.78 | −0.43 | −0.50 | −0.36 | −0.54 | −0.91 |
| Change in ret. capital | −0.30 | | −0.27 | −0.50 | −0.38 | −0.30 | −0.13 | −0.25 | −0.43 |
| Risk of unempl. index | 1.57 | | 1.73 | 1.68 | 1.46 | 1.38 | 1.32 | 1.68 | 1.25 |
| Prob. of reempl. index | −0.62 | | −0.57 | −0.56 | −0.77 | −0.64 | −0.77 | −0.96 | −0.75 |
| Percent w/ outcome = 1 | 18.81 | 18.81 | 18.74 | 19.81 | 13.40 | 18.81 | 18.60 | 18.81 | 18.97 |
| *B) Out of labor force after 4 years* | | | | | | | | | |
| Closure w/ bankruptcy | 9.57 | 10.22 | 9.27 | 8.99 | 8.69 | 9.52 | 9.28 | 9.62 | 13.89 |
| Return on capital | −0.46 | | −0.40 | −0.88 | −0.29 | −0.40 | −0.20 | −0.41 | −0.81 |
| Change in ret. capital | −0.39 | | −0.34 | −0.56 | −0.48 | −0.39 | −0.05 | −0.33 | −0.68 |
| Risk of unempl. index | −0.30 | | −0.16 | −0.89 | −0.36 | −0.41 | −0.06 | −0.75 | −1.40 |
| Prob. of reempl. index | −2.23 | | −2.40 | −2.76 | −1.91 | −2.24 | −2.02 | −2.59 | −1.58 |
| Percent w/ outcome = 1 | 16.23 | 16.23 | 16.11 | 16.06 | 12.57 | 16.23 | 15.61 | 16.23 | 15.04 |
| *C) Permanent disability program after 6 years* | | | | | | | | | |
| Closure w/ bankruptcy | 1.23 | 1.36 | 1.35 | 1.86 | 1.62 | 1.20 | 1.32 | 1.27 | 3.24 |
| Return on capital | −0.11 | | −0.07 | −0.16 | −0.06 | −0.09 | −0.02 | −0.10 | −0.32 |
| Change in ret. capital | −0.09 | | −0.09 | −0.07 | −0.04 | −0.09 | −0.01 | −0.08 | −0.01 |
| Risk of unempl. index | 0.23 | | 0.23 | 0.02 | 0.15 | 0.24 | 0.07 | 0.33 | 0.44 |
| Prob. of reempl. index | −0.29 | | −0.29 | −0.38 | −0.13 | −0.29 | −0.18 | −0.29 | −0.22 |
| Percent w/ outcome = 1 | 3.77 | 3.77 | 3.71 | 3.94 | 2.59 | 3.77 | 3.74 | 3.77 | 6.17 |

evidence from Sweden indicates that it significantly increases the risk of hospitalization due to alcohol-related conditions (Eliason and Storrie, 2009a). There is also empirical evidence showing that the mental distress associated with unemployment typically is more severe for men than for women; see Waters and Moore (2002), McKee-Ryan et al. (2005), or Kuhn et al. (2009). More generally, recent empirical studies find that work tends to be a healthy activity, particularly for workers with illnesses that are responsible for the majority of disability insurance claims in advanced economies, such as musculoskeletal pain and mental disorder; see, e.g., Waddel (2004), Waddell and Burton (2006), and OECD (2008). Markussen et al. (2013) show that continued work during episodes of long-term illness in most cases improves future labor market prospects.

**Table 9**
Estimated percentage point impacts of employment opportunities on mortality six years after base year. Average marginal effects (robust standard errors in parentheses).

| | Men | Women |
|---|---|---|
| Closure with bankruptcy | 0.26 | 0.06 |
| | (0.10) | (0.15) |
| No downsizing (<10%) | Ref. | Ref. |
| 10–20% downsizing | 0.11 | 0.13 |
| | (0.05) | (0.07) |
| 20–35% downsizing | 0.09 | 0.09 |
| | (0.06) | (0.07) |
| 35–99% downsizing | 0.09 | 0.06 |
| | (0.05) | (0.06) |
| Liquidation | 0.08 | −0.10 |
| | (0.08) | (0.08) |
| Takeover | 0.10 | 0.01 |
| | (0.05) | (0.06) |
| Initial rate of return on capital[a] | 0.00 | 0.00 |
| | (0.02) | (0.02) |
| Change in return on capital[a] | −0.05 | 0.01 |
| | (0.02) | (0.02) |
| Risk of unemployment[a] | 0.13 | −0.06 |
| | (0.07) | (0.08) |
| Probability of reemployment[a] | 0.02 | −0.04 |
| | (0.05) | (0.07) |
| Percent with outcome = 1 | 1.09 | 0.58 |
| Observations | 527,684 | 174,781 |

[a] Marginal effects are calculated as the effect of a one standard deviation change in the explanatory variable. See also notes to Table 5.

To check for possible health effects of job loss, we extend our samples and include workers who otherwise satisfy initial sample criteria (e.g., age 20–63 in the base year), but who died during the six-year outcome period. We next estimate the impacts of employment opportunities on mortality in exactly the same manner as we have estimated the impacts on other outcome measures. The results reported in Table 9 show that displacement appears to raise mortality for men but not for women. Adjusting the estimated bankruptcy effect for contamination bias caused by inclusion of displaced employees in the control group (non-downsizing firms), we find that displacement raises the six-year mortality rate for men by 0.33 percentage points (34 percent). This implies that around 10% of the deaths among male workers in our data can be attributed to job displacement. A general deterioration of local industry-specific risk of unemployment also tends to raise mortality among men. For women, coefficient estimates of the downsizing variables (without closure) are similar in size to those for men. The latter is consistent with large literature indicating that the uncertainty associated with organizational change adversely affects the health of retained employees; see, e.g., Ferrie (2001) and Røed and Fevang (2007), the latter for recent Norwegian evidence. [21]

## 6. Concluding remarks

We have shown in this paper that negative shifts in employment opportunities explain significant shares of non-participation and disability insurance dependency in Norway. The causal relationship

[21] The causal link between displacement and mortality risk has also been studied in other countries. For example, Eliason and Storrie (2009b) and Sullivan and von Wachter (2009) report mortality effects among displaced male workers in Sweden and Pennsylvania that are larger than those of the present study. Martikainen et al. (2007) uncover an association between unemployment and mortality risk in Finland, but argue that there is no excess mortality among displaced workers.

between employment opportunities and disability program entry is particularly strong for male workers. According to our baseline estimates, job loss more than doubles the risk of subsequent program entry for men, while raising enrollment by approximately 50% for women. These effects are considerably larger than those of prior studies. We find that the conventional measures of downsizing and firm closures used in employer–employee data impart attenuation bias in estimates, which explains the discrepancy across studies.

For men, we have uncovered evidence that a portion of the job loss effect can be explained by adverse health consequences. For women, no such health effects have been identified. These findings are in accordance with previous evidence indicating that the adverse health impacts of job loss are indeed more severe for men than for women. For both genders, we have found that the impacts of job loss on subsequent disability program entry are larger the worse are local labor market conditions. Moreover, the development of local labor market conditions as well as of the current employer's profitability have distinct impacts on the employees' risk of disability program entry. A probable explanation is that management may coerce workers to apply for disability insurance benefits as a way of cutting costs without having to resort to layoffs, and that their incentives for pursuing such strategies rise in times of low profitability and adverse local economic conditions.

Taken together, the evidence presented in this paper points to a considerable element of substitutability between unemployment and disability insurance. Our findings suggest that the process of reallocating redundant workers from old to new employers is far from seamless, and that many displaced workers permanently change status from supporting the welfare state to becoming supported by it. Significant human capital resources are squandered in this process. The finding that loss of employment is among the major causes of disability program entry – whether it stems from genuine health effects or from adverse shocks to the expected value of labor market participation for given health levels – suggests that appropriate solutions to the "disability problem" should address strategies for improving the employment opportunities of potential claimants rather than focus exclusively on income insurance. If job loss and unemployment are among the root causes of the rising disability problem, it is probable that provision of employment opportunities is among its remedies.

## Acknowledgments

We are grateful to Oddbjørn Raaum, Mari Rege, Kjell Gunnar Salvanes, Kjetil Telle, and two anonymous referees for valuable comments. We acknowledge funding from the Norwegian Research Council (grant #s 173583 and 202513). The paper is part of the research activities of the Centre of Equality, Social Organization, and Performance (ESOP), University of Oslo. Data made available by Statistics Norway have been essential for the research project.

## Appendix A

List of explanatory variables used in the baseline model:

Age in base year: 44 dummy variables; one for each age 20–63.
Marital status in base year: 4 dummy variable; single, married, divorced, widow(er).
Children: 3 dummy variables; No children, 1–2 children, 3 + children.
Spouse/family situation: 3 dummy variables; spouse home, spouse home * 1–2 children, spouse home * 3 + children.
Education/industry: 21 dummy variables: low/primary, low/manufacturing, low/retail, low/hotel/restaurant, low/transport, low/finance, low/education, low/health, low/other, medium/primary, medium/manufacturing, medium/retail, medium/hotel/restaurant, medium/transport, medium/finance, medium education, medium/health, medium/other, bachelor degree, graduate school, education missing.
Work experience: 6 dummy variables; 1–5 years, 6–10 years, 11–15 years, 16–20 years, 21–25 years, >25 years.
Earnings: Two scalar variables; log earnings in base year, difference in log earnings from the year before the base year to the base year.
Early retirement eligibility: 2 dummy variables; eligible or not eligible for early retirement benefits during the four-year period in question (eligibility depends on age and on the firm's affiliation to the early retirement program).
Immigrant status: 8 dummy variables; OECD, East Europe, Middle East/North Africa, Other Africa, South East Asia, South America, not immigrant.
Place of residence: 90 dummy variables; corresponding to travel-to-work-areas defined by Statistics Norway.
Size of municipality: 5 dummy variables; <2000, 2–5000, 5–10,000, 10–50,000, >50,000.
Firm size in base year: 4 dummy variables; 11–25, 26–50, 51–200, >200.
Firm turnover in base year: 5 dummy variables; No turnover, 0.1–10%, 10–15%, 15–20%, >20%.
Downsizing: 4 dummy variables; No downsizing <10%, 10–20%, 20–35%, 35–99.9%,
Closure: 5 dummy variables; No closure, closure with bankruptcy, liquidation, takeover.
Firm profitability: 2 scalar variables; Return on capital in base year, change in return on capital from base year (t) to year t+3.
Labor market tightness: 2 scalar variables from auxiliary regression; risk of unemployment and probability of reemployment.
Time: 3 dummy variables, one for each of the three periods in the dataset.

## References

Autor, D., Duggan, M.G., 2003. The rise in the disability rolls and the decline in unemployment. Q. J. Econ. 118, 157–205.
Barth, E., 1997. Firm-specific seniority and wages. J. Labor Econ. 15 (3), 495–506.
Bartus, T., 2005. Estimation of marginal effects using margeff. Stata J. 5 (3), 309–329.
Black, D., Daniel, K., Sanders, S., 2002. The impact of economic conditions on participation in disability programs: evidence from the coal boom and bust. Am. Econ. Rev. 92, 27–50.
Bound, J., Burkhauser, R.V., 1999. Economic analysis of transfer programs for people with disabilities. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, 3C. Elsevier.
Bratberg, E., 1999. Disability retirement in a welfare state. Scand. J. Econ. 101 (1), 97–114.
Bratsberg, B., Raaum, O., Røed, K., 2010. When minority labor migrants meet the welfare state. J. Labor Econ. 28 (3), 633–676.
Campolieti, M., 2004. Disability insurance benefits and labor supply: some additional evidence. J. Labor Econ. 22 (4), 863–889.
Couch, K.A., Placzek, D.W., 2010. Earnings losses of displaced workers revisited. Am. Econ. Rev. 100 (1), 572–589.
Dahl, S.-Å., Nilsen, Ø.A., Vaage, K., 2000. Work or retirement? Exit routes for Norwegian elderly. Appl. Econ. 32 (14), 1865–1876.
De Jong, P., Lindeboom, M., van der Klauw, B., 2011. Screening disability insurance applicants. J. Eur. Econ. Assoc. 9 (1), 106–129.
Diamond, P., Sheshinski, E., 1995. Economic aspects of disability benefits. J. Public Econ. 57 (1), 1–23.
Eliason, M., Storrie, D., 2009a. Job loss is bad for your health — Swedish evidence on cause-specific hospitalization following involuntary job loss. Soc. Sci. Med. 68 (8), 1396–1406.
Eliason, M., Storrie, D., 2009b. Does job loss shorten life? J. Hum. Resour. 44 (2), 277–302.
Fallick, B., 1996. A Review of the recent empirical literature on displaced workers. Ind. Labor Relat. Rev. 50 (1), 5–16.
Ferrie, J.E., 2001. Is job insecurity harmful to health? J. R. Soc. Med. 94, 71–76.
Fevang, E., Røed, K., 2006. Veien til uføretrygd i Norge. Rapport 10/2006.Stiftelsen Frischsenteret for samfunnsøkonomisk forskning (www.frisch.uio.no/pdf/rapp06_10.pdf).
Gaure, S., Røed, K., 2007. How tight is the labour market? Sources of changes in the aggregate exit rate from unemployment across the business cycle. J. Bus. Cycle Meas. Anal 3 (2), 155–174.
Gibbons, R., Katz, L., 1991. Layoffs and lemons. J. Labor Econ. 9 (4), 351–380.

Gruber, J., Kubik, J., 1997. Disability insurance rejection rates and the labor supply of older workers. J. Public Econ. 64 (1), 1–23.
Hallock, K.F., 2009. Job loss and the fraying of the implicit employment contract. J. Econ. Perspect. 23 (4), 69–93.
Hamermesh, D., 1987. The costs of worker displacement. Q. J. Econ. 102 (1), 51–76.
Heckman, J.J., Robb, R., 1985. Alternative methods for evaluating the impact of interventions: an overview. J. Econ. 30, 239–267.
Henningsen, M., Hægeland, T., 2008. Downsizing as a sorting device. Are Low-productive Workers More Likely to Leave Downsizing Firms? Discussion Paper, 543. Statistics Norway.
Huttunen, K., Møen, J., Salvanes, K.G., 2011. How destructive is creative destruction? Effects of job loss on mobility, withdrawal and income. J. Eur. Econ. Assoc. 9 (5), 840–870.
Jacobson, L.S., Lalonde, J.T., Sullivan, D.G., 1993. Earnings losses of displaced workers. Am. Econ. Rev. 83 (4), 685–709.
Kasl, S., Jones, B., 2002. The impact of job loss and retirement on health. In: Kawachi, B.A. (Ed.), Social Epidemiology. Oxford University Press.
Kletzer, L.G., 1998. Job displacement. J. Econ. Perspect. 12 (1), 115–136.
Kuhn, P.J. (Ed.), 2002. Losing Work, Moving On. International Perspectives on Worker Displacement. W.E. Upjohn Institute for Employment Research, Kalamazoo, Michigan.
Kuhn, A., Lalive, R., Zweimüller, J., 2009. The public health costs of job loss. J. Heal. Econ. 28, 1099–1115.
Lazear, E.P., 1981. Agency, earnings profiles, productivity, and hours restrictions. Am. Econ. Rev. 71, 606–620.
Markussen, S., Mykletun, A., Røed, K., 2013. The case for presenteeism — evidence from Norway's sickness insurance program. J. Public Econ. 96, 959–972.
Martikainen, P., Maki, N., Jantti, M., 2007. The effects of unemployment on mortality following workplace downsizing and workplace closure: a register-based follow-up study of Finnish men and women during economic boom and recession. Am. J. Epidemiol. 165 (9), 1070–1075.
McKee-Ryan, F.M., Song, Z., Wanberg, C.R., Kinicki, A.J., 2005. Psychological and physical well-being during unemployment: a meta-analytic study. J. Appl. Psychol. 90 (1), 53–76.
Moulton, B.R., 1986. Random group effects and the precision of regression estimates. J. Econ. 32 (3), 385–397.
Neal, D., 1995. Industry-specific human capital: evidence from displaced workers. J. Labor Econ. 13 (4), 653–677.
OECD, 2008. Are All Jobs Good for Your Health? The Impact of Work Status and Working Conditions on Mental Health. Chapter 4 in OECD Employment Outlook 2008, Paris.
OECD, 2010. Sickness, disability and work — breaking the barriers. A Synthesis of Findings Across OECD Countries.OECD, Paris.
Rege, M., Telle, K., Votruba, M., 2009. The effect of plant downsizing on disability pension utilization. J. Eur. Econ. Assoc. 7 (5), 754–785.
Røed, K., 2012. Active social insurance. IZA J. Labor Policy 1 (8). http://dx.doi.org/10.1186/2193-9004-1-8.
Røed, K., Fevang, E., 2007. Organizational change, absenteeism and welfare dependency. J. Hum. Resour. 42 (1), 156–193.
Røed, K., Westlie, L., 2012. Unemployment insurance in welfare states: the impacts of soft duration constraints. J. Eur. Econ. Assoc. 10 (3), 518–554.
Ruhm, C.J., 1991. Are workers permanently scarred by job displacements? Am. Econ. Rev. 81 (1), 319–324.
Salvanes, K.G., 1997. Market rigidities and labour market flexibility: an international comparison. Scand. J. Econ. 99 (2), 315–333.
Sullivan, D., von Wachter, T., 2009. Job displacement and mortality: an analysis using administrative data. Q. J. Econ. 124 (3), 1265–1306.
von Wachter, T., Song, J., Manchester, J., 2009. Long-term earnings losses due to job displacement during the 1982 recession: an analysis using longitudinal administrative Data from 1974 to 2004. Mimeo, Columbia University.
Waddel, G., 2004. The Back Pain Revolution, Second edition. Elsevier Health Sciences.
Waddell, G., Burton, A.K., 2006. Is work good for your health and well-being? Report for Department for Work and Pensions. TSO (The Stationery Office), London.
Waters, L.E. And, Moore, K.A., 2002. Predicting self-esteem during unemployment: the effect of gender, financial deprivation, alternate roles, and social support. J. Employ. Couns. 39, 171–189.

# Chapter 5:

Vulnerable groups and labor market performance.
Towards more sorting in the labor market?

January, 2019

# Vulnerable groups and labor market performance. Towards more sorting in the labor market?

Elisabeth Fevang, *The Ragnar Frisch Centre for Economic Research, Oslo, Norway*[*]

## Abstract

Based on Norwegian register data, with information about all individuals born in the period 1967-1984, I explore how employment propensities and labor earnings of vulnerable groups have developed relative to the population at large. Vulnerable groups are defined as individuals having either poor health, low cognitive ability or a disadvantaged family background. My main indicator of poor health is low birth weight, which is observed for both men and women. In addition, I use information about height, Body Mass Index (BMI) and cognitive ability measured at age 18-19 for men entering the military service. Family background is described in terms of socioeconomic class, which is defined by parents' earnings rank during their age 50-54. For men, lower birth weight and underweight at age 18-19 has become a stronger predictor of non-employment and low earnings, while there is quite constant effect of height and obesity. For women, where birth weight is the only health-measure I have, I do not find any evidence of changing impact. The importance of cognitive skills on labor market success has diminished over time, which is due to decreasing returns to high ability. The most striking finding, however, is that poor social background has become a steadily more important determinant of non-employment and low earnings.

*Keywords*: Health and inequality, employment, cognitive ability
*JEL classification*: D63, I14, J21, J24

---

[*] This paper is part of the Frisch Centre project 1178 "Egalitarism under pressure? New perspectives on inequality and social cohesion" and 4134 "Disentangling absence patterns" financed by the Norwegian Research Council. The data made available by Statistics Norway, Medical Birth Registry of Norway and the Norwegian Armed Forces has been essential for the paper. Thanks to Knut Røed, Astrid Grasdal, Oddbjørn Raaum, Øyvind Anti Nielsen, Ole-Petter Hansen, Otto Lillebø, Atle Midttun, Erik Reinert, and a number of conference and seminar participants for valuable comments and suggestions. Correspondence to: Elisabeth Fevang, the Ragnar Frisch Centre for Economic Research, Gaustadalléen 21, 0349 Oslo, Norway. E-mail: elisabeth.fevang@frisch.uio.no

# 1 Introduction

In Norway, as well as in other OECD-countries, policy makers seek to promote a high labor market participation rate and an inclusive labor market. According to the European Commision (2017) "Labor markets are inclusive when everyone of working age can participate in paid work, especially vulnerable and disadvantaged people". At the same time many would argue that the labor market has become tougher because of increased focus on reorganizations and productivity. Empirical studies have shown that mass layoffs and organizational changes raise the probability of permanent exclusion from the labor market (Rege *et al.*, 2009; Bratsberg *et al.*, 2013), and it may also adversely affect the health of the employees (Ferrie, 2001; Kiwimäki *et al.*, 2001). However, there is no clear evidence that these events have become more prevalent over time, or that they are implemented in a less inclusive fashion. Moreover, there have also been changes that have improved working conditions for many employees, for instance access to better physical aids, less manual work and more flexible work schedules.

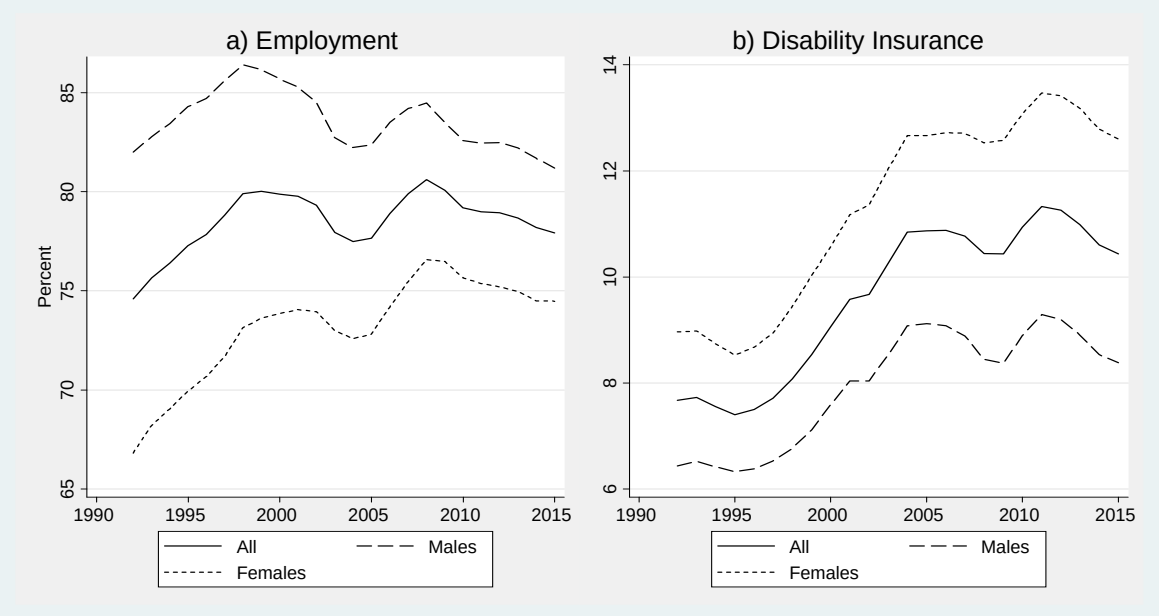



***Figure 1. Employment rate (panel a) and disability rate (panel b) in Norway 1992-2015. Prime-age (ages 25-54).***

*Note; Individuals are defined as employed if annual earnings exceed NOK186,000 ( ≈ US$23,000) measured in 2017 NOK which is 2 times the "basic-amount" (often called G) in the Norwegian public pension system. Earnings encompass wages and business income, including sickness benefits and parental leave payments. The fraction with disability insurance is measured at the end of March each year. Disability insurance include temporary disability insurance and permanent disability insurance. Temporary disability insurance corresponds to work assessment allowance, which was introduced in 2010. Work assessment allowance replaced medical and vocational rehabilitation benefits (1992-2010) and time-limited disability insurance (2004-2010).*
*Source: own calculations based on administrative registers.*

Figure 1 illustrates the development in employment and recipients of disability insurance in Norway. From 1992 to 2015, the prime-age employment rate increased from 74.6 percent to 77.9 percent, which is driven by an increasing number of women entering the labor market. During this period, there has been some fluctuations in the employment rate and it peaked in 2008, where 80.6 percent were employed. Since 2008, the employment rates for both genders have decreased. There was a tremendous increase in the level of individuals claiming disability insurance from the middle of the 1990's to the middle of 2000's, since then the level has been quite stable around 10-11 percent. Taken at face value, this high, and for many years rising, share of individuals receiving health-related benefits may indicate that the labor market has become *less* inclusive, and possibly that health problems have played an increasingly important role in limiting labor market participation in Norway. Alternatively, the rising participation in disability insurance programs simply mirrors that the welfare state now takes better care of those who (in any case) fall outside the labor market.

The aim of this paper is to examine this question empirically. The analysis covers all children born in Norway from 1967 to 1984, and labor market performance is evaluated during age 27-31; i.e., between 1994 and 2015. Over this time period, I investigate whether labor market success has become more or less causally influenced by factors determined early in life (or even prior to birth), such as social background, innate health, and cognitive ability (IQ). In particular, I am interested in studying the extent to which the labor market has become increasingly or decreasingly open for persons who got a particularly bad starting point in life along any of these dimensions. Have our society succeeded in facilitating labor market participation even for these vulnerable groups, or have we moved toward a more selective and less inclusive labor market?

As I explain in the next section, there is already a large empirical literature on the separate relationships between health, ability, and social background, on the one hand, and labor market performance, on the other. However, there is little systematic evidence on the way the influence of these factors have changed over time. In particular, there is to the best of my knowledge, no existing evidence at all built on analyses where all these factors have been incorporated simultaneously. Since health, ability, and social background are likely to be highly correlated, such a simultaneous analysis may be important in order to understand what is going on in the labor market.

Poor health is already known to be associated with greater risk of being outside the labor market. From a research point of view, the key challenge is to identify the direction of causality

and to disentangle the correlation arising from other common determinants (third factors) from the causal mechanisms of interest. One way to identify the effect of health on employment is to use health indicators measured before people reach working age, and then control for any factors that may have affected both this measure and adult employment. The main indicator for poor health used in the present paper is low birth weight (<2,500 gr), which has the advantage of being indisputably predetermined relative to any attempts at labor market entry, but the disadvantage that it is a very imprecise proxy for the health status relevant for labor market success.

To establish the social background of all newborns, I use administrative registers to collect information about their parents' labor-related earnings during their age 50-54, and apply that information to obtain a rank-based socioeconomic status (SES) indicator. My main indicator for low SES is that the sum of the two parents' earnings during this period belongs to the lowest vigintile (5 percent bin) in their generation.[1]

While the indicators for low birth weight and low SES are available for all children born in Norway during the period covered by the analysis in this paper, data on other vulnerability indicators that I use are observed only for men conscripted to military service. As additional proxies for health status at a time more relevant with respect to labor market performance, I utilize information about height and Body Mass Index (BMI), both measured at age 18-19 for men entering the military service. In addition, to establish a proxy for low cognitive ability, I use scores from IQ tests, also performed at the time of military conscription. As a result, the analyzes in this paper are largely divided into two main parts; one encompassing *all* children born between 1967 and 1984, with a focus on the influence of low birth-weight and low SES, and one covering boys only, with a focus on the influence of a wider range of health indicators, SES, and cognitive ability.

Most of the analyzes in this paper are performed by means of ordinary least squares (OLS) regressions, with employment and log earnings at age 27-31 as the main dependent variables. Identification of causality is largely based on a control function approach, with primary focus on the changes over time in the estimated impacts. In addition, I estimate models based on siblings fixed effects, as well as twin fixed effects. However, although these models clearly solve some identification challenges, I will argue that they raise some serious concerns related to external validity, and that they are poorly suited for examination of time trends in causal effects.

[1] Results based on alternative ranking criteria are presented in a separate robustness analysis.

A key finding in this paper is that the labor market has become gradually more sorted with respect to socioeconomic status (SES); and being born into a low SES family has become an ever stronger predictor for poor labor market performance as adult, both in terms of non-employment and low earnings. This is the case for both men and women. Although the social gradient in labor market performance becomes slightly weaker when birth weight, height, BMI and cognitive ability are controlled for, such controls do not alter the conclusion with respect to the rising influence of social background. For men, I also find indications of increased sorting with respect to health. Although this relationship is not as robust as that for social background, most of the results point in the direction that men with poor health have experienced a steady decline in relative employment prospects. In particular, men with low birth weight and underweight at age 18-19 are to an increasing extent non-employed at prime age. For women, where birth weight is the only available health-measure, I do not find any changing impact on employment.

Consistent with earlier research, I find that the impact of cognitive skills on labor market performance has diminished over time. This finding is primarily driven by decreasing returns to very high ability; the influence of having low – relative to medium – ability has been more stable. Given that my indicators for social background, health, and ability are all noisy, and at the same time highly interrelated, it is clear that I cannot provide a complete decomposition of how these different factors affect adult labor market performance. In particular, it is probable that my health indicators are too crude to fully account for the relationship between social background and health. Another challenge for the identification of health effects is that the correlation between the various health indicator and actual health in adulthood may have changed over time. This is particularly the case for low birth weight, as huge improvements in neonatal health care both have raised the probability of surviving with poor health (implying a stronger relationship between low birth weight and poor adult health) and limited the adverse consequences of low birth weight, given survival (implying a weaker relationship). To assess this challenge, I return to an analysis of the relationship between birth weight and alternative adult health measures after having presented the main results. It essentially fails to identify significant changes in the relationship between birth weight and adult health, although it cannot rule out such changes either.

The rest of the paper is organized as follows. Section 2 provides an overview of related literature, section 3 describes the data and section 4 presents some descriptive statistics. Section 5 explains the empirical strategy. Section 6 presents the main results. Section 7 shows reliability and

robustness checks, including comparison with family fixed effects. Finally, section 8 concludes the paper.

## 2 Related literature

The paper relates to several strands of empirical literature examining the relationships between health, ability, social background, and labor market outcomes. In particular, a vast literature has documented that neonatal health, most commonly proxied by birth weight, influences health outcomes later in life, cognitive development, educational attainment and employment/earnings. According to the World health Organization (WHO), birth weight less than 2,500 grams (gr) represents a significant health disadvantage, and is therefore used as the primary threshold for low birth weight. Babies with low birth weight are either preterm born (gestational length less than 37 weeks) and/or having low fetal growth rate, both affecting children's health. This cut-off point is often used in studies of neonatal health, however most of the recent studies also investigates differences within the normal range of birth weight. For example, children with lower birth weight are more likely to have behavioral problems like attention deficit hyperactivity disorder (ADHD) (Linnet *et al*., 2006) and poorer cognitive development (Figlio *et al*., 2014).[2] Effect of birth weight on mental disorders and cognitive development is also observed in adulthood; see Flensborg-Madsen and Mortensen (2017) and Abel *et al.* (2010).[3] Furthermore, lower birth weight is associated with higher risk of obesity, diabetes and coronary heart disease in adulthood (Barker, 2006).[4] Although the positive effect of gaining weight is not restricted to low birth weight (birth weight <2,500 gr), the effect is less clear at the top of the weight distribution. In recent studies, there has been more focus on potential negative effects of high birth weight (>4,500 gr) which may be a risk factor of overweight, obesity and diabetes (Cnattingius *et al*., 2012; Harder *et al*., 2007).

---

[2] The study of Linnet *et al*. (2006) includes all children (n=834) born in the period 1980-1994 registered with ADHD in the Danish Psychiatric Central Register. 20,100 children are used as controls. They control for socioeconomic status of the parents, familial psychopathology and maternal smoking during pregnancy. Figlio *et al*. (2014) use data about all children born in Florida 1992-2002 1.3 million singletons and almost 15,000 twin pairs. Sibling and twin fixed effects as well as cross-sectional analysis with control for infant birth month and year, maternal age, education and birth order, give almost the same results

[3] The article by Flensborg-Madsen and Mortensen (2017) is based on data from individuals born in Copenhagen 1959-1961 while Abel *et al*. (2010) include all children born in Sweden (1973-1984) and Denmark (1979-1986). The studies control for birth year, mothers age at birth, parental economic class and birth order.

[4] Barker (2006) is a meta-analysis and the referred analysis have controlled for confounding variables.

Apart from Figlio *et al*. (2014) all the studies described in the previous section are medical studies relying on cross-sectional data with control for confounding variables like social background. To identify the causal effect of birth weight, the predominant solution in economics has been to use sibling or twin fixed effects. A large economic literature documents that siblings/twins of lower birth weight have worse outcomes in terms of schooling attainment, test scores, IQ, height, disability insurance dependency, wages, employment (see f.ex. Bharadwaj *et al*., 2018; Figlio *et al*., 2014; Royer, 2009; Black *et al*., 2007; Currie and Moretti, 2007; Behrman and Rosenzweig, 2004).

The papers referred to above have information about health at birth, measured by birth weight, but they do not have any information about health after birth. A number of recent studies have focused on childhood health (both physical and mental) and outcomes later in life. Based on data about all children born in Great Britain one week in March 1958 combined with interviews at different ages, Case *et al*. (2005) find that children with chronic conditions in childhood have lower educational attainment, wages, employment probabilities and poorer health as adults. Furthermore, they show that uterine environment (measured as mothers' smoking habits during pregnancy and low birth weight) also affects adult outcomes, even when conditioning on childhood health. Using retrospective measures about perceived health in childhood from the Panel Study of Income Dynamics (PSID), Smith (2009) shows that siblings with poorer childhood health have lower earnings.

There is some evidence suggesting that the adverse consequences of mental health problems experienced during adolescence is more severe than those of physical health problems. Goodman *et al.* (2011) utilize the same data source from Great Britain as Case *et al.* (2005). Their study shows that mental health problems experienced by the age of 16 were associated with a 28 percent reduction in household income by the age of 50, while there were only minor effects of physical health problems. Currie *et al.* (2010) utilize administrative register data about children born 1979-1987 in the Canadian province of Manitoba with information about birth weight and diagnoses (f.ex. asthma, major injuries, ADHD/conduct disorders) during childhood. The authors show that young adults having poor health in early childhood are more likely to be on social assistance and having lower educational attainment. This effect is present mainly because poor health in childhood tends to persist in adulthood. Physical health problems that disappear later on have little impact on future outcomes, while children once diagnosed with ADHD or conduct disorder experience more

adverse outcomes as adults. Finally, the study shows that lower birth weight is a significant predictor of future use of social assistance and school outcomes, also when conditioning on the other observed health measures.

Evensen *et al*. (2016) and Evensen *et al.* (2017) utilize Norwegian health survey data about 8,000 adolescents linked with register data about subsequent labor market outcomes. Using sibling fixed effects the studies show that externalising problems, such as attention or conduct problems, have a clear negative impact on later educational attainment and earnings. Internalizing problems, such as anxiety or depression, reduce adult earnings, but it does not seem to have any negative effect on educational attainment. Furthermore, Evensen *et al*. (2017) find that individuals in the bottom of the earnings distribution are more negatively affected by mental health problems.

Adult height is partly determined by genetics, but it also serves as an indicator of an individual's infectious and dietary disease history during childhood, see Elo and Preston (1992). Height is a variable that has been used as a proxy for health in several studies, see f.ex. van den Berg *et al*. (2014) and Black *et al*. (2007), and being taller is associated with higher earnings in many studies (see f.ex. Lundborg *et al*., 2014; Case and Paxson, 2008; Persico *et al*., 2004). Possible explanations for the positive relationship between height and labor market performance are that height is positively correlated with cognitive and non-cognitive skills, muscular strength, and favourable family background, and that taller people are subjected to positive discrimination. In Sweden, all males entering the military service conduct tests to measure cognitive and non-cognitive skills and muscular strengths. Lundborg *et al*. (2014) use these test data of 450,000 Swedish men entering the military service at age 18-19 during the period 1984-1997, combined with register data on earnings in 2003, and show that the height premium is partly explained by these variables. The estimated height premium was further reduced when controlling for family characteristics by using sibling fixed effects, and the remaining effect was concentrated to the very shortest individuals having low earnings. The authors suggest that this remaining effect could be interpreted as a discrimination of short people in the labor market.

WHO classifies the Body Mass Index (BMI) as underweight, normal weight, overweight and obesity. Overweight and obesity are major risk factors for many chronic diseases, including diabetes, cardiovascular diseases and several types of cancer (see f.ex. Aune *et al*; 2016). Both underweight, overweight and obesity are associated with increased risk of dying, but less is known about the risk factors of underweight. Whether underweight, in itself, is a risk factor or whether

there is a reverse causation in the sense that underweight is a result of preexisting illness and smoking is not clear (see f.ex. Aune *et al.*; 2016; Roh *et al*., 2014). Based on data from Switzerland, Roh *et al*. (2014) find that higher mortality risk among underweight individuals are mainly caused by a higher risk of external causes of deaths (e.g. accidents, suicides). Using survey-data from England, Kelly *et al*. (2010) find that underweight individuals are more likely to smoke, to drink alcohol, and to be inactive. In addition, the study find that both underweight and overweight/obesity are associated with a higher prevalence of respiratory disease, less physical activity and poorer mental health.

Several studies have shown that obesity is associated with lower wage and increased risk of non-employment among women, while the effect is less pronounced among men, (see f.ex Johansson *et al*., 2009; Atella *et al*., 2008; Morris, 2007). The effect is present also after controlling for confounding variables, which may indicate that employers discriminate against obese individuals. Less is known about underweight and outcomes in the labor market. However, a recent study by Hughes and Kamari (2017) find a U-formed relationship between BMI and unemployment, where underweight and obesity are positively correlated with unemployment and overweight is negatively correlated.

Until now, I have focused on literature documenting the effect of various observed health measures on later outcomes and the references have typically controlled for social background in order to find causal effects of the different measures. Some studies have also focused on the importance of social background on children's health, which in turn affects outcomes later in life. Based on data from US (Case *et al*., 2002) and Canada (Currie and Stabile, 2003), the studies find that low household income is associated with poorer health among children, as measured by perceived health status (reported by a parent and/or a physician) and chronic health conditions. These children are more likely to have poorer health in adulthood and/or lower educational attainment which in turn affect earnings and employment. This literature suggests that part of the intergenerational transmission of socioeconomic status may be related to health in childhood.[5]

A large literature documents that there has been increasing returns to skills as technological changes have increased the demand for high-skilled workers. In US, the steadily increasing reward

[5] The explanation is also confirmed by other studies, see f.ex. Currie, 2009; Johnston *et al*., 2011.

of skills has been documented since the 1960's (Bound and Johnson, 1992; Juhn *et al*., 1993; Acemoglu and Author, 2011). After 2000 there has been little or no growth in cognitive skill-intensive occupations (Acemoglu and Author, 2011; Beaudry *et al*., 2014). By comparing two surveys (NLSY79 and NLSY97) Deming (2017) shows that there has been a strong growth in social skill-intensive jobs between 1980 and 2012, while there has been a decrease in jobs requiring high cognitive skills and less social skills (STEM jobs). Similar findings are documented using Swedish register data. During the period 1992-2013, Edin *et al*. (2017) find an increase in the wage reward of non-cognitive skills whereas the returns to cognitive skills have been quite stable. Deming (2017) points out one main explanation for these findings: Technology is to an increasing extent substituting for cognitive skills, while social skills are more difficult to replace.

When it comes to whether the impact of health and social background has changed over time, empirical evidence is limited. Bharadwaj *et al*. (2018) compare Swedish twin pairs born 1974 and onwards with twin pairs born 1926-1958 and their twin fixed effects show that the estimated effects of birth weight on earnings and high school completion are quite similar across the groups of cohorts.[6] Their study, however, does not explore whether individuals are employed or not and their sample is restricted to twins. Moving on to the role of social background, Markussen and Røed (2017) show that being born into the poorest families have become a stronger predictor of non-employment and lower earnings.

## 3 Data

In this paper, I utilize administrative register data covering the whole population of Norway. The data contains information about age, gender, level of education, employment status, annual earnings and social security transfers. Most of the datasets are available for the period 1992-2015, while information about earnings is given back to 1967. The datasets include information on family-ties, making it possible to identify social background. I study labor market outcomes for all persons born in Norway from 1967 through 1984.

[6] One other study has focused on the development in the *correlation* between health and employment. Based on British data from the period 1973-2009, Minton et al (2012) find that the difference in employment rates among individuals with and without limiting long term illness has grown substantially during the period.

### *3.1 Definition of outcomes*

*Earnings.* Since I am interested in development of labor market outcomes and the first cohort in my sample is born in 1967, I have to focus on earnings at an early stage in life. I choose earnings at age 27-31. Earnings above 99th percentile are censored. In order to include individuals with zero earnings I add a small amount of earnings to every individual when using log-specification of earnings.[7] Earnings encompass wages and business income, including sickness benefits and parental leave payments.

*Employment*. I define employment as having average annual earnings above 2 times the "basic-amount" (often called G) in the Norwegian public pension system during age 27-31. In 2017 2G corresponds to NOK186,000( ≈ US$23,000), which is approximately 1/3 of average full-time earnings in Norway.[8] With this definition, a person will be defined as employed if he/she has relatively low earnings over, say, 4-5 years or high earnings over just one or two years. Since I have information on employment until 2015, I will be able to cover people born in the period 1967-1984 in these analyzes. This means that I will identify changes during a period of 18 years.

*Mortality.* I will distinguish between infant mortality (mortality within the first year of life) and mortality the next years.

*Education.* I focus on whether the individuals have completed high school or not at the age of 27. High school completion is defined as having at least 12 years of education.

*Disability.* I define a person as disabled if he/she receives permanent disability insurance some time during the calendar year the person turns 19. The reason I am interested in considering disability at this early stage of life is that my usage of this outcome is primarily intended to assess whether the relationship between birth weight and adult health has been stable over time. I therefore wish to avoid transitions into the disability insurance program that may have been caused by labor market developments.

[7] This amount is 1 G=93,000 NOK during the whole period age 27-31, which correspond to 0.2 G per year.
[8] The basic amount is adjusted each year approximately corresponding to the general wage growth.

### *3.2 Health*

Information about birth weight is taken from The Medical Birth Registry, which covers all births in Norway since 1967. By this register, I also have information about year and month of birth, gestational length, and whether it is a single or a multiple birth.

For men, I also use information obtained from the Norwegian military records. Every able men are obliged to attend the military service, and at the age of 18-19. Before actually entering the service, they are called in for an examination of whether they are liable for the service. From this examination I use information about height and weight, which are also combined into a body mass index (BMI).

### *3.3 Ability*

From the military records, I also add results from cognitive ability tests. Men who are physically or psychologically disabled, in addition to people living abroad, are exempted from the tests. The general ability score (IQ) is a measure building on results from three time-limited tests: Arithmetic (25 min), Word Similarities (8 min) and Figures (20 min). The Arithmetic test measures not only arithmetic ability, but also elementary algebraic and logical reasoning ability. The test is quite similar to the Arithmetic test in Wechsler Adult Intelligence Scale (WAIS). The Word Similarities Test is similar to the Vocabulary Test in WAIS and the Figures Test is similar to the Raven Progressive Matrices. In the middle of the 1990's the Arithmetic Test changed from having open answers to multiple choice, apart from that all the test have remained unchanged. General ability is measured on a stanine ("standard nine") score, which scale scores on a nine-point scale with a mean of 5 and a standard deviation of 2. For a more detailed description of general ability, see Sundet *et al*. (2004).

### *3.4 Social background*

Markussen and Røed (2017) find that annual earnings obtained around age fifty are most highly correlated with lifetime earnings. In order to obtain a consistent measure of parents' earnings which is representative for lifetime earnings, I have chosen the sum of the two parents' earnings during their age 50-54. Since I have access to earnings data in the period 1967-2015, I am able to identify every parent born in the period 1917-1958. Some parents are born after 1958 and for those parents I use the years closest to age 50-54. Earnings are adjusted for general wage growth. Finally, each

offspring's parental earnings rank is identified cohort by cohort by ranking parental earnings in 20 categories, which is done separately for male and female offspring.

### *3.5 Sample*

Since family-ties often are missing among immigrants, I will focus on Norwegian (native-born) residents in the analyses. I will also exclude people with missing birth weight or birth weight less than 500 gr. Since I focus on employment age 27-31 my sample is also restricted to individuals living in Norway when they are between 27 and 31 years old.[9]

## 4 Descriptive Statistics

In this section, I present some descriptive statistics about the main sample. Since the analyses is based on individuals born between 1967 and 1984, the descriptive statistics are also limited to this group. In order to look at developments over time, while maintaining sufficient statistical power, I divide the sample in three-years- intervals.

[9] 1,032,543 individuals are born between 1967 and 1984, out of which 974,149 individuals are natives (5.7 percent are excluded). Furthermore, 1,722 individuals are excluded due to missing birth weight or birth weight less than 500 gr. 1.0 percent died within the first year of life and 1.5 percent died between 1 and 31 years of age and additional 3.5 percent do not live in Norway at age 27-31. The final sample consists of 914,448 persons: 467,865 men and 446,583 women.

***Table 1a). Descriptive statistics for the main sample. Males****

| Year of birth | 1967-69 | 1970-72 | 1973-75 | 1976-78 | 1979-81 | 1982-84 |
|---|---|---|---|---|---|---|
| *N* | *91,656* | *88,533* | *80,517* | *70,577* | *68,664* | *67,918* |
| **Individual and background characteristics** | | | | | | |
| Birth weight (in gr) | 3,556 (553) | 3,562 (556) | 3,561 (553) | 3,581 (555) | 3,591 (558) | 3,579 (567) |
| Birth weight categories (per cent) | | | | | | |
| <2,500 gr | 3.41 | 3.50 | 3.44 | 3.34 | 3.24 | 3.49 |
| 2,500-3,000 gr | 9.98 | 9.63 | 9.68 | 8.89 | 8.79 | 8.19 |
| 3,000-3,500 gr | 30.07 | 29.82 | 29.94 | 29.49 | 28.55 | 28.65 |
| 3,500-4,500 gr | 52.55 | 53.13 | 53.10 | 54.03 | 55.11 | 54.33 |
| >4,500 gr | 3.98 | 3.93 | 3.84 | 4.25 | 4.31 | 4.33 |
| Percent multiple births | 1.77 | 1.73 | 1.72 | 1.68 | 1.72 | 1.97 |
| **Family characteristics** | | | | | | |
| Mother's age at birth | 26.51 (5.82) | 26.02 (5.38) | 25.95 (5.08) | 26.24 (4.99) | 26.62 (4.97) | 27.05 (4.99) |
| Number of siblings | 1.97 (1.27) | 1.83 (1.19) | 1.72 (1.12) | 1.70 (1.11) | 1.73 (1.10) | 1.77 (1.10) |
| **Military data** | | | | | | |
| Height (centimeters) | 179.82 (6.47) | 179.91 (6.45) | 179.86 (6.49) | 179.95 (6.54) | 179.96 (6.52) | 180.03 (6.51) |
| Missing height | 2.04 | 2.78 | 4.03 | 2.83 | 2.86 | 4.52 |
| Height group (percent) | | | | | | |
| <170 cm | 5.21 | 4.94 | 5.15 | 5.09 | 5.19 | 4.98 |
| 170-180 cm | 42.06 | 41.88 | 42.18 | 41.54 | 41.19 | 41.31 |
| 180-190 cm | 45.85 | 46.09 | 45.44 | 45.75 | 46.34 | 46.34 |
| >190 cm | 6.88 | 7.09 | 7.24 | 7.62 | 7.29 | 7.37 |
| BMI | 22.01 (2.97) | 22.29 (3.09) | 22.46 (3.16) | 22.59 (3.29) | 22.58 (3.46) | 23.02 (3.90) |
| BMI classification (percent) | | | | | | |
| Underweight (<18.5) | 7.08 | 6.00 | 5.25 | 5.08 | 6.18 | 5.95 |
| Normal weight (18.5-25) | 80.45 | 79.14 | 78.86 | 77.74 | 75.85 | 71.63 |
| Overweight (25-30) | 10.42 | 12.25 | 12.98 | 13.71 | 13.90 | 16.40 |
| Obesity (>30) | 2.05 | 2.61 | 2.90 | 3.47 | 4.07 | 6.02 |
| IQ (stanines) | 5.08 (1.83) | 5.16 (1.84) | 5.26 (1.83) | 5.24 (1.76) | 5.15 (1.72) | 5.12 (1.73) |
| Missing IQ (percent) | 6.28 | 7.00 | 11.00 | 6.02 | 7.42 | 14.06 |
| IQ classification (percent) | | | | | | |
| Low IQ (1-3) | 19.36 | 18.52 | 16.79 | 15.83 | 16.89 | 17.03 |
| Middle IQ (4-6) | 58.38 | 57.79 | 58.16 | 60.43 | 61.51 | 61.72 |
| High IQ (7-9) | 22.26 | 23.69 | 25.06 | 23.74 | 21.59 | 21.25 |
| **Outcomes** | | | | | | |
| Percent employed when 27-31 | 90.51 | 90.69 | 88.19 | 87.80 | 87.60 | 86.04 |
| Average annual earnings when 27-31 (NOK, 2017 value) | 489,638 (210,925) | 505,102 (220,734) | 482,477 (227,160) | 475,494 (229,152) | 484,036 (236,916) | 469,545 (240,795) |
| Percent with completed high school | 71.64 | 74.88 | 78.73 | 80.38 | 80.23 | 79.51 |
| Percent receiving permanent disability insurance when 19 | | | 0.41 | 0.53 | 0.63 | 0.62 |

*The main sample is described in section 2.3. Standard deviations are given in parenthesis.

***Table 1b). Descriptive statistics for the main sample. Females****

| Year of birth | 1967-69 | 1970-72 | 1973-75 | 1976-78 | 1979-81 | 1982-84 |
|---|---|---|---|---|---|---|
| *N* | *87,642* | *84,659* | *77,174* | *67,344* | *65,555* | *64,209* |
| **Individual and background characteristics** | | | | | | |
| Birth weight (in gr) | 3,425 (527) | 3,435 (522) | 3,432 (525) | 3,454 (529) | 3,467 (528) | 3,457 (536) |
| Birth weight categories (percent) | | | | | | |
| <2,500 gr | 4.18 | 3.91 | 3.99 | 3.92 | 3.65 | 3.90 |
| 2,500-3,000 gr | 13.75 | 13.38 | 13.47 | 12.56 | 12.00 | 12.58 |
| 3,000-3,500 gr | 36.54 | 36.26 | 36.68 | 35.40 | 35.37 | 35.02 |
| 3,500-4,000 gr | 43.53 | 44.53 | 43.94 | 46.05 | 46.77 | 46.34 |
| >4,500 gr | 1.99 | 1.91 | 1.93 | 2.06 | 2.20 | 2.16 |
| Percent multiple births | 1.94 | 1.73 | 1.81 | 1.99 | 1.89 | 1.94 |
| **Family characteristics** | | | | | | |
| Mother's age at birth | 26.52 (5.82) | 26.03 (5.36) | 25.94 (5.08) | 26.20 (5.00) | 26.61 (4.97) | 27.03 (4.96) |
| Number of siblings | 1.98 (1.27) | 1.84 (1.19) | 1.73 (1.13) | 1.69 (1.10) | 1.73 (1.10) | 1.76 (1.09) |
| **Outcomes** | | | | | | |
| Percent employed when 27-31 | 73.76 | 77.22 | 77.88 | 80.18 | 82.11 | 81.18 |
| Average annual earnings when 27-31 (NOK, 2017 value) | 317,309 (186,933) | 334,705 (188,546) | 337,832 (185,261) | 344,704 182,009) | 357,470 (183,032) | 351,167 (181,798) |
| Percent with completed high school | 69.65 | 75.93 | 81.51 | 85.64 | 86.75 | 85.92 |
| Percent receiving perma nent disability insurance when 19 | | | 0.30 | 0.44 | 0.50 | 0.54 |

*The main sample is described in section 3.5. Standard deviations are given in parenthesis.

Table 1a) and 1b) show descriptive statistics for men and women. During the period 1967-1984 there has been a gradual decline in the number of births in Norway, which is reflected in a shrinking number of observations in the tables. Mean birth weight among men (women) born in the period 1967-69 is 3,556 (3,425) gr while the mean is approximately 30 gr higher among men and women born in the latest cohorts. The fraction with low birth weight (<2,500 gr) has been quite stable among men (3.25-3.50 percent). Among women, the fraction with low birth weight has fluctuated between 4.2 and 3.7 percent (the fraction was highest for women born in the period 1967-69 and lowest for those born 1979-81). Being born with high birth weight (>4,500 gr) has become slightly more common during the period 1967-1984; it accounts for 4.0 (2.0) percent of men

(women) born in the beginning of the period while the fraction was 4.3 (2.2) percent at the end of the period.

Mother's age at birth has increased by about 0.5 years while the number of siblings has decreased from almost 2.0 to 1.8.[10] Military data shows that mean BMI for men has increased from 22 to 23 and the fraction with obesity has almost tripled; among the earliest cohorts 2.1 percent are classified as obese while it accounts for 6 percent among the latest cohorts. When it comes to ability, the mean stanine score is highest among men born in the middle of the 1970's while the mean among the earliest and latest cohorts is almost the same.[11] I do not have information about BMI and IQ for every men in Norway and one reason is that some of them are permanently disabled. According to Eide *et al*. (2005), 1.4 percent are classified as disabled because of chronic disease or birth defects. The fraction not appearing at the draft board is somewhat higher, especially among the last cohorts.[12] Some men are registered with information about BMI, but not with IQ – approximately 7 percent among the group of cohorts born 1973 -75 and 10 percent among men born 1982-84, among the other groups of cohorts it accounts for about 4 percent. Very few (0.1-0.5 percent) are registered with IQ but not with BMI.

The fraction employed and the development over time is quite different for men and women. Among men born in the period 1967-69, 90.5 percent were employed at the age of 27-31 and the rate decreased to 86.0 percent among men born between 1982 and 1984. Among women the fraction employed has increased from 73.8 percent to 81.2 percent, but the fraction employed was even higher (82.1 percent) among the cohorts born in 1979-81. The fraction with completed high school is remarkably higher among cohorts born in the middle of the 1970's compared to those born at the end of the 1960's, but the fraction has remained quite stable after that. The probability of receiving permanent disability insurance at age 19 increases from 0.4 (0.3) percent for

[10] Siblings are identified via mothers. Every link to children is identified, also for children born before 1967 child, meaning that I count every siblings having the same mother.

[11] An increase in intelligence test scores is observed in many countries and is known as the Flynn effect, but based on Norwegian military data, Sundet *et al*. (2004) conclude that the Flynn effect may have come to an end in Norway since the means of the scores stopped to increase in the mid1990s (corresponding to men born in the mid-1970s).

[12] The fraction of men in our sample with no information from the military service is 1.9 percent for cohorts born 1967-69, 2.6 percent for cohorts born 1970-72, 2.5 percent for cohorts born 1973-75, 2.2 percent for cohorts born 1976-78, 2.1 percent for cohorts born 1979-81 and 3.7 percent for cohorts born 1982-84.

men (women) born in the period 1973-75 to 0.6 (0.5) percent for men (women) born in the period 1982-84.[13]

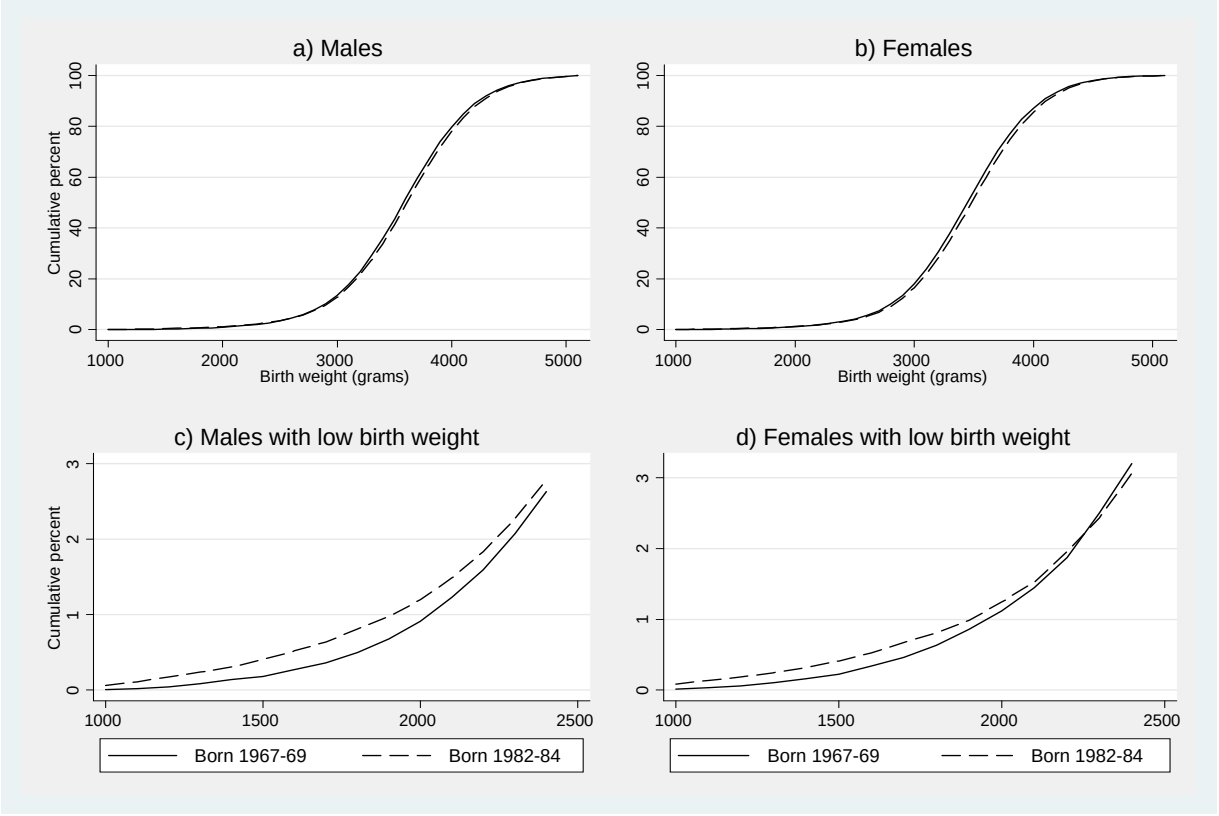



***Figure 2. Cumulative distribution of birth weight among natives living in Norway at the age of 27-31.***

*Note; Birth weight is divided in 100 gr- categories and birth weight below 1000 gr and above 5000 gr are grouped. Low birth weight is birth weight <2,500 gr.*

Figure 2 shows the cumulative distribution of birth weight among the earliest and latest cohorts; panel a) and b) show the whole distribution while I zoom in on individuals born with low birth weight in panel c) and d). In the middle of the distribution the latest cohorts lie to the right of the earliest cohorts meaning that people in the latest cohorts are somewhat heavier at birth, but the difference is quite small. On the very bottom of the birth weight distribution the development is

[13] Brage and Thune (2015) show that persons granted permanent disability insurance at this stage in life are dominated by persons with mental retardation and congenital malformations, and part of the recent rice of young benefit receivers is attributed to an increasing fraction of the adult population with severe and chronic health impairment

different. As already shown in table 1, the overall fraction with low birth weight (<2,500 gr) has not increased. However, being born with very low birth weight (<1,500 gr) accounts for 0.18 (0.23) percent of men (women) belonging to the earliest cohorts while the fraction increased to 0.41 percent among both men and women belonging to the youngest cohorts. This is due to improvements in neonatal health care making it possible for lighter babies to survive.[14]

***Table 2a) Parental earnings rank by health and ability indicators. Males***

| Year of birth | 1967-69 | 1970-72 | 1973-75 | 1976-78 | 1979-81 | 1982-84 |
|---|---|---|---|---|---|---|
| *N* | *91,656* | *88,533* | *80,517* | *70,577* | *68,664* | *67,918* |
| **Birth weight categories** | | | | | | |
| <2,500 gr | 9.55 | 9.67 | 9.76 | 9.41 | 9.37 | 9.44 |
| 2,500-3,000 gr | 9.79 | 9.82 | 9.70 | 9.76 | 9.66 | 9.69 |
| 3,000-3,500 gr | 10.29 | 10.23 | 10.27 | 10.26 | 10.20 | 10.27 |
| 3,500-4,500 gr | 10.71 | 10.68 | 10.74 | 10.74 | 10.72 | 10.69 |
| >4,500 gr | 10.84 | 10.83 | 10.79 | 10.80 | 10.83 | 10.99 |
| Missing height | 8.84 | 8.78 | 8.93 | 8.83 | 8.56 | 8.28 |
| **Height group** | | | | | | |
| <170 cm | 9.22 | 9.23 | 9.33 | 9.14 | 9.12 | 9.31 |
| 170-180 cm | 10.09 | 10.12 | 10.20 | 10.18 | 10.14 | 10.20 |
| 180-190 cm | 10.88 | 10.81 | 10.86 | 10.84 | 10.84 | 10.84 |
| >190 cm | 11.36 | 11.35 | 11.29 | 11.34 | 11.24 | 11.49 |
| **BMI classification** | | | | | | |
| Underweight (<18.5) | 10.23 | 10.28 | 10.18 | 10.24 | 10.11 | 9.98 |
| Normal weight (18.5-25) | 10.66 | 10.67 | 10.75 | 10.73 | 10.73 | 10.81 |
| Overweight (25-30) | 9.72 | 9.76 | 9.79 | 9.86 | 9.90 | 10.14 |
| Obesity (>30) | 8.74 | 8.70 | 8.66 | 8.72 | 8.76 | 9.07 |
| Missing IQ (percent) | 8.87 | 8.91 | 9.53 | 8.93 | 9.18 | 9.36 |
| **IQ classification** | | | | | | |
| Low IQ (1-3) | 8.29 | 8.19 | 8.32 | 8.08 | 8.17 | 8.28 |
| Middle IQ (4-6) | 10.51 | 10.44 | 10.40 | 10.44 | 10.48 | 10.59 |
| High IQ (7-9) | 12.68 | 12.65 | 12.54 | 12.55 | 12.57 | 12.60 |

*Note; Parental earnings rank ranges from 1 to 20. The unconditional mean is 10.5.*

[14] When looking at the birth weight distribution among newborn children (and not only those surviving and living in Norway at the age of 27-31) the fraction of babies with very low birth weight (<1,500 gr) is quite stable during the period 1967-1984, while the fraction with low birth weight (<2,500 gr) decreased.

| *Table 2b) Parental earnings rank by health indicators. Females* | | | | | | |
|---|---|---|---|---|---|---|
| Year of birth | 1967-69 | 1970-72 | 1973-75 | 1976-78 | 1979-81 | 1982-84 |
| *N* | *87,642* | *84,659* | *77,174* | *67,344* | *65,555* | *64,209* |
| **Birth weight categories** | | | | | | |
| <2,500 gr | 9.67 | 9.51 | 9.52 | 9.31 | 9.30 | 9.44 |
| 2,500-3,000 gr | 9.91 | 9.87 | 9.78 | 9.85 | 9.66 | 9.80 |
| 3,000-3,500 gr | 10.38 | 10.43 | 10.35 | 10.38 | 10.39 | 10.35 |
| 3,500-4,500 gr | 10.77 | 10.77 | 10.82 | 10.77 | 10.85 | 10.88 |
| >4,500 gr | 10.98 | 10.84 | 10.85 | 10.73 | 10.93 | 10.66 |

*Note; Parental earnings rank ranges from 1 to 20. The unconditional mean is 10.5.*

There is a clear social gradient in the various health and ability indicators I use in this paper. Table 2a) and 2b) show that parental earnings rise with birth weight across the four lightest birth weight categories, while there is only minor differences between the two heaviest birth weight groups. Health – and ability measures (available only for men) also show the same; being shorter, deviations from normal weight as well as lower ability are associated with lower parental earnings. In addition, mean parental earnings among men with missing information from the military service is lower compared to men registered by the military service, which is plausible since disabled men are not obliged to attend. The gradient of the different indicators is quite stable over time.

Figures 3 and 4 illustrate how the fraction with presumably poor health and ability varies with parental earnings and how this relationship has changed from the earliest to the latest cohorts. Lower birth weight is more common the lower the parental earnings are, and belonging to the poorest 5 percent families compared to the 5 percent richest families approximately doubles the risk of being born with birth weight below 2,500 gr. The indicators measured by the military service show differences by social background to various extent (figure 4). For men born in the period 1967-69, the prevalence of low ability is 5 percent for men with the richest parents compared to nearly 35 percent of men with the poorest ones. The fraction of men with low ability is somewhat lower for the last cohorts, but the gradient is quite similar. For men belonging to the earliest cohorts, overweight accounts for almost 14 percent of men with the least favourable social background and 7 percent of the richest ones. The prevalence of overweight is approximately 6 percentage point higher for men born in the last cohorts and this increase is quite constant across social background. Also, being obese has become more common among every group of men, but measured in percentage point the increase has been highest in the lowest economic class. When it comes to underweight

the difference between economic classes is not very large, but the social gradient seems to be somewhat clearer for the latest cohorts.

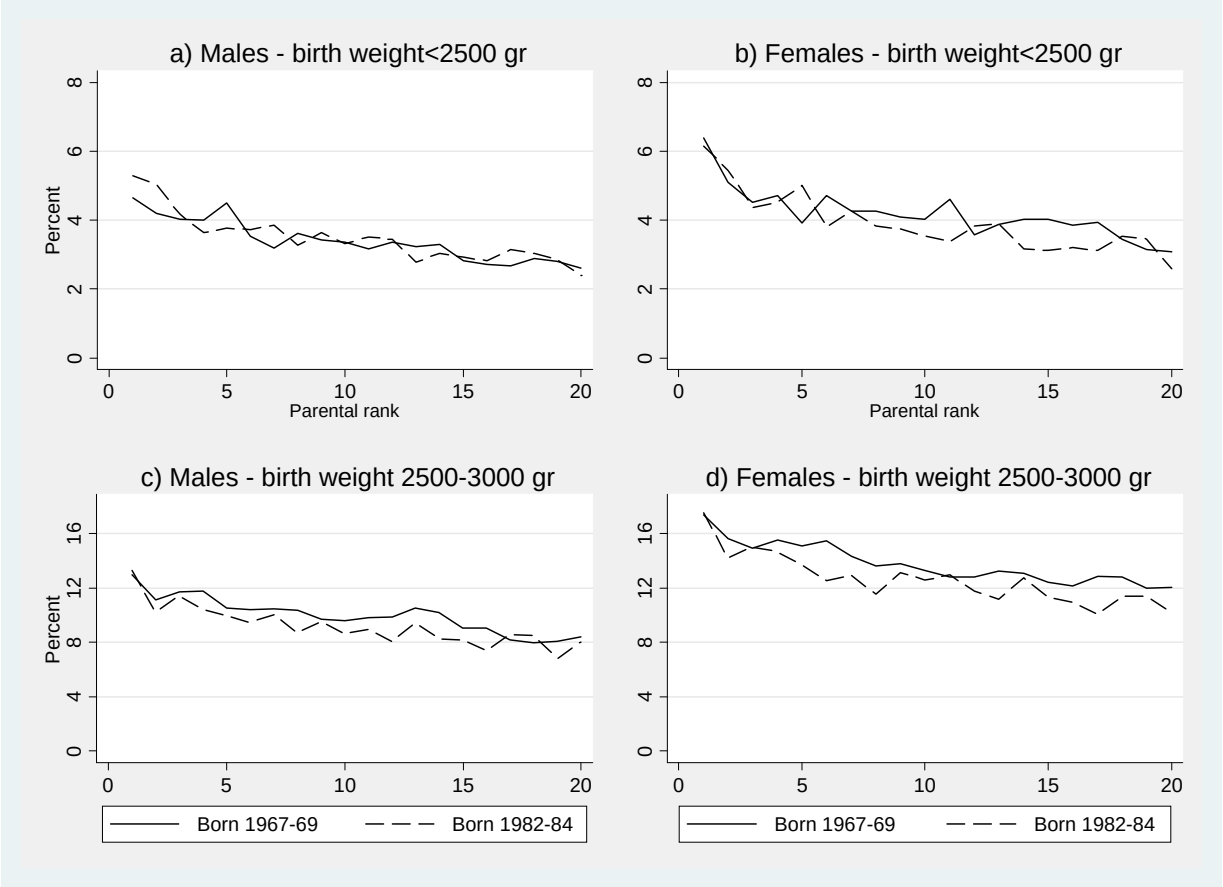



***Figure 3. Incidence of lower birth weight by parental earnings rank (divided in 20 bins).***

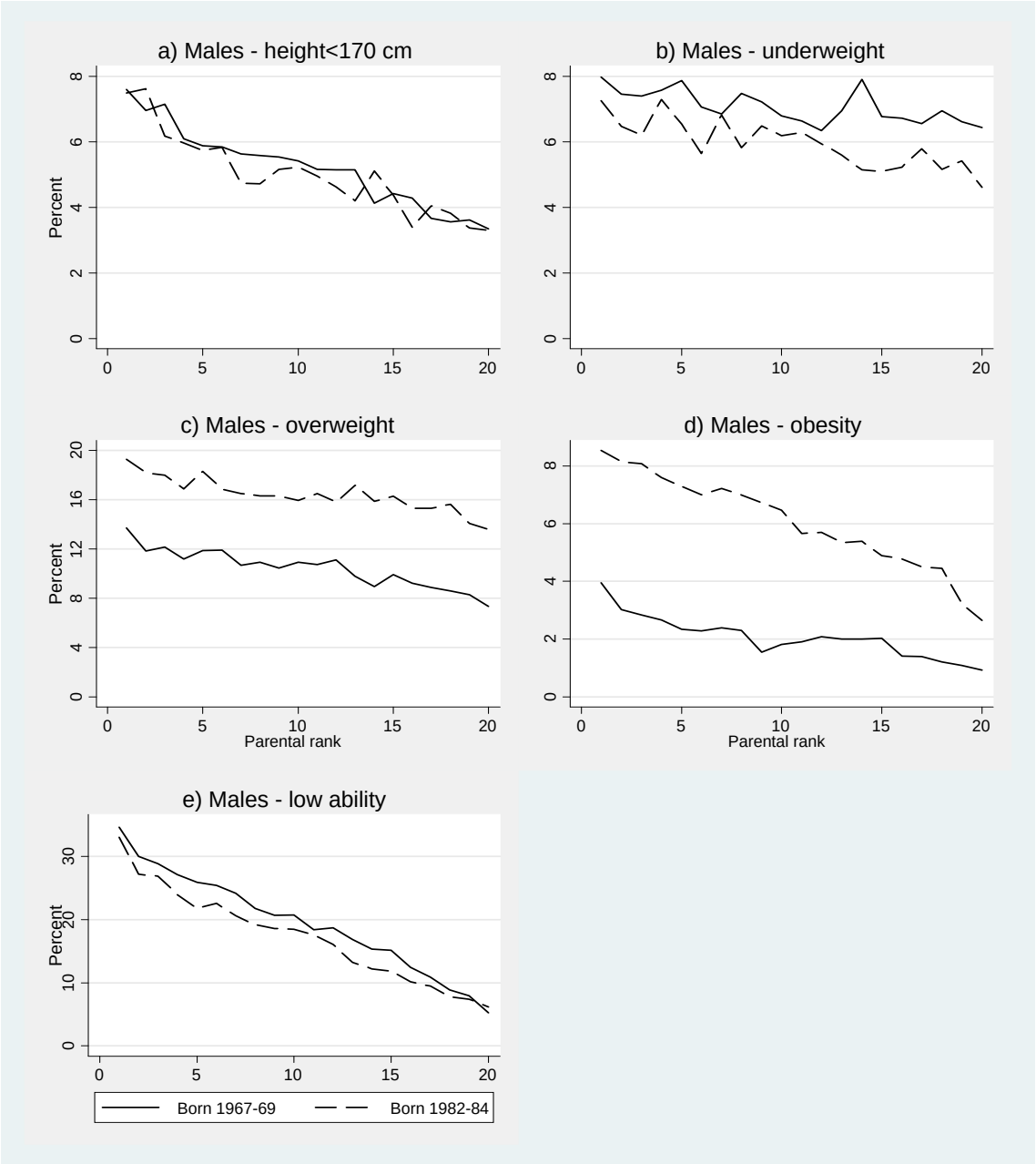



***Figure 4. Incidence of "poor" health/ability by parental earnings rank (divided in 20 bins). Males only.***

*Note; underweight=BMI<18.5-25, overweight=BMI 25-30, obesity=BMI>30, low ability=stanine score 1-3.*

A social gradient in the health- and ability- indicators is also present when using parental education as measure of social background, see appendix table 1a) and 1b). Furthermore, the tables also show problems arising when comparing the education level across different birth cohorts since educational attainment increases substantially over time.

## 5 Empirical strategy

The focus in this article is whether health, ability and social background have become more crucial in order to be successfully integrated in the labor market. My main dependent variables ($y_i$) are log earnings and employment by young adulthood (age 27-31). Other outcome measures are mortality, high school completion and disability. My baseline specification is as follows:

$$y_{ic} = \alpha_c + \beta_c H_i + \delta_c x_i + \varepsilon_i, \quad (1)$$

where $i$ indexes the individual and c indexes birth cohort. $H_i$ is a vector of indicators of health, cognitive ability and social background. $x_i$ is a vector of individual specific control variables like birth order, month of birth, multiple birth, number of siblings and mothers age at birth.[15] The models are estimated separately for males and females because the control variables may have heterogeneous impact across genders due to differences in mean outcomes and birth weight. Equation (1) is estimated using Ordinary Least Squares (OLS).

In order to explore whether the impacts of health, ability, and social background have changed over time, I assume that the cohort-effect, c, is constant over three-year intervals, meaning that I run 6 separate regressions; 1967-69, 1970-72, 1973-75, 1976-78, 1979-81 and 1982-84. Doing this, I avoid the imposition of arbitrary functional form restrictions on the way causal impacts have developed over time. To check whether any estimated changes of the coefficients are significant I also run a regression where all cohorts are included in the same model.

In robustness analyses, I estimate models with siblings and twins fixed effects. Many studies about birth weight have used sibling fixed effects in order to control for socioeconomic background, but within siblings there may still be genetic factors which potentially explain differences in birth weight. In order to deal with this problem, a number of studies have used twin fixed effects (Bharadwaj *et al*., 2018; Figlio *et al*., 2014; Royer, 2009; Black *et al*., 2007; Almond *et al*., 2005; Behrman and Rosenzweig, 2004) where the variation in birth weight between twins are taken as random and suggested due to variation in nutritional intake. Family fixed effects give the following specification of the model:

[15] Higher birth order is associated with higher birth weight (Seidman *et al.*, 1988) while higher birth order is negatively correlated with outcomes later in life (Black *et al*., 2005). Number of siblings and age of mother are also suggested to be associated with adult outcomes (see f.ex. Currie and Hyson, 1999b)

$$y_{icj} = \alpha_c + \beta_c H_i + \delta_c x_i + \mu_j + \varepsilon_i, \quad (3)$$

where $j$ indexes mothers and $y_{icj}$ is thus the outcome of one individual $i$ born to mother $j$ in cohort $c$. $\mu_j$ is the sibling/twin fixed effects. I estimate models for twins and non-twins separately and when I refer to siblings I mean siblings which are not twins. When estimating sibling fixed effects the control variables are birth order, month and year of birth. When estimating twin fixed effects $\beta$ is identified based on differences in $H$ within twin pairs meaning that any family and birth-specific confounders are controlled out.

# 6 Main results

In this section, I present the estimated effects of health, ability, and social background on labor market performance at age 27-31. I start by presenting models where the vulnerability indicators include birth weight and social background (SES). Thereafter, I show results from models where health and ability-measures from the Norwegian military service are included. The latter results are available for men, only.

Since the functional form of the relationship between birth weight and earnings/employment is a priori unknown, I will use dummies for different birth weight categories and aggregate data into 500 gr strata. Birth weight less than 2,500 gr and birth weight more than 4,500 gr are grouped and in the analyzes I also merge birth weight 3,500-4,000 gr with 4,000-4,500 gr since there are no statistical significant differences between those groups.[16] Hence, I have 5 categories: <2,500 gr, 2,500-3,000 gr, 3,000-3,500 gr, 3,500-4,500 gr and >4,500 gr. Birth weight 3,500-4,500 gr are chosen as the reference group, intended to represent the group with the most favorable health condition. Similarly, social background is coded by means of 20 dummy variables, each representing the vigintile in the parental earnings distribution (see section 3). Here, I use the two medium vigintiles (10-11) as references, and to avoid too long tables, I report the estimated coefficients

[16] I started out by running models with 6 different birth weight-categories:<2,500 gr, 2,500-3,000 gr, 3,000-3,500 gr, 3,500-4,000 gr, 4,000-4,500 gr and >4,500 gr. Models on employment/earnings does not find any statistical differences between the groups 3,500-4,000 and 4,000-4,500 gr.

associated with growing up in the four lowest vigintiles only (an overview of the complete social gradient estimates is provided in the appendix).

In this section, the results are presented as follows; I start with tables reporting detailed estimation results, and based on this results I present the main results graphically, with a focus on the effects of belonging to the groups with lowest birth weight and lowest socioeconomic status (SES). Additionally, I show results from tests of whether changes in the effects of poor health and low SES are significant.

The labor market outcomes I focus on are log labor earnings and employment. As the causal variables of interest turn out to have a significant influence on employment, an analysis of earnings conditional on employment will be hard to interpret. In the main part of the analysis, I therefore focus on an unconditional earnings analysis, including also the non-employed. To circumvent the resultant problem with zero earnings, I add a small earnings amount for everyone (see also section 2.1).[17]

[17] When adding a smaller amount (0.1 G per year instead of 0.2 G per year) the point estimates become somewhat higher compared to what I present in the tables.

***Table 3. Estimation results on log earnings***

| Year of birth | 1967-69 | 1970-72 | 1973-75 | 1976-78 | 1979-81 | 1982-84 |
|---|---|---|---|---|---|---|
| **Males** | | | | | | |
| *N* | *91,656* | *88,533* | *80,516* | *70,576* | *68,664* | *67,918* |
| **Birth weight** | | | | | | |
| <2,500 gr | -0.161*** | -0.146*** | -0.145*** | -0.186*** | -0.180*** | -0.203*** |
| | (0.013) | (0.014) | (0.015) | (0.018) | (0.019) | (0.019) |
| 2,500-3,000 gr | -0.080*** | -0.080*** | -0.082*** | -0.095*** | -0.063*** | -0.112*** |
| | (0.008) | (0.008) | (0.010) | (0.011) | (0.012) | (0.012) |
| 3,000-3,500 gr | -0.042*** | -0.042*** | -0.044*** | -0.045*** | -0.050*** | -0.043*** |
| | (0.005) | (0.005) | (0.006) | (0.007) | (0.007) | (0.008) |
| 3,500-4,500 gr | Ref. | Ref. | Ref. | Ref. | Ref. | Ref. |
| >4,500 gr | 0.021* | 0.001 | 0.037*** | 0.033** | 0.047*** | 0.015 |
| | (0.012) | (0.012) | (0.014) | (0.015) | (0.016) | (0.017) |
| **Parental earnings rank** (coefficients reported for the four lowest ranks only) | | | | | | |
| 1 | -0.393*** | -0.438*** | -0.492*** | -0.511*** | -0.566*** | -0.604*** |
| | (0.013) | (0.013) | (0.015) | (0.017) | (0.017) | (0.018) |
| 2 | -0.240*** | -0.231*** | -0.255*** | -0.246*** | -0.288*** | -0.327*** |
| | (0.013) | (0.013) | (0.015) | (0.017) | (0.017) | (0.018) |
| 3 | -0.161*** | -0.149*** | -0.225*** | -0.201*** | -0.247*** | -0.260*** |
| | (0.013) | (0.013) | (0.015) | (0.016) | (0.017) | (0.018) |
| 4 | -0.143*** | -0.140*** | -0.161*** | -0.174*** | -0.233*** | -0.201*** |
| | (0.013) | (0.013) | (0.015) | (0.017) | (0.017) | (0.018) |
| **Females** | | | | | | |
| *N* | *87,642* | *84,658* | *77,174* | *67,344* | *65,554* | *64,208* |
| **Birth weight** | | | | | | |
| <2,500 gr | -0.192*** | -0.156*** | -0.198*** | -0.181*** | -0.174*** | -0.190*** |
| | (0.016) | (0.016) | (0.016) | (0.017) | (0.018) | (0.018) |
| 2,500-3,000 gr | -0.105*** | -0.088*** | -0.096*** | -0.093*** | -0.091*** | -0.071*** |
| | (0.009) | (0.009) | (0.009) | (0.010) | (0.010) | (0.010) |
| 3,000-3,500 gr | -0.044*** | -0.044*** | -0.045*** | -0.030*** | -0.031*** | -0.033*** |
| | (0.005) | (0.006) | (0.007) | (0.007) | (0.007) | (0.007) |
| 3,500-4,500 gr | Ref. | Ref. | Ref. | Ref. | Ref. | Ref. |
| | 0.005 | -0.017 | -0.014 | -0.008 | -0.014 | -0.010 |
| | (0.012) | (0.021) | (0.022) | (0.022) | (0.022) | (0.022) |
| **Parental earnings rank** | | | | | | |
| 1 | -0.447*** | -0.515*** | -0.527*** | -0.564*** | -0.614*** | -0.648*** |
| | (0.016) | (0.016) | (0.016) | (0.017) | (0.017) | (0.018) |
| 2 | -0.270*** | -0.259*** | -0.294*** | -0.298*** | -0.326*** | -0.366*** |
| | (0.016) | (0.015) | (0.016) | (0.017) | (0.017) | (0.018) |
| 3 | -0.202*** | -0.214*** | -0.238*** | -0.219*** | -0.274*** | -0.318*** |
| | (0.016) | (0.016) | (0.016) | (0.017) | (0.017) | (0.018) |
| 4 | -0.194*** | -0.169*** | -0.200*** | -0.219*** | -0.227*** | -0.253*** |
| | (0.016) | (0.016) | (0.016) | (0.017) | (0.017) | (0.018) |

*Note; *p<0.10, **p<0.05, ***p<0.01. Standard error in parenthesis. Parental earnings are divided in 20 bins (vigintiles). Vigintile 10 and 11 is the reference group. The following control variables are included in the models (number of categories for categorical variables in parentheses): multiple birth (1), birth order (4), siblings (4), mother's age at birth (6), month of birth (12).*

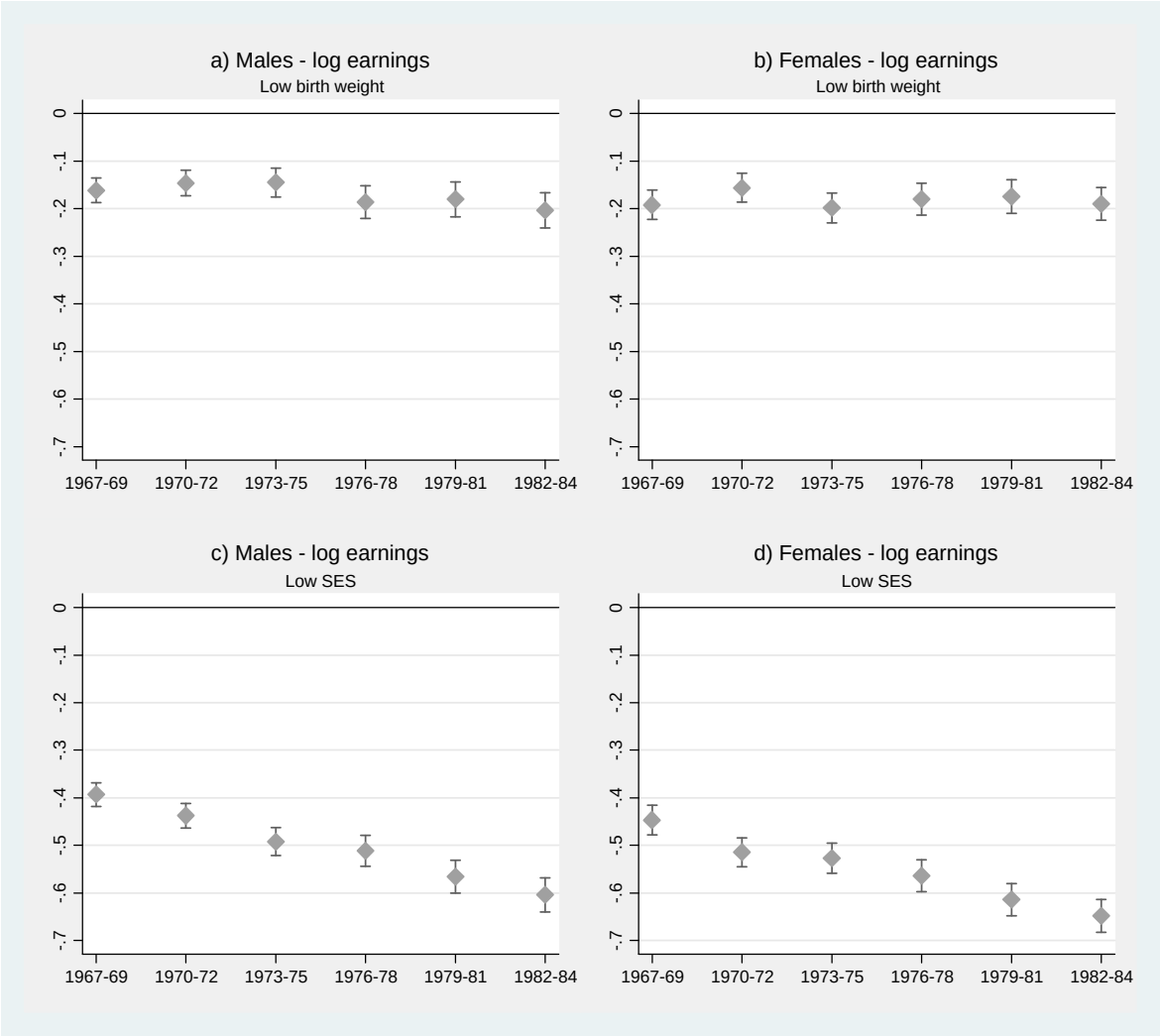



***Figure 5. Estimated effects of low birth weight (<2,500 gr) and low SES. Dependent variable is log earnings at age 27-31.***

*Note; The reference group for birth weight is 3,500-4,500 gr. Socioeconomic status (SES) is divived in 20 vigintiles based on parental earnings' rank. Vigintile 10 and 11 is the reference group. Low SES is vigintile 1, which is the five percent with lowest parental earnings. Point estimates with confidence interval are shown in the figure. The list of control variables is described in the note to table 3.*

***Table 4. t-test of significance of low birth weight (<2,500 gr) and low SES between each group of cohorts. Dependent variable is log earnings, as shown in figure 5.***

| | Males | | | | | Females | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Year of birth | 1970-72 | 1973-75 | 1976-78 | 1979-81 | 1982-84 | 1970-72 | 1973-75 | 1976-78 | 1979-81 | 1982-84 |
| **Low birth weight** | | | | | | | | | | |
| 1967-69 | 0.73 | 0.73 | -1.11 | -0.84 | -1.88* | 1.67* | - 0.28 | 0.50 | 0.75 | 0.09 |
| 1970-72 | | 0.02 | -1.79* | -1.50 | -2.55** | | -1.87* | -1.05 | -0.76 | -1.42 |
| 1973-75 | | | -1.79* | -1.50 | -2.53** | | | 0.73 | 0.97 | 0.34 |
| 1976-78 | | | | 0.24 | -0.72 | | | | 0.25 | -0.37 |
| 1979-81 | | | | | -0.95 | | | | | -0.60 |
| **Low SES** | | | | | | | | | | |
| 1967-69 | -2.23 ** | -4.80 *** | -5.51 *** | -7.99 *** | -9.78 *** | -3.12 *** | -3.60 *** | -5.04 *** | -7.14 *** | - 8.55 *** |
| 1970-72 | | -2.60 *** | -3.39 *** | -5.88 *** | -7.66 *** | | -0.55 | -2.09 ** | -4.20 *** | -5.62 *** |
| 1973-75 | | | -0.88 | -3.33 *** | -5.05 *** | | | -1.52 | -3.60 *** | -4.99 *** |
| 1976-78 | | | | -2.38 ** | -4.07 *** | | | | -2.03 ** | -3.39 *** |
| 1979-81 | | | | | -1.68* | | | | | -1.36 |

*Note; *p<0.10, **p<0.05, ***p<0.01.*

Table 3 shows that earnings for both men and women increase with birth weight, at least up to around 4,500 gr. Concentrating on the two lowest birth weight categories, figure 5 illustrates that being born with birth weight <2,500 gr, compared to the reference category of birth weight 3,500-4,500 gr, decreases earnings by 15-20 percent for both genders. For men, the negative effect of low birth weight is stronger among cohorts born after the middle of the 1970's compared to those born earlier, and according to table 4 several of the differences between the latest and earliest cohorts are statistically significant.

Estimates from the four lowest parental earnings ranks are presented in table 4 while all point estimates for parental earnings ranks among the last cohorts (1982-84) and the earliest cohorts (1967-69) are illustrated in appendix figure 1. There is a clear social gradient in earnings, particularly in the bottom of the parental earnings rank where the impact of the gradient also has become more prominent over time. For example, the estimated coefficients show that coming from the very poorest economic class, compared to middle class decreases male (female) earnings by 32 (36)

percent for the earliest cohorts while the difference has increased to 45 (48) percent for the last cohorts.[18]

Figure 5 and table 3 present results from models where all individuals are included, also those having earnings below the threshold to be classified as employed (including individuals with zero earnings).[19] When restricting the sample to employed individuals the effect on earnings is much smaller compared to what I find using the whole sample and there is no particular change in the effect over time. According to the conditional point estimates reported in appendix figure 2, being born with low birth weight decreases earnings by only 2-4 percent among both genders, while being born into the 5 percent poorest families decreases earnings by 8-12 percent among men and 7-9 percent among women.[20]

[18] The log coefficient is transformed to percentage changes by using the formula ($exp^{\beta}$ - 1)

[19] 1.4 percent of men born in the period 1967-69 have zero earnings. The fraction increases over time and among men born 1982-84 3.0 percent have zero earnings. Among women the fraction is quite constant at 3 percent.

[20] This is consistent with findings from Currie and Madrian (1999). In their literature review, they show that a number of studies have found that health is a more important determinant on hours worked than on wages.

| *Table 5. Estimation results on employment* | | | | | | |
|---|---|---|---|---|---|---|
| Year of birth | 1967-69 | 1970-72 | 1973-75 | 1976-78 | 1979-81 | 1982-84 |
| **Males** | | | | | | |
| *N* | *91,656* | *88,533* | *80,516* | *70,576* | *68,664* | *67,918* |
| *Empl. rate ref. grp* | *91.40* | *91.56* | *89.14* | *88.76* | *88.66* | *87.06* |
| **Birth weight** | | | | | | |
| <2,500 gr | -0.051*** | -0.043*** | -0.043*** | -0.061*** | -0.069*** | -0.063*** |
| | (0.006) | (0.006) | (0.006) | (0.007) | (0.007) | (0.008) |
| 2,500-3,000 gr | -0.021*** | -0.023*** | -0.024*** | -0.030*** | -0.023*** | -0.033*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| 3,000-3,500 gr | -0.014*** | -0.012*** | -0.015*** | -0.013*** | -0.016*** | -0.012*** |
| | (0.002) | (0.002) | (0.003) | (0.003) | (0.003) | (0.003) |
| 3,500-4,500 gr | Ref. | Ref. | Ref. | Ref. | Ref. | Ref. |
| >4,500 gr | 0.006 | -0.004 | 0.011* | 0.010 | 0.013*** | 0.010 |
| | (0.005) | (0.005) | (0.006) | (0.006) | (0.006) | (0.007) |
| **Parental earnings rank** | | | | | | |
| 1 | -0.132*** | -0.149*** | -0.168*** | -0.178*** | -0.185*** | -0.201*** |
| | (0.005) | (0.005) | (0.006) | (0.007) | (0.007) | (0.007) |
| 2 | -0.080*** | -0.079*** | -0.084*** | -0.079*** | -0.092*** | -0.105*** |
| | (0.005) | (0.005) | (0.006) | (0.007) | (0.007) | (0.007) |
| 3 | -0.046*** | -0.045*** | -0.079*** | -0.064*** | -0.076*** | -0.085*** |
| | (0.005) | (0.005) | (0.006) | (0.007) | (0.007) | (0.007) |
| 4 | -0.045*** | -0.041*** | -0.061*** | -0.051*** | -0.066*** | -0.060*** |
| | (0.005) | (0.005) | (0.006) | (0.007) | (0.007) | (0.007) |
| **Females** | | | | | | |
| *N* | *87,642* | *84,658* | *77,174* | *67,344* | *65,554* | *64,208* |
| *Empl. rate ref. grp* | *73.37* | *78.61* | *79.48* | *81.52* | *83.39* | *82.60* |
| **Birth weight** | | | | | | |
| <2,500 gr | -0.063*** | -0.049*** | -0.076*** | -0.061*** | -0.056*** | -0.071*** |
| | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) |
| 2,500-3,000 gr | -0.039*** | -0.031*** | -0.033*** | -0.036*** | -0.031*** | -0.025*** |
| | (0.005) | (0.004) | (0.005) | (0.005) | (0.005) | (0.005) |
| 3,000-3,500 gr | -0.018*** | -0.016*** | -0.015*** | -0.008*** | -0.010*** | -0.011*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| 3,500-4,500 gr | Ref. | Ref. | Ref. | Ref. | Ref. | Ref. |
| | 0.006 | -0.008 | -0.008 | -0.009 | 0.004 | -0.003 |
| | (0.011) | (0.010) | (0.011) | (0.011) | (0.010) | (0.011) |
| **Parental earnings rank** | | | | | | |
| 1 | -0.183*** | -0.208*** | -0.207*** | -0.225*** | -0.245*** | -0.252*** |
| | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) |
| 2 | -0.113*** | -0.107*** | -0.126*** | -0.128*** | -0.142*** | -0.156*** |
| | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) |
| 3 | -0.086*** | -0.092*** | -0.101*** | -0.095*** | -0.113*** | -0.128*** |
| | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) |
| 4 | -0.083*** | -0.072*** | -0.082*** | -0.091*** | -0.096*** | -0.107*** |
| | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) |

*Note; *p<0.10, **p<0.05, ***p<0.01. Standard error in parenthesis. Parental earnings is divided in 20 bins (vigintiles). Vigintile 10 and 11 is the reference group. The control variables are the same as those reported for table 3.*

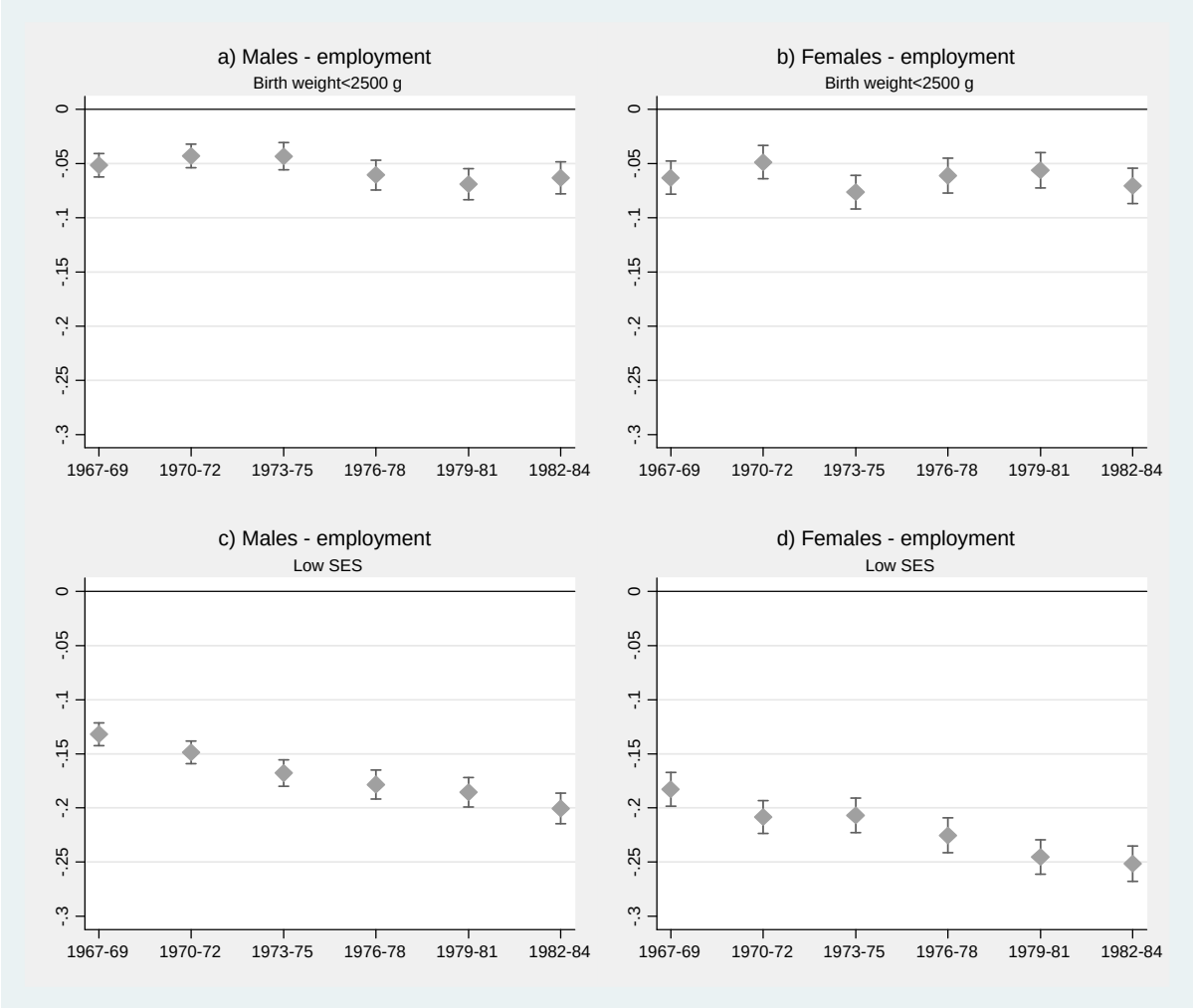



***Figure 6. Estimated effects of low birth weight (<2,500 gr) and low SES. Dependent variable employment.***
*Note; The reference group for birth weight is 3,500-4,500 gr. Socioeconomic status (SES) is divived in 20 vigintiles based on parental earnings' rank. Vigintile 10 and 11 is the reference group. Low SES is vigintile 1, which is the five percent with lowest parental earnings. Point estimates with confidence interval are shown in the figure. The control variables are the same as those reported for table 3.*

***Table 6. t-test of significance of low birth weight (<2,500 gr) and low SES between each group of cohorts. Dependent variable is employment, as shown in figure 6.***

| | Males | | | | | Females | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Year of birth | 1970-72 | 1973-75 | 1976-78 | 1979-81 | 1982-84 | 1970-72 | 1973-75 | 1976-78 | 1979-81 | 1982-84 |
| **Low birth weight** | | | | | | | | | | |
| 1967-69 | 1.02 | 0.93 | -1.00 | -1.89* | -1.30 | 1.37 | -1.24 | 0.14 | 0.59 | -0.67 |
| 1970-72 | | -0.07 | -1.95* | -2.82*** | -2.24** | | -2.53** | -1.12 | -0.64 | -1.90* |
| 1973-75 | | | -1.84* | -2.69*** | -2.12** | | | 1.28 | 1.69* | 0.48 |
| 1976-78 | | | | -0.85 | -0.28 | | | | 0.43 | -0.76 |
| 1979-81 | | | | | 0.57 | | | | | -1.16 |
| **Low SES** | | | | | | | | | | |
| 1967-69 | -2.04 ** | -4.29 *** | -5.33 *** | -6.12 *** | -7.84 *** | -2.40 ** | -2.21 ** | -3.75 *** | -5.46 *** | -5.97 *** |
| 1970-72 | | -2.28 *** | -3.39 *** | -4.19 *** | -5.90 *** | | 0.13 | -1.48 | -3.19 *** | -3.71 *** |
| 1973-75 | | | -1.18 | -1.98 ** | -3.63 *** | | | -1.57 | -3.25 *** | -3.76 *** |
| 1976-78 | | | | -0.78 | -2.41 ** | | | | -1.65 * | -2.15 ** |
| 1979-81 | | | | | -1.62 | | | | | -0.51 |

*Note; *p<0.10, **p<0.05, ***p<0.01*

Moving on to the effects on employment propensity, Table 5 shows that the probability of being employed increases for both genders across the 4 lowest birth weight categories, while the effect of gaining weight at birth is less clear when birth weight exceeds 4,500 gr. This is consistent with what I find with earnings. Among men it seems like being born into the lowest birth weight category influences employment to a greater extent over time. However, the influence fluctuates a bit and across the first three groups of cohorts. The effect of low birth weight actually declines somewhat from the first to the second group of cohorts; men born in the period 1967-69 have 5.2 percentage points lower propensity to be employed compared to the reference group while the effect is 4.3 percentage points for men born 1970-72 and 1973-75. Among male groups of cohorts born later on, the estimated impact of low birth weight has increased to 6.1-6.9 percentage points and most of the differences between the first and last three groups of cohorts are significant (table 6).

For economic classes above the mean, there are no or only minor differences in male employment, while there is a social gradient in female employment for cohorts born 1967-69 (see appendix figure 4). For the last cohorts of women, the gradient is more like that of men with small

differences in employment at the middle and top of the distribution. For economic classes below the mean, there is a strong association between parental earnings and employment. Additionally, the effect of coming from the very poorest becomes significantly more important over time. Compared to having parents with median earnings, being born into the 5 percent poorest families implies a 13 (18) percentage point reduction in the propensity to be employed for men (women) in the earliest cohorts (1967-69). For men (women) born in the latest group of cohorts (1982-84) the effect has increased to 20 (25) percentage point. The findings of social background are in line with the study of Markussen and Røed (2017).

***Table 7. Estimation results. Effect of health and ability on log earnings at age 27-31. Males only.***

| Year of birth | 1967-69 | 1970-72 | 1973-75 | 1976-78 | 1979-81 | 1982-84 |
|---|---|---|---|---|---|---|
| *N* | *91,656* | *88,533* | *80,516* | *70,576* | *68,664* | *67,918* |
| **Height (cm)** | | | | | | |
| <170 | -0.128*** | -0.107*** | -0.101*** | -0.100*** | -0.160*** | -0.122*** |
| | (0.010) | (0.011) | (0.013) | (0.014) | (0.015) | (0.016) |
| 170-180 | -0.029*** | -0.022*** | -0.029*** | -0.029*** | -0.033*** | -0.037*** |
| | (0.005) | (0.005) | (0.006) | (0.006) | (0.007) | (0.007) |
| 180-190 | Ref. | Ref. | Ref. | Ref. | Ref. | Ref. |
| >190 | -0.003 | 0.010 | -0.027*** | 0.014 | -0.005 | 0.001 |
| | (0.009) | (0.009) | (0.011) | (0.012) | (0.012) | (0.013) |
| **Body Mass Index(BMI)** | | | | | | |
| Underweight (BMI <18.5) | -0.118*** | -0.136*** | -0.158*** | -0.164*** | -0.179*** | -0.205** |
| | (0.009) | (0.010) | (0.012) | (0.014) | (0.013) | (0.014) |
| Normal weight | Ref. | Ref. | Ref. | Ref. | Ref. | Ref. |
| Overweight (BMI25-30) | 0.009 | 0.004 | 0.006 | 0.011 | 0.052*** | 0.036*** |
| | (0.007) | (0.007) | (0.008) | (0.009) | (0.009) | (0.009) |
| Obesity (BMI >30) | -0.114*** | -0.104*** | -0.103*** | -0.070*** | -0.065*** | -0.066*** |
| | (0.016) | (0.015) | (0.016) | (0.016) | (0.016) | (0.014) |
| **Ability** | | | | | | |
| Low ability (stanine score 1-3) | -0.168*** | -0.145*** | -0.184*** | -0.178*** | -0.183*** | -0.168*** |
| | (0.006) | (0.006) | (0.008) | (0.009) | (0.009) | (0.010) |
| Medium ability | Ref. | Ref. | Ref. | Ref. | Ref. | Ref. |
| High ability (stanine score 4-6) | 0.065*** | 0.075*** | 0.040*** | 0.026*** | 0.015* | 0.017* |
| | (0.006) | (0.006) | (0.007) | (0.007) | (0.008) | (0.009) |
| **Birth weight** | | | | | | |
| <2,500 gr | -0.088*** | -0.075*** | -0.092*** | -0.119*** | -0.093*** | -0.115*** |
| | (0.013) | (0.013) | (0.015) | (0.017) | (0.018) | (0.019) |
| 2,500-3,000 gr | -0.036** | -0.041*** | -0.050*** | -0.057*** | -0.020* | -0.065*** |
| | (0.008) | (0.008) | (0.009) | (0.011) | (0.011) | (0.012) |
| 3,000-3,500 gr | -0.021*** | -0.022*** | -0.028*** | -0.028*** | -0.028*** | -0.022*** |
| | (0.005) | (0.005) | (0.006) | (0.007) | (0.007) | (0.008) |
| 3,500-4,500 gr | Ref. | Ref. | Ref. | Ref. | Ref. | Ref. |
| >4,500 gr | 0.003 | -0.008 | 0.030** | 0.017 | -0.028 | 0.008 |
| | (0.011) | (0.012) | (0.015) | (0.015) | (0.016) | (0.016) |
| **Parental earnings rank** | | | | | | |
| 1 | -0.313*** | -0.360*** | -0.423*** | -0.424*** | -0.469*** | -0.495*** |
| | (0.012) | (0.013) | (0.014) | (0.016) | (0.017) | (0.018) |
| 2 | -0.198** | -0.187** | -0.224** | -0.208** | -0.240** | -0.278** |
| | (0.012) | (0.013) | (0.014) | (0.016) | (0.017) | (0.017) |
| 3 | -0.128*** | -0.118*** | -0.194*** | -0.170*** | -0.203*** | -0.212*** |
| | (0.012) | (0.012) | (0.014) | (0.016) | (0.017) | (0.017) |
| 4 | -0.113*** | -0.117*** | -0.139*** | -0.142*** | -0.204*** | -0.159*** |
| | (0.012) | (0.012) | (0.012) | (0.016) | (0.017) | (0.017) |

*Note; *p<0.10, **p<0.05, ***p<0.01. Standard error in parenthesis. Parental earnings are divided in 20 bins (vigintiles). Vigintile 10 and 11 is the reference group. The control variables are the same as those reported for table 3.*

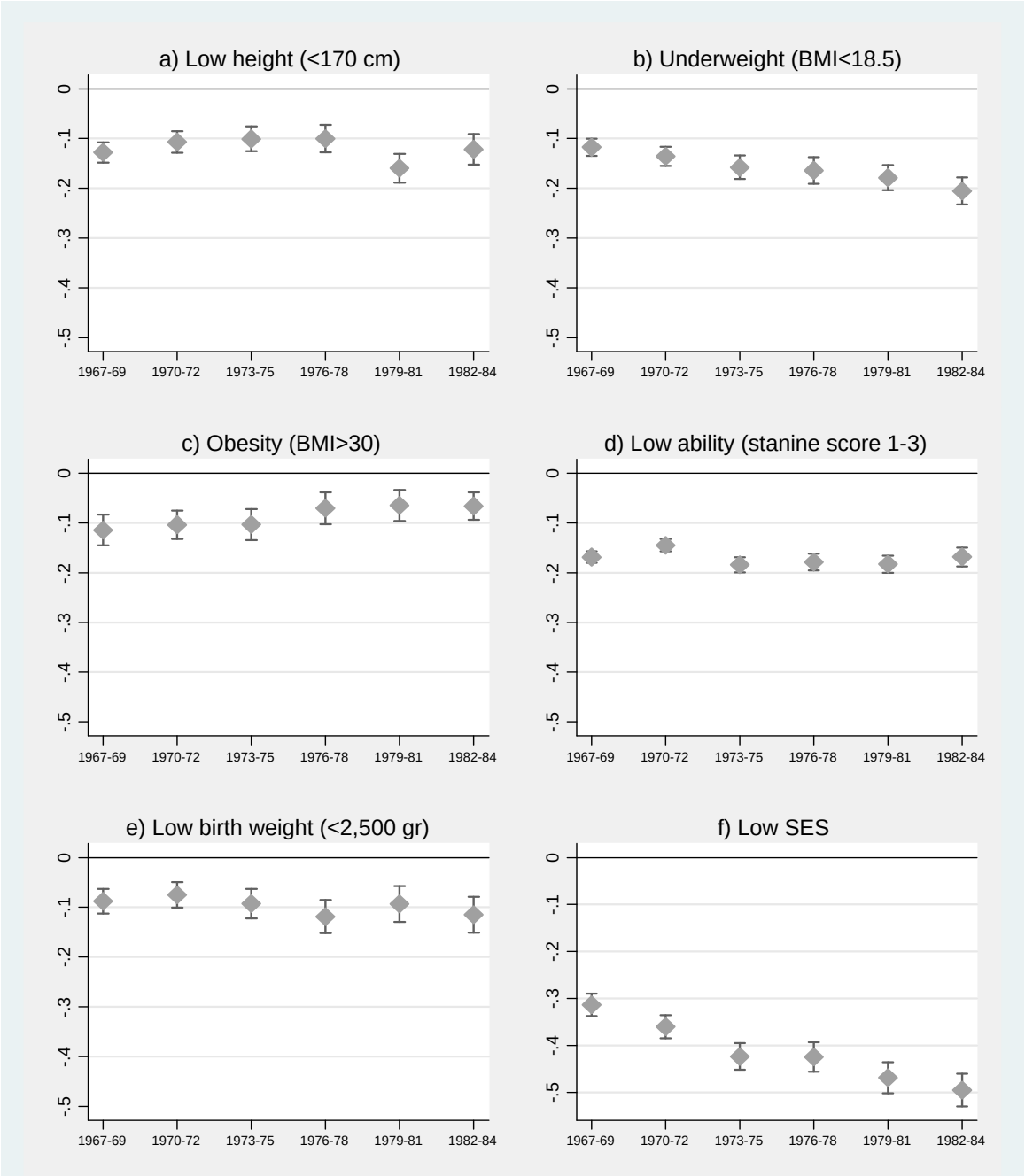



***Figure 7. Estimated effects of various health-indicators on log earnings at age 27-31. Males only.***
*Note; Reference groups are the following; height:180-190 cm, BMI-group: normal weight (BMI 18.5-25), ability: medium ability (stanine score 4-6), birth weight: 3,500-4,500 gr. Socioeconomic status (SES) is divived in 20 vigintiles based on parental earnings' rank. Vigintile 10 and 11 is the reference group. Low SES is vigintile 1, which is the five percent with lowest parental earnings. Point estimates with confidence interval are shown in the figure. The control variables are the same as those reported for table 3.*

***Table 8. t-test of significance of various health indicators between each group of cohorts. Dependent variable is log earnings, as shown in figure 7.***

| Year of birth | 1970-72 | 1973-75 | 1976-78 | 1979-81 | 1982-84 |
|---|---|---|---|---|---|
| **Low height** | | | | | |
| 1967-69 | 1.23 | 1.56 | 1.55 | -1.77* | 0.33 |
| 1970-72 | | 0.35 | 0.37 | -2.89*** | -0.80 |
| 1973-75 | | | 0.04 | -3.17*** | -1.10 |
| 1976-78 | | | | -3.10*** | -1.11 |
| 1979-81 | | | | | 1.94* |
| **Underweight** | | | | | |
| 1967-69 | -1.24 | -2.60*** | -2.86*** | -3.93*** | -5.53*** |
| 1970-72 | | -1.38 | -1.68* | -2.66*** | -4.22*** |
| 1973-75 | | | -0.35 | -1.21 | -2.73*** |
| 1976-78 | | | | -0.82 | -2.28** |
| 1979-81 | | | | | -1.54 |
| **Obesity** | | | | | |
| 1967-69 | 0.44 | 0.46 | 1.86* | 2.16** | 2.23** |
| 1970-72 | | 0.02 | 1.49 | 1.80* | 1.86* |
| 1973-75 | | | 1.47 | 1.78* | 1.83* |
| 1976-78 | | | | 0.27 | 0.21 |
| 1979-81 | | | | | -0.08 |
| **Low ability** | | | | | |
| 1967-69 | 2.44** | -1.49* | -0.91 | -1.34 | 0.02 |
| 1970-72 | | -3.73*** | -3.08*** | -3.51*** | -2.13** |
| 1973-75 | | | 0.50 | 0.10 | 1.35 |
| 1976-78 | | | | -0.39 | 0.83 |
| 1979-81 | | | | | 1.22 |
| **Low birth weight** | | | | | |
| 1967-69 | 0.63 | -0.22 | -1.43 | -0.25 | -1.25 |
| 1970-72 | | -0.84 | -2.01** | -0.83 | -1.83* |
| 1973-75 | | | -1.19 | -0.04 | -1.01 |
| 1976-78 | | | | 1.08 | 0.16 |
| 1979-81 | | | | | -0.92 |
| **Low SES** | | | | | |
| 1967-69 | -2.41** | -5.53*** | -5.37*** | -7.45*** | -8.69*** |
| 1970-72 | | -3.15*** | -3.08*** | -5.17*** | -6.41*** |
| 1973-75 | | | -0.05 | -2.12** | -3.34*** |
| 1976-78 | | | | -2.00** | -3.18*** |
| 1979-81 | | | | | -1.18 |

*Note; *p<0.10, **p<0.05, ***p<0.01.*

I now turn to regression models where I also include indicators for cognitive ability and additional health variables measured at a time closer to potential labor market entry. More specifically, I include information about height, weight and cognitive ability from the Norwegian military records. I will use four indicators for height; <170 cm, 170-180 cm, 180-190 cm and >190 cm. By height and weight I construct Body Mass Index (BMI), which is weight in kilograms divided by height in meters squared. I utilize WHO's cut- off points for underweight (BMI <18.5), normal weight (BMI 18.5-25), overweight (BMI 25-30) and obesity (BMI >30). For cognitive ability, I use three categories (low, medium and high) depending on IQ measured in stanine score. Since these data are based on military records, they are available only for men.

Controlling for the other variables, the relationship between height and earnings seems to be nonlinear (table 7). Being shorter than the reference group (180-190 cm) implies lower earnings, while being taller does not affect earnings significantly. The negative effect is most pronounced among men being shorter than 170 cm, which is consistent with findings in Lundborg *et al* (2014). The effect of low height (<170 cm) fluctuates somewhat, especially among the last three groups of cohorts, but there is no sign of gradual change in one direction or another.

Relative to normal weight, both underweight and obesity are a statistically significant determinants of male earnings. Overweight actually implies somewhat higher earnings, which is surprising given that overweight is supposed to be a risk factor for morbidity and mortality. One other striking feature is that underweight affects employment more negatively than obesity. However, the results of overweight can be supported by several studies (see f.ex. Flegal *et al*; 2013) where overweight implies lower mortality risk than normal weight. Aune *et al*. (2016) argue that studies finding this relationship may be driven by lack of important control variables, like smoking habits and prevalent disease.[21] In their meta-analysis, Aune *et al*. (2016) find that when smokers and individuals with prevalent disease are excluded both overweight, obesity and underweight result in higher mortality. In my study, I do not have access to neither smoking habits nor diseases, which probably means that I underestimate the adverse influence of higher BMI. One other reason why individuals defined as overweight have no higher risk of non-employment may be that the BMI-measure has an ambiguous interpretation, since the measure is not able to distinguish between body

[21] Both smoking and many chronic diseases may indicate lower weight and increased risk of mortality

fat and fat-free mass. Johansson *et al*. (2009) use measures of both BMI, weight, fat mass and waist circumference to investigate how obesity influence labor market participation. They find that obesity, defined by BMI, does not influence employment significantly, while obesity defined by fat mass, implies significant lower propensity to be employed.

Underweight is found to have a more negative influence on earnings over time. Table 7 shows that the point estimates of underweight on log earnings change significantly, from -0.12 for the first cohorts to -0.21 for the last. When it comes to the effect of obesity on earnings the effect has become significantly less important over time.

Low ability compared to medium ability has negative effect on earnings, but there is no particular trend in the estimated effects. There are, however, declining differences between men with high and normal ability. Men born in the period 1967-69 have around 7 percent higher earnings compared to men with medium ability, while men with high ability born in the latest cohorts have only 1.7 percent higher earnings. When running a model with IQ as a continuous variable I find that IQ affect employment to a lesser extent over time. This is in line with findings in Deming (2017) and Edin *et al*. (2017) which indicate stable or decreasing returns to cognitive skills during the last decades.

The effect of birth weight is lower compared to what I find in table 3, meaning that the effect of lower birth weight is partly explained by increased risk of being underweight at age 18-19, being shorter and being registered with low ability. But also with control for these indicators there is an effect of birth weight, and the tendency that lower birth weight affects employment to somewhat greater extent among the last groups of cohorts is still present.

Controlling for the variables from the military records reduces the estimated effects of social background somewhat. In particular, the coefficient for the lowest economic class is cut by about one fourth, meaning that part of the social gradient is explained by lower cognitive skills, lower height, underweight and obesity measured at age 18-19.

Appendix figure 5 illustrates results for log earnings condition on employment. The magnitude of the coefficients for all vulnerable groups in the figure is lower compared to what I find when utilizing the whole sample.

***Table 9. Estimation results. Effect of health and ability on employment. Males.***

| Year of birth | 1967-69 | 1970-72 | 1973-75 | 1976-78 | 1979-81 | 1982-84 |
|---|---|---|---|---|---|---|
| *N* | *91,656* | *88,533* | *80,516* | *70,576* | *68,664* | *67,918* |
| **Height (cm)** | | | | | | |
| <170 | -0.042*** | -0.037*** | -0.034*** | -0.031*** | -0.052*** | -0.039*** |
| | (0.004) | (0.005) | (0.005) | (0.006) | (0.006) | (0.006) |
| 170-180 | -0.006*** | -0.004*** | -0.006*** | -0.007*** | -0.010*** | -0.009*** |
| | (0.002) | (0.002) | (0.002) | (0.003) | (0.003) | (0.003) |
| 180-190 | Ref. | Ref. | Ref. | Ref. | Ref. | Ref. |
| >190 | -0.003 | 0.003 | -0.007 | 0.003 | -0.006 | -0.004 |
| | (0.004) | (0.004) | (0.005) | (0.005) | (0.005) | (0.005) |
| **Body Mass Index (BMI)** | | | | | | |
| Underweight (BMI <18.5) | -0.034*** | -0.036*** | -0.054*** | -0.049*** | -0.058*** | -0.060*** |
| | (0.004) | (0.004) | (0.005) | (0.006) | (0.005) | (0.006) |
| Normal weight | Ref. | Ref. | Ref. | Ref. | Ref. | Ref. |
| Overweight (BMI 25-30) | 0.008** | 0.008** | 0.002 | 0.009** | 0.017*** | 0.011*** |
| | (0.003) | (0.003) | (0.003) | (0.004) | (0.004) | (0.004) |
| Obesity (BMI >30) | -0.022*** | -0.020*** | -0.031*** | -0.013* | -0.021*** | -0.017*** |
| | (0.007) | (0.006) | (0.007) | (0.007) | (0.006) | (0.006) |
| **Ability** | | | | | | |
| Low ability (stanine score 1-3) | -0.049*** | -0.042*** | -0.057*** | -0.052*** | -0.053*** | -0.047*** |
| | (0.003) | (0.003) | (0.003) | (0.004) | (0.004) | (0.004) |
| Medium ability | Ref. | Ref. | Ref. | Ref. | Ref. | Ref. |
| High ability (stanine score 7-9) | 0.006** | 0.006*** | 0.002 | -0.003 | -0.010*** | 0.004 |
| | (0.002) | (0.002) | (0.003) | (0.003) | (0.003) | (0.004) |
| **Birth weight** | | | | | | |
| <2,500 gr | -0.028*** | -0.021*** | -0.026*** | -0.039*** | -0.042*** | -0.036*** |
| | (0.005) | (0.005) | (0.006) | (0.007) | (0.007) | (0.007) |
| 2,500-3,000 gr | -0.008** | -0.012*** | -0.015*** | -0.018*** | -0.010** | -0.019*** |
| | (0.003) | (0.003) | (0.004) | (0.004) | (0.005) | (0.005) |
| 3,000-3,500 gr | -0.008*** | -0.006*** | -0.010*** | -0.009*** | -0.009*** | -0.007*** |
| | (0.002) | (0.002) | (0.003) | (0.003) | (0.003) | (0.003) |
| 3,500-4,500 gr | Ref. | Ref. | Ref. | Ref. | Ref. | Ref. |
| >4,500 gr | 0.001 | -0.006 | 0.009 | 0.005 | 0.008 | (0.009) |
| | (0.005) | (0.005) | (0.006) | (0.006) | (0.006) | (0.006) |
| **Parental earnings' rank** | | | | | | |
| 1 | -0.107*** | -0.125*** | -0.146*** | -0.151*** | -0.155*** | -0.165*** |
| | (0.005) | (0.005) | (0.006) | (0.007) | (0.007) | (0.007) |
| 2 | -0.067*** | -0.066*** | -0.075*** | -0.068*** | -0.077*** | -0.090*** |
| | (0.005) | (0.005) | (0.006) | (0.007) | (0.007) | (0.007) |
| 3 | -0.037*** | -0.035*** | -0.069*** | -0.055*** | -0.062*** | -0.070*** |
| | (0.005) | (0.005) | (0.006) | (0.007) | (0.007) | (0.007) |
| 4 | -0.036*** | -0.035*** | -0.054*** | -0.041*** | -0.057*** | -0.047*** |
| | (0.005) | (0.005) | (0.006) | (0.007) | (0.007) | (0.007) |

*Note; *p<0.10, **p<0.05, ***p<0.01. Standard error in parenthesis. Parental earnings are divided in 20 bins (vigintiles). Vigintile 10 and 11 is the reference group. The control variables are the same as those reported for table 3.*

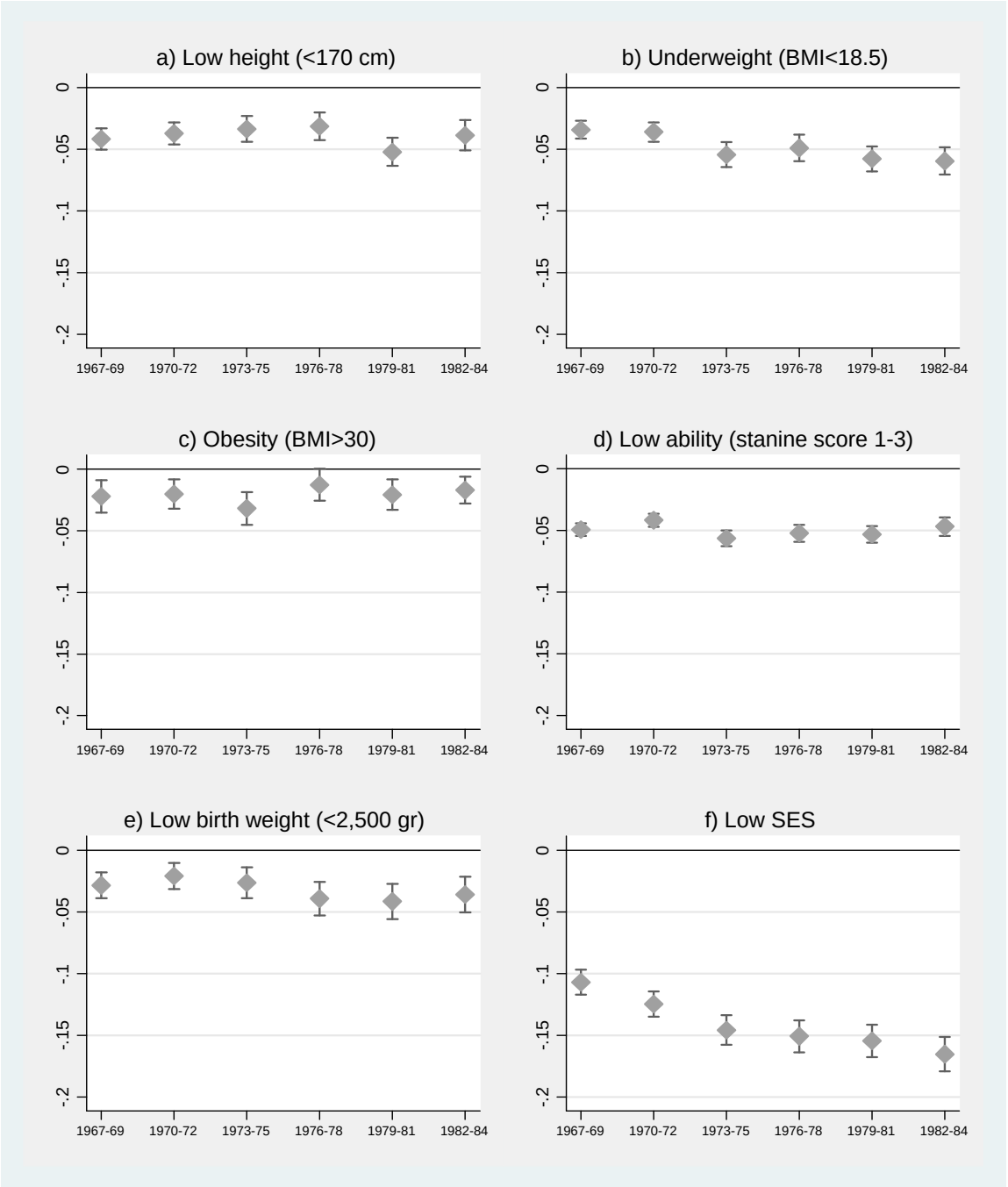



***Figure 8. Estimated effects of various health-indicators on employment. Males only.***
*Reference groups are the following; height:180-190 cm, BMI-group: normal weight (BMI 18.5-25), ability: medium ability (stanine score 4-6), birth weight: 3,500-4,500 gr. Socioeconomic status (SES) is divived in 20 vigintiles based on parental earnings' rank. Vigintile 10 and 11 is the reference group. Low SES is vigintile 1, which is the five percent with lowest parental earnings. Point estimates with confidence interval are shown in the figure. The control variables are the same as those reported for table 3.*

***Table 10. t-test of significance of various health indicators between each group of cohorts..***
***Dependent variable is employment, as shown in figure 8.***

| | 1970-72 | 1973-75 | 1976-78 | 1979-81 | 1982-84 |
|---|---|---|---|---|---|
| **Low height** | | | | | |
| 1967-69 | 0.63 | 1.15 | 1.41 | -1.43 | 0.41 |
| 1970-72 | | 0.52 | 0.80 | -1.99** | -0.18 |
| 1973-75 | | | 0.29 | -2.44** | -0.66 |
| 1976-78 | | | | -2.65*** | -0.92 |
| 1979-81 | | | | | 1.69* |
| **Underweight** | | | | | |
| 1967-69 | -0.31 | -3.16*** | -2.20** | -3.69*** | -3.88*** |
| 1970-72 | | -2.76*** | -1.85* | -3.26*** | -3.46*** |
| 1973-75 | | | 0.75 | -0.47 | -0.71 |
| 1976-78 | | | | -1.21 | -1.42 |
| 1979-81 | | | | | -0.25 |
| **Obesity** | | | | | |
| 1967-69 | 0.20 | -1.00 | 0.96 | 0.14 | 0.56 |
| 1970-72 | | -1.26 | 0.81 | -0.07 | 0.37 |
| 1973-75 | | | 2.08** | 1.24 | 1.77* |
| 1976-78 | | | | -0.91 | -0.53 |
| 1979-81 | | | | | 0.46 |
| **Low ability** | | | | | |
| 1967-69 | 1.90* | -1.65* | -0.66 | -0.87 | 0.55 |
| 1970-72 | | -3.38*** | -2.36** | -2.57*** | -1.13 |
| 1973-75 | | | 0.87 | 0.69 | 1.99** |
| 1976-78 | | | | -0.19 | 1.10 |
| 1979-81 | | | | | 1.29 |
| **Low birth weight** | | | | | |
| 1967-69 | 0.92 | 0.23 | -1.22 | -1.44 | -0.84 |
| 1970-72 | | -0.66 | -2.07** | -2.28** | -1.68* |
| 1973-75 | | | -1.41 | -1.62 | -1.03 |
| 1976-78 | | | | -0.23 | 0.36 |
| 1979-81 | | | | | 0.58 |
| **Low SES** | | | | | |
| 1967-69 | -2.22** | -4.74*** | -5.15*** | -5.55*** | -6.79*** |
| 1970-72 | | -2.56** | -3.05*** | -3.46*** | -4.70*** |
| 1973-75 | | | -0.57 | -0.99 | -2.21** |
| 1976-78 | | | | -0.41 | -1.59 |
| 1979-81 | | | | | -1.18 |

*Note; *p<0.10, **p<0.05, ***p<0.01*

Investigating the effect of the various health- and ability measures on employment (table 9, figure 8 and table 10) gives much of the same pattern as that of log earnings.

Low height (height <170 cm) is associated with 3-5 percentage point lower propensity to be employed compared to the reference group with height 180-190 cm. For the first groups of cohorts, being underweight (compared to having normal weight) implies approximately 3.5 percentage point lower probability of being employed. For the last group of cohorts the negative effect has increased significantly; to around 6 percentage point. Obesity, compared to normal weight, is associated with around 2 percentage point lower employment propensity.

Low ability compared to medium ability has negative effect on employment, but there is no particular change over time. There are only minor differences in the employment propensity between men with high and medium ability.

As with earnings, the magnitude of the estimate of low birth weight and low socioeconomic class becomes lower when controlling for the variables measured by the military service. According to the point-estimates low birth weight (compared with birth weight 3,500-4,500 gr) decreases the probability to be employed by 2.9 percentage points for the first group of cohorts while the effect has increased somewhat (to 3.6 percentage points) for the last group of cohorts. Controlling for the variables from the military records reduces the magnitude of the coefficient of lowest economic class by about forty percent. But still, belonging to the poorest economic class decreases the employment probability considerably, and the effect becomes significantly stronger over time. For cohorts born in the period 1967-69 men born into the 5 percent poorest families has 10.7 percentage point lower employment probability (compared to median earnings families), while the effect has increased to 15.5 percentage point for cohorts born in the period 1982-84.

Taken together, I find strong evidence of steadily increasing negative impact on labor market performance for young adults of being born into the very poorest families. There are also some indications of stronger negative impact of poor health among males, but the results are sensitive to what kind of health indicator I use. Low birth weight and underweight at age 18-19 have become a stronger predictor of low earnings and non-employment, while there is quite constant effect of low height and obesity. One possible explanation is that underweight individuals to a greater degree face common mental health problems, such as anxiety and depression. With more emphasize on reorganizations and efficiency it is reasonable to assume that good mental health is more strongly

rewarded in the labor market. When it comes to cognitive ability, the influence on employment and earnings has decreased over time, which is driven by decreasing return to high ability.

# 7 Reliability and robustness

I start this section by comparing the main results based on the control function approach with models with family – or twin fixed effects. Furthermore, I check whether low birth weight seems to be a stable predictor of health. Finally, I investigate whether the estimates are robust with respect to some sample selection issues and some choices of variable definitions.

## *7.1 Comparison with family fixed effects*

Several studies, particularly in economics, have used twin or sibling fixed effects in order to explore the effect of health on different outcomes.

In the main estimations above, I have divided the sample into six groups of cohorts. With both sibling and twin-fixed effects, the sample size is strongly restricted and estimating models for each group yields large standard errors. In order to receive more precise estimates with family fixed effects, I will instead divide the sample into three groups of cohorts corresponding to birth year 1967-72, 1973-78 and 1979-84. By estimating models for men and women separately, the sample with family fixed effects are restricted to same-sex siblings being born within the same group of cohort. Appendix table 3a) - 3d) give descriptive statistics of this sibling (being born as singletons) and twin sample.[22] The most striking observation in these tables is the tremendous difference in birth weight between singletons and twins, where the mean birth weight among twins is approximately 900 gr less than singletons. Approximately 40 percent of female twins and 32-36 percent of male twins are born with low birth weight (<2,500 gr), while it accounts for less than 3 percent of singletons. A small fraction of twins are represented in the reference category 3,500-4,500 gr; 6-7 percent of male twins and 4 percent of female twins.

[22] In this context siblings refer to same-sex siblings being born within the same cohort-group.

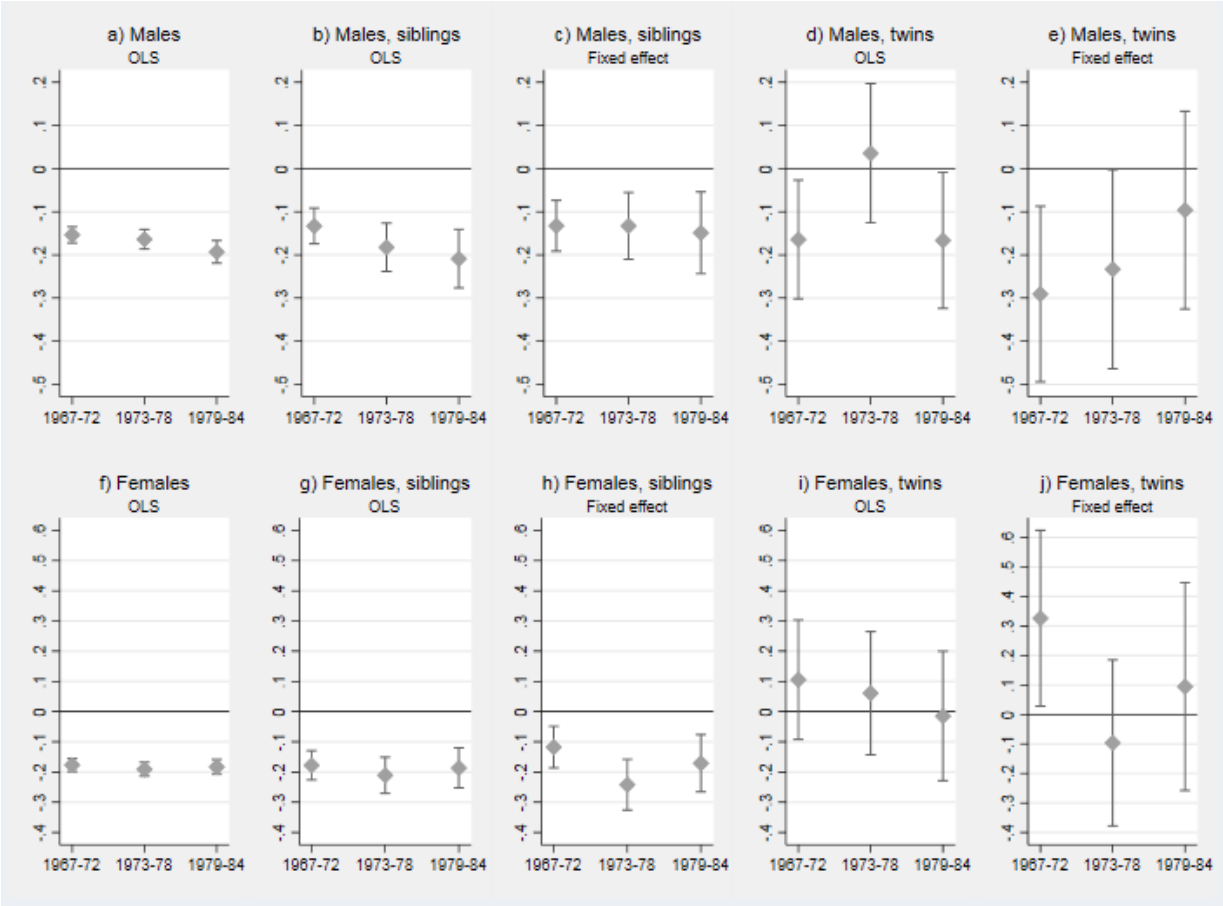



***Figure 11. Estimated effects of low birth weight (<2,500 gr) on log earnings at age 27-31. Comparison with family fixed effects.***

*Note; The reference group for birth weight is 3,500-4,500 gr. Point estimates with confidence interval are shown in the figure. Control variables for OLS; see note to table 3. For sibling fixed effects; birth order (4 dummy variables), month of birth (12 dummy variables). For twin fixed effects; none.*

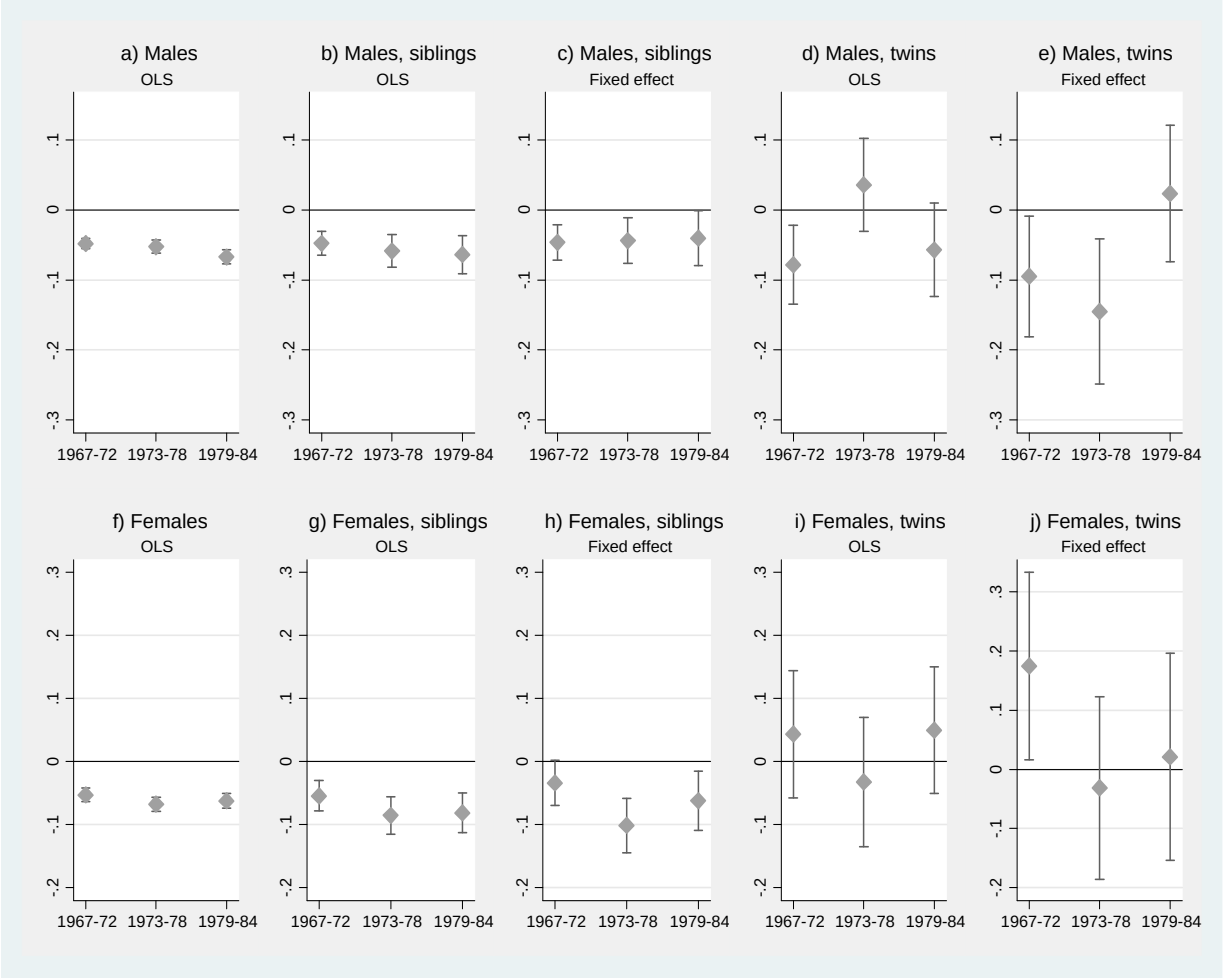



***Figure 12. Estimated effects of low birth weight (<2,500 gr) on employment. Comparison with family fixed effects.***

*Note; The reference group for birth weight is 3,500-4,500 gr. Point estimates with confidence interval are shown in the figure. Control variables for OLS; see note to table 3. For sibling fixed effects; birth order (4 dummy variables), month of birth (12 dummy variables). For twin fixed effects; none.*

The data sources utilized in Black *et al*. (2007) are similar to what I use in my study, but they include all twins in their models (not only same-sex twins) and estimate one regression for both genders. Furthermore, they use a loglinear specification of birth weight. When I run a regression on log earnings with the same specification as theirs, twin fixed effects give an estimate of 0.18 with standard error 0.06. This is somewhat higher compared to their estimate of 0.12.[23]

[23] The deviation may be due to their sample restrictions, which only contains persons working full-time. Other reasons for deviations are differences in how earnings are measured; they measure earnings in 2002, while I look at earnings age 27-31.

Figure 11 and 12 show estimation results with sibling and twin fixed effects, where the outcome measure is log earnings and employment, respectively. In order to facilitate comparisons of the different models I have also included estimates from models without family fixed effects based on the exact same samples. With twin fixed effects, the standard errors are large, and the point estimates fluctuates a lot. When log earnings is the outcome measure (see figure 11) the point estimates of being born with low birth weight (<2,500 gr) compared to the reference group of 3,500-4,500 gr ranges from minus 0.29 to plus 0.33. Using a loglinear specification of birth weight (not shown in figure or tables) also gives unstable results, f.ex., male twins born in the period 1967-73 experience that being 10 percent heavier than their twin raises the probability of being employed by 2 percentage points, while female twins born in the same period experience that a 10 percent increase in birth weight actually *decreases* the propensity to be employed by almost 3 percentage points. The estimates with sibling fixed effects are more similar to what I find by running OLS for the whole sample, but the standard errors are quite large and the estimates do not show any particular trend in the importance of birth weight.

Given the large difference in weight-distributions between twins and non-twins, there are good reasons to question the external validity of findings based on twin-data. As earlier mentioned there have also been major improvements in neonatal health care and since twins are overrepresented with lower birth weight and complications at birth, they may be more sensitive to given changes in technology, which in turn may affect outcomes later in life. The small sample sizes and the highly ambiguous results also suggest that twin fixed effects are not the appropriate strategy in order to identify trends in the effects of poor health.

Another problem with family fixed effects is that there may be spillover effects among siblings which affect the outcome variables (see f.ex. Fletcher *et al*., 2012; Breining, 2014; Black *et al*., 2017). This effect is probably strongest when one of the siblings has severe illnesses, which is more common the lower the birth weight is. If this is the case the fixed effects models give results that may not be generalizable to other groups. Furthermore, parents may invest in activities that reinforce or compensate initial health differences between siblings/twins (Almond and Mazumder, 2013; Yi *et al*., 2015), which in turn may affect the outcome variable.

Earlier research on the effect of height, BMI and ability has not focused on family-fixed effects, but there are (at least) one exception – namely Lundborg *et al*. (2014) who tried to separate

out the effect of height from social background. Figure 13 and 14 illustrate estimates on log earnings and employment from models with family fixed effects when incorporating health- and ability measures from the military records. Models based on twin fixed effects give very imprecise estimates, making it difficult to draw any conclusion about the magnitude of the changes in the effects of the different variables. Compared to the estimates presented in figure 7 and 8, sibling fixed effects give more imprecise estimates of the different health- and ability- measures. The results are, however, more precise than the twin-fixed estimates. The size of the point-estimates with sibling fixed effects also differ somewhat from the OLS-estimates. When it comes to development over time, it seems like being underweight influences earnings and employment steadily more, while it is no particular change in any of the other variables.

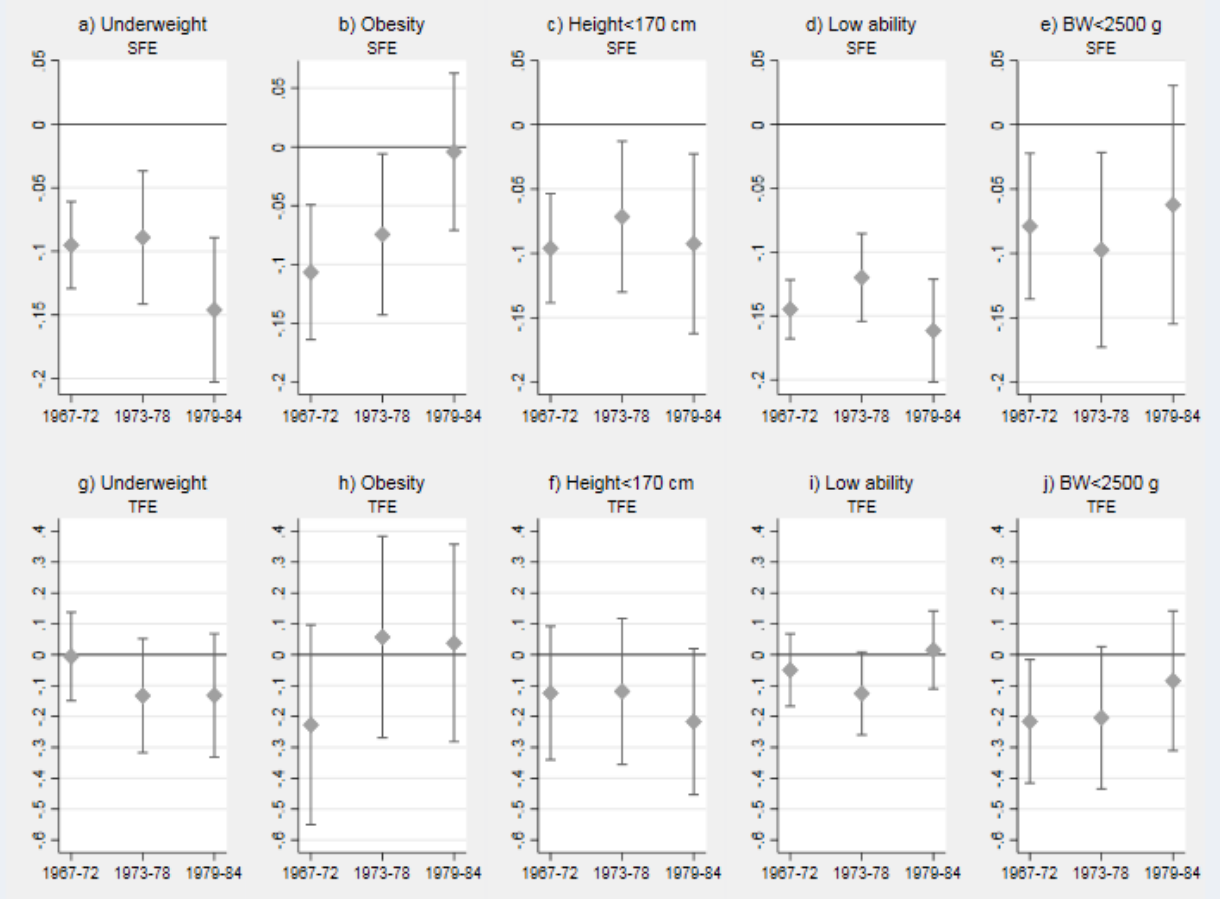



***Figure 13. Estimated effects of health- and ability indicators on log earnings at age 27-31. Family fixed effects. Males only.***

*Note; SFE=sibling fixed effects, TFE=twin fixed effects. Reference groups are the following; height:180-190 cm, BMI-group: normal weight (BMI 18.5-25), ability: medium ability (stanine score 4-6), birth weight: 3,500-4,500 gr. Point estimates with confidence interval are shown in the figure. Control variables for sibling fixed effects; birth order (4 dummy variables), month of birth (12 dummy variables). For twin fixed effects; none.*

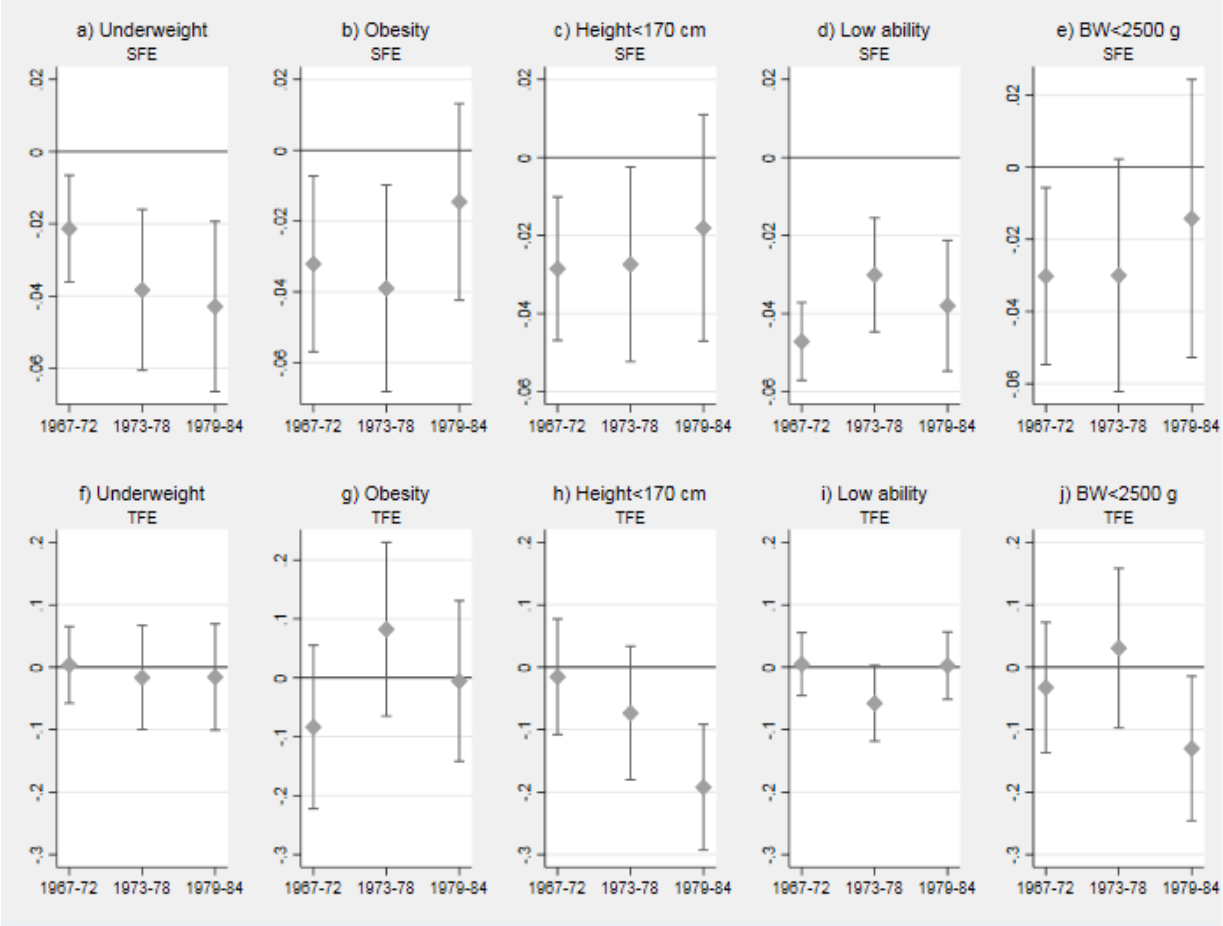



***Figure 14. Estimated effects of health- and ability indicators on employment. Family fixed effects. Males only.***
*Note; SFE=sibling fixed effects, TFE=twin fixed effects. Reference groups are the following; height:180-190 cm, BMI-group: normal weight (BMI 18.5-25), ability: medium ability (stanine score 4-6), birth weight: 3,500-4,500 gr. Point estimates with confidence interval are shown in the figure. Control variables for sibling fixed effects; birth order (4 dummy variables), month of birth (12 dummy variables). For twin fixed effects; none.*

### *7.2 The effect of birth weight on other outcomes*

The reliability of the results of the various indicators depends on whether the indicators capture the same degree of vulnerability over time. When it comes to birth weight, a stable relationship may be questionable and my primary concern is whether health of the individuals with birth weight less than 2,500 gr has changed. Many babies in this group are in need of extensive health care at birth and improvements in neonatal health care have resulted in substantially lower infant mortality rates. This may have affected health of the group with low birth weight, but it is not clear in which direction. Health may have deteriorated since the group will consist of a steadily increasing number

of individuals who were on the margin of survival as infants. However, Bharadwaj *et al*. (2013) have shown that additional treatment at birth also has positive effect on test scores and grades in school. This may imply that health of the whole group with low birth weight has improved over time.

In this section I test the assumption of stable relationship between low birth weight and adult health by exploring how low birth weight affects other potentially health-related outcomes and the degree to which this effect has changed over time. For both genders, I am able to explore the effect on mortality, high-school dropout and take-up of disability insurance. For men, I also investigate the impact on health- and ability variables from the military service.

Figure 15 shows how mortality (infant mortality and mortality at age 1-31), permanent disability insurance at age 19 and high school drop-out are influenced by low birth weight compared to the reference category with birth weight 3,500-4,500 gr.[24] When mortality is the outcome measure, every native being alive at birth or at age 1, respectively, are included, while the sample is identical to what is explained in section 2.3 when studying the effect on the other outcomes.

As expected, the figure shows that low birth weight affects infant mortality to a lesser extent over time. When surviving the first year of life, being born with low birth weight is also associated with higher risk of mortality afterwards for both genders, but there is no change in this effect over time. Furthermore, low birth weight is associated with higher risk of receiving disability benefits at age 19 and high-school dropout, which is consistent with findings in previous literature. For both genders, there has been an increase in the importance of low birth weight on take up of disability insurance, and the effect is significantly stronger among the last group of cohorts compared to the earliest group.[25] The association between low birth weight and high-school dropout has decreased over time for women, while there is no particular change for men.

Taken together, the results are somewhat ambiguous. An increasing effect on disability insurance may reflect that health of the group with low birth weight, at least for some individuals, has deteriorated, According to Brage and Thune (2015) the group of individuals being granted permanent disability insurance before age 20 is dominated by people with mental retardation and

[24] High-school dropout is defined as not completed high school by the age of 27.

[25] Since information about disability insurance is available from 1992, I am not able to explore this outcome for cohorts born before 1972. Being granted permanent disability insurance at age 19 is a rare outcome, accounting for 0.3-0.6 percent of the population (see table 1a and 1b)

congenital malformations, suggesting that disability at this early stage in life is not influenced by changes in the labor market. On the other hand, weaker importance of low birth weight on female education may be interpreted as a sign of better health among women with low birth weight. Another interpretation is that poor health has become less important for girls in order to complete high school.[26]

[26] This result is not consistent with results from Black et al. (2007). Using variation within twins, they find that cohorts born 1977-86 compared to cohorts born 1967-76 have stronger return of birth weight on education. Deviations in the results may be due to their restriction of the sample, which only contain twins.

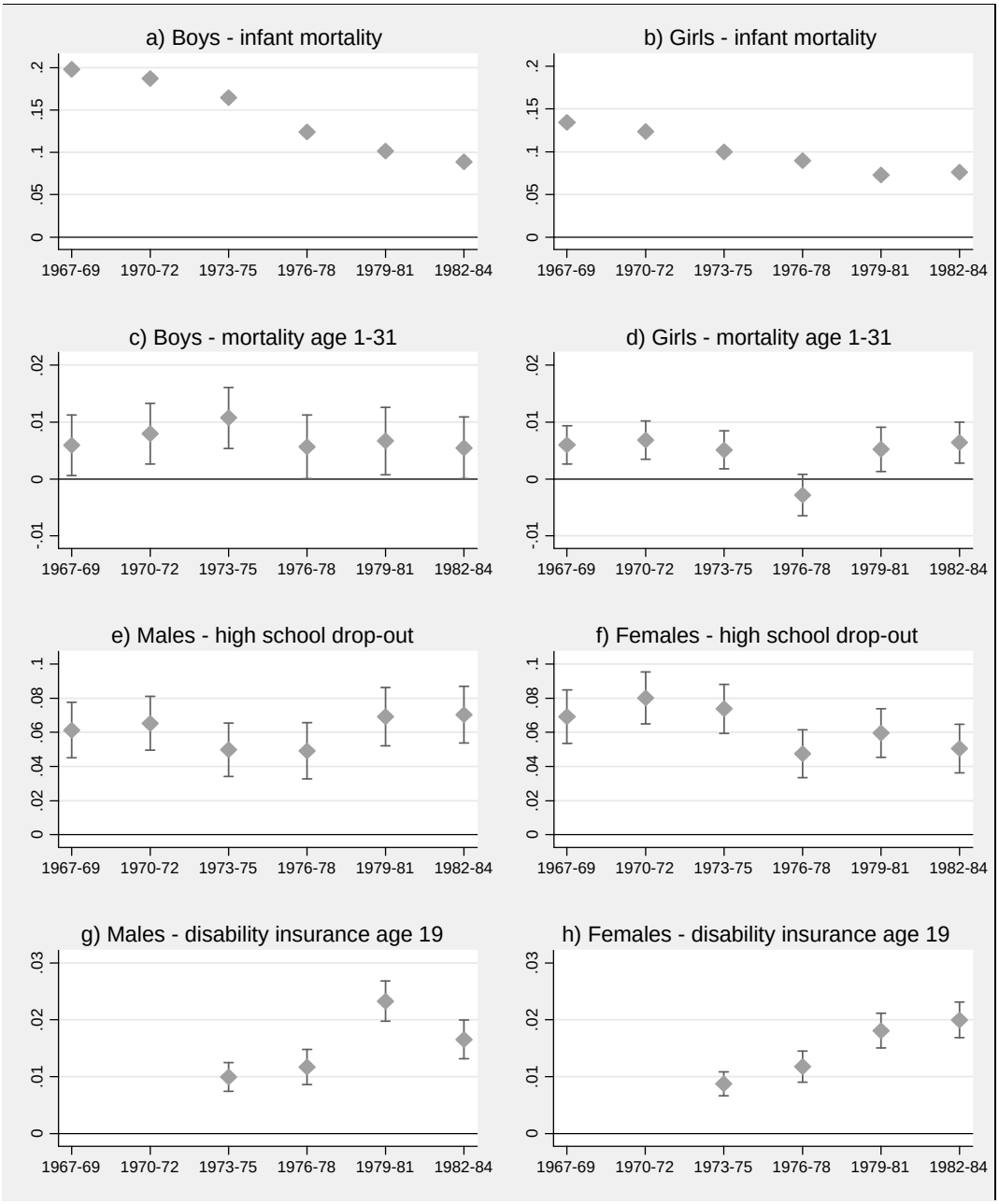



***Figure 15. Estimated effects of low birth weight (<2,500 gr) on various outcomes.***
*Note; The reference group is birth weight 3,500-4,500 gr. High-school dropout is defined as not completed high school at the age of 27. Point estimates with confidence interval are shown in the figure. The control variables are the same as those reported for table 3.*

Furthermore, figure 16 illustrates the impact of low birth weight on the health measures from the military service. Health and ability- measures are registered for the great majority of men and using these as outcome-variables, I am able to investigate whether the effect of birth weight on those measures has changed over time. Since entering the military service is compulsory for every

able men, men who are not registered with any information from the military service are probably negatively selected. Results from estimations (not shown in any tables or figures) show that men not registered with information from the military service have substantially lower earnings and are less likely to be employed. Hence, I will use missing information from the military service as an outcome measure indicating poor health.

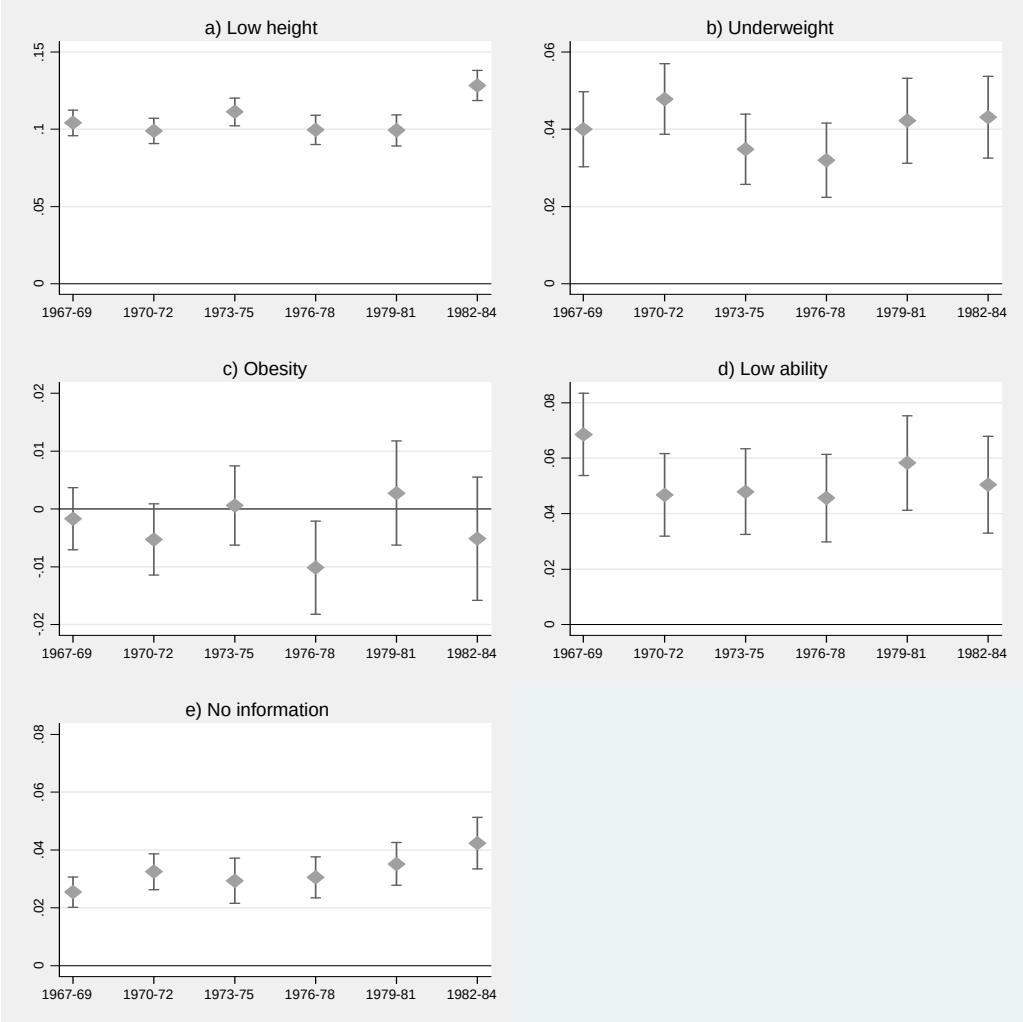



***Figure 16. Estimated effects of low birth weight (<2,500 gr) on health outcomes registered by the military service. Males only.***
*Note; The reference group is birth weight 3,500-4,500 gr. Low height=height<170 cm, underweight=BMI<18.5, obesity=BMI>30, low ability=stanine score 1-3, no information=not registered with any information from the military service. Point estimates with confidence interval are shown in the figure. The control variables are the same as those reported for table 3.*

In line with earlier studies, figure 16 shows that low birth weight increases the risk of low height and low ability. Low birth weight is also associated with higher risk of being underweight,

which is consistent with findings in Black *et al*. (2007).[27] Furthermore, I find that low birth weight does not imply higher risk of obesity. Given the numerous studies detecting that low birth weight increases the risk of obesity, it is a surprising result. However, a study from Switzerland (Jornayvaz *et al*., 2016) find a U-shaped relationship between birth weight and body mass index, and a study using Norwegian register data (Eide *et al*., 2005) finds that higher birth weight implies higher adult weight. My models show that men with birth weight 2,500-3,500 gr have the lowest risk of being obese while men with high birth weight (>4,500 gr) have the highest risk. The latter effect also becomes substantially stronger over time. This may reflect that an increasing proportion of men with high birth weight have mothers with health problems, like diabetes and obesity.[28]

Figure 16 does not show any clear evidence of changing impact of low birth weight on the various health measures, but there are some estimators that deviate from the others. Low birth weight has a significantly stronger impact on low height for men born in the last period (1982-84) compared to earlier periods, while low birth weight affect low ability most for men born in the earliest period (1967-69). Black *et al*. (2007) find that the impacts of birth weight on different outcomes (height, BMI, IQ and high school completion) increase over time when comparing estimates based on twin fixed effects from the cohorts 1977-1986 with cohorts 1967-1976. Their study is also based on Norwegian register data, but their estimates are not directly comparable to my estimates since they only use twins. Bharadwaj *et al*. (2018) also investigate changes in the impact of birth weight, but they compare cohorts further apart than I do. They compare Swedish cohorts born 1974 and onwards with cohorts born 1926-1958 and their twin fixed effects give quite constant impact of birth weight on outcomes like earnings and high school completion.

Using missing information from the military service as an indicator of poor health, I do find that the effect of low birth weight has become stronger over time. This may imply that the proportion of individuals with severe health problems in the group with low birth weight have increased, which is consistent with the findings on disability (see figure 17). The overall fraction with missing

[27] Apart from Black *et al*. (2007) I am not aware of any studies that focus on birth weight and underweight later in life.

[28] Estimating models where I switch the reference category to be 2,500-3,000 g, I find that men with birth weight <2,500 g use to be somewhat (0-1 percentage point) more likely to be obese. Men with high birth weight (>4,500 g) have the highest risk of being obese, which also become substantially stronger over time. Compared to the reference category men with high birth weight born in the period 1967-69 have 1.3 percentage point higher risk of being obese while the risk has increased to 3.2 percentage point for men born the period 1982-84.

information has also increased considerably, hence, relative to the size of this group, the impact of low birth weight has not become a more important determinant.

### *7.3 Robustness checks*

In this section I test whether my results of low birth weight are sensitive to how I define employment and social background, and whether the estimates are driven by changes in the composition of people with lower birth weight. Results from the robustness checks of low birth weight are reported in figure 17 (results on log earnings) and figure 18 (employment).

Panels a) and b) in figure 17 and 18 show results from models where individuals with perceived severe health problems are excluded, i.e. individuals born with very low birth weight (<1,500 gr) and individuals being disabled at age 19. As expected, excluding those individuals give lower point estimates for the group with birth weight below 2,500 gr. Furthermore, the increasing impact of low birth weight among men is no longer significant, which may be attributed to the fact that the proportion of people born with very low birth weight has increased (see figure 2). This also seem to be the case when excluding individuals receiving permanent disability insurance at age 19, meaning that the increasing impact of low birth weight on males earnings and employment is sensitive to whether groups with very poor health are excluded.

Social background could be measured in different ways and I have used two alternative definitions. Measuring social background by parents' education level only (alternative 1) gives stronger association between birth weight and earnings and employment which indicate that controlling for parents' earnings is essential to separate health from social background. Having children with poor health may affect parents' earnings, which implies that parents' earnings at age 50-54 could be an endogenous variable. To circumvent this problem I use parents' earnings at birth instead (alternative 2). This definition also shows stronger impact of birth weight on employment, which may imply that I control away part of the effect of birth weight on earnings and employment when I use parents' earnings at their early fifties. However, since earnings in early fifties are more representative for lifetime earnings than earnings earlier in life (see Markussen and Røed; 2017) I will argue that this is a better measure of social background than earnings at birth. The estimated time trends are in any case very similar.

For men, the results are to some degree sensitive to how I define employment (see figure 18, panel e and f). The stronger the criteria for being classified as employed the stronger is the

effect of low birth weight. When identifying the development over time there are also some differences and the weakest definition of employment gives the strongest changes in the importance of low birth weight. With the strictest definition of employment, the tendency of increasing impact of low birth weight is less clear.

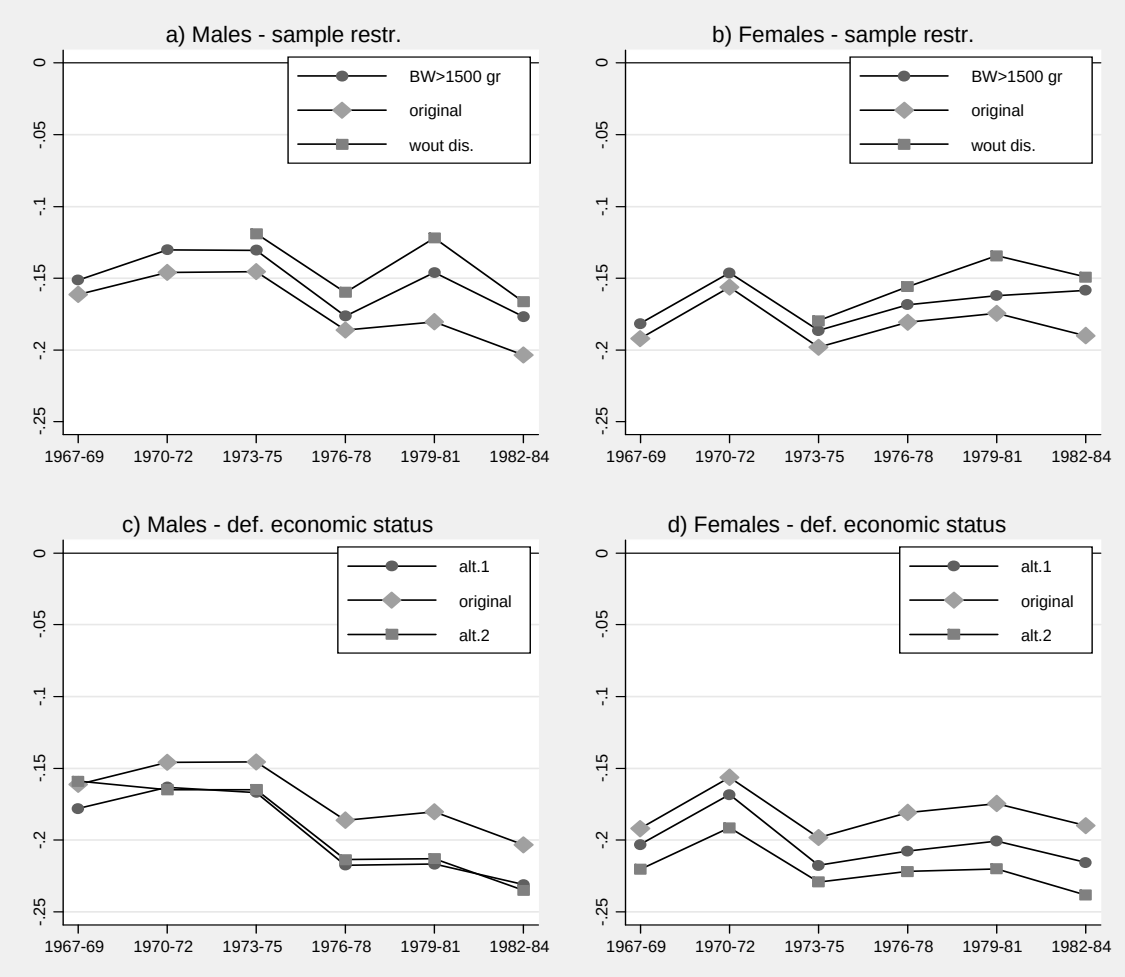



***Figure 17. Estimated effects of low birth weight (<2,500 gr) on log earnings at age 27-31. Robustness checks.***

*Note; The reference group is birth weight 3,500-4,500 gr. Point estimates are shown in the figure. The point estimates from figure 5, panel a) for males and panel b) for females are depticted in every panels and labeled by "original." Panel a) and b) compare the "original" estimate with estimates from models where indididuals with birth weight less than 1,500 gr and persons receiving disability insurance at age 19 are excluded. Panel c) and d) compare the "original" estimate with estimates from models where social background is defined by parents' education level only (alternativ 1) and parents' earnings rank at birth of the child (alternativ 2).*

*The control variables are the same as those reported for table 3.*

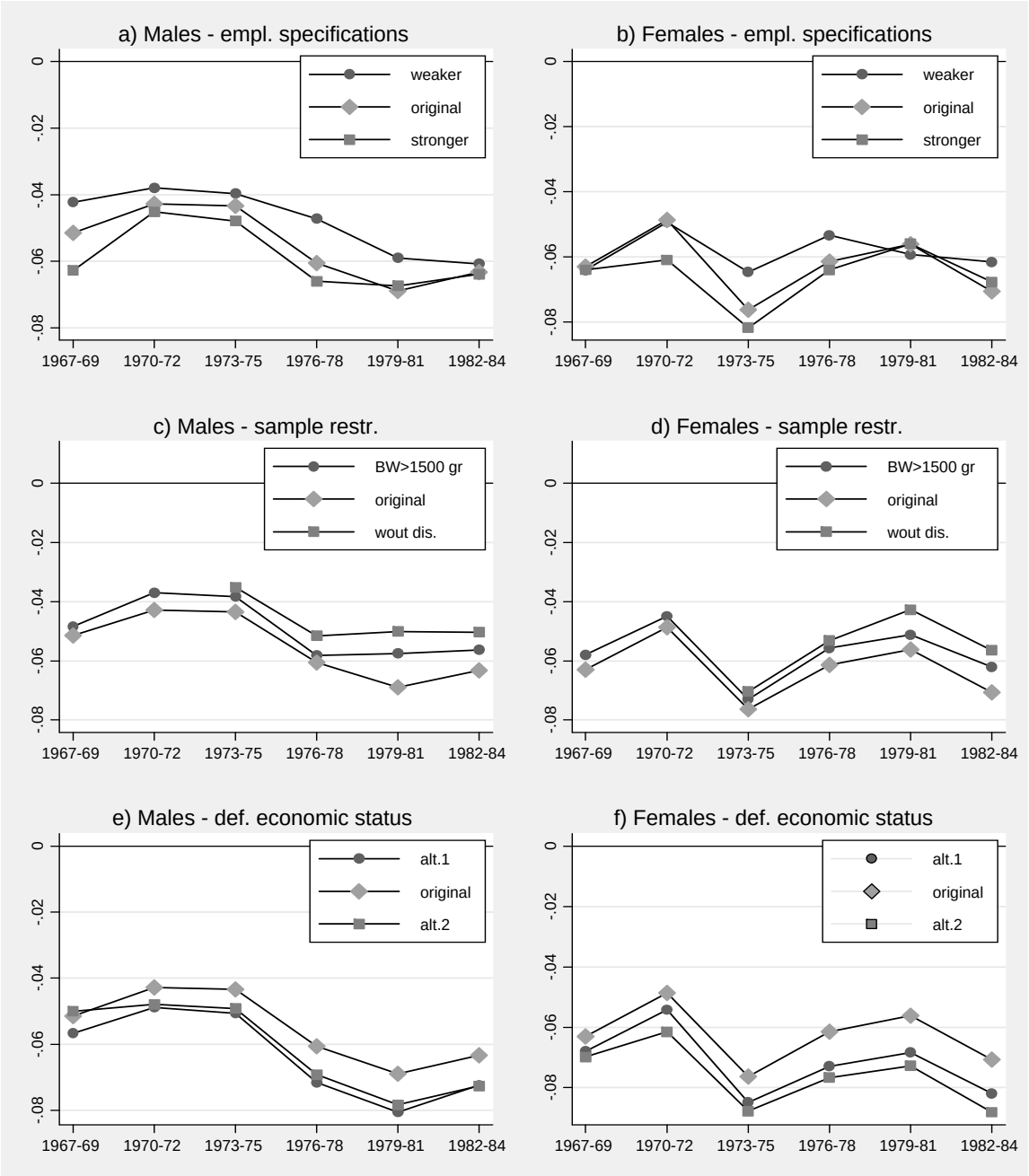



***Figure 18. Estimated effects of low birth weight (<2,500 gr) on employment. Robustness checks.***
*Note; The reference group is birth weight 3,500-4,500 gr. Point estimates are shown in the figure. The point estimates from figure 6, panel a) for males and panel b) for females are depticted in every panels and labeled by "original." The "weaker" definition of employent in panel a) and b) refers to employment defined as having average annual earnings above NOK93,000 ($\approx$ US$11,500) at age 27-31 while the stronger definition refers to having average annual earnings above NOK279,000 ($\approx$ US$44,500) at age 27-31. Panel c) and d) compare the "original" estimate with estimates from models where indididuals with birth weight less than 1,500 gr and persons receiving disability insurance at age 19 are excluded. Panel e) and f) compare the "original" estimate with estimates from models where social background is defined by parents' education level only (alternativ 1) and parents' earnings rank at birth of the child (alternativ 2).*
*The control variables are the same as those reported for table 3.*

# 8 Concluding remarks

The intention with this paper is to investigate whether it has become more difficult for vulnerable groups to be part of the labor market in Norway. I define vulnerable groups as individuals having either poor health, low cognitive ability or coming from low socioeconomic classes. My health indicators are birth weight which is registered for almost all residents in addition to height and BMI measured at age 18-19 for men entering the military service, while cognitive ability is also measured for men entering the military service. Socioeconomic class is defined according to parental earnings rank. Employment is defined as having average annual earnings above a certain amount during the age 27-31.[29]

Using variation across cohorts born during the period 1967-84 I find clear evidence of steadily increasing negative impact on labor market performance of being born into the very poorest families. This conclusion is robust with respect to the inclusion of health indicators, such as birth weight, and indicators for cognitive ability. Hence, low socioeconomic position appears to have become a more important handicap in the labor market over time, beyond its influence on health and ability.

I also find some indications of stronger impact of health on labor market performance among males, but the results are sensitive to what kind of health indicators I use. Lower birth weight and underweight affect employment and earnings to a greater extent over time, while there is no evidence of increasing impacts of low height or obesity. One interpretation of the results is that some aspects of health has become more important, while other aspects have become less important. Moreover, the health indicators I use are obviously far from capturing all dimensions of health, and they are also likely to be correlated with other characteristics that are valued in the labor market. For example, height has been found to be correlated with cognitive and non-cognitive skills, muscular strength and social background. In my study, I control for cognitive skills and social background while I have no information about non-cognitive skills and muscular strength, indicating that some of the effect of low height is affected by these factors.

I find no evidence whatsoever that having low cognitive ability, as measured by IQ tests at age 18/19, has become a larger handicap in the labor market. Overall, the influence of cognitive

[29] This certain amount is NOK186,000 (≈US$23,000)

ability appears to have declined, as the employment and earnings premium associated with having very high ability decreased over time.[30]

According to the point estimates, poor social background is a much stronger predictor of low earnings and non-employment than any of the health variables. However, as the health indicators used in this study are likely to be noisy, and thus only weakly correlated with the actual health status relevant for labor market performance, this does not rule out an important role for health. To the contrary, it is probable that parts of the association between social background and earnings/employment may be explained by poor health among children from poor socioeconomic class, as well as class-related differences in health at higher ages.

Has the labor market become less inclusive and subjected to more sorting? Most of the evidence provided in this paper suggests that the answer to this question is yes. In terms of family background, it is clearly the case that the social gradient in labor market performance has become steeper, particularly at the bottom of the socioeconomic status distribution. This pattern is only slightly modified by controlling for observed health indicators and cognitive ability. When it comes to the influence of health itself, the evidence is a bit more mixed. While the findings reported in this paper point toward an increased importance of health for men, there is no evidence of increased importance for women.

[30] The effect of low ability compared to medium ability has been quite stable.

## References


Abel, K. M., Wicks, S., Susser, E. S., Dalman, C., Pedersen, M. G., Mortensen, P. B. and Webb, R. T. (2010). Birth Weight, Schizophrenia, and Adult Mental Disorder. Is Risk Confined to the Smallest Babies? *Archives of General Psychiatry,* Vol. 67(9), 923-930

Acemoglu, D. and Author, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. *Handbook of Labor Economics*, 4 (2011), 1043-1171.

Almond, D and Mazumder, B. (2013). Fetal Origins and Parental Responses. *Annual Review of Economics,* Vol. 5, 37-56.

Almond, D., Chay, K. and Lee, D. (2005). The Costs of Low Birth Weight. *Quarterly Journal of Economics,* Vol. 120(3), 1031-1083.

Atella, V., Pace, N. and Vuri, D (2008). Are employers discriminating with respect to Weight? European Evidence using Quantile Regression. *Economics and Human Biology,* Vol. 6(3), 305-329.

Aune, D., Sen, A., Prasad, A., Norat, T., Janszky, I., Tonstad, S., Romundstad, P. and Vatten, L. J. (2016). BMI and all cause Mortality: Systematic Review and Non-Linear Dose-Response Meta-Analysis of 230 Cohort Studies with 3.74 Million Deaths among 30.3 Million Participants. *BMJ* 2016;353:i2156.

Barker, D. J (2006). Adult Consequences of Fetal Growth Restriction. *Clinical Obstetrics and Gynecology*, Vol. 49(2), 270-283.

Beaudry, P., Green, D. A. and Sand, B. M. (2014). The Declining Fortunes of the Young since 2000. *American Economic Review, Papers and Proceedings,* Vol. 104(5), 381-386.

Behrman, J. R. and Rosenzweig, M. R. (2004). Returns to Birthweight. *Review of Economics and Statistics,* Vol. 86(2), 586-601.

Van den Berg, G., Lundborg, P., Nystedt, P and Rooth, D.-O. (2014). Critical Periods during Childhood and Adolescence. *Journal of the European Economic Association,* Vol. 12(6), 1521-1557.

Bharadwaj, P., Lundborg, P. and Rooth, D.-O. (2018). Birth Weight in the Long Run. *Journal of Human Resources,* Vol. 53(4), 189-231.

Bharadwaj, P., Løken, K. V. and Neilson, C. (2013). Early Life Health Interventions and Academic Achievement. *American Economic Review*, Vol. 103(5), 1862-1891.

Black, S. E., Breining, S., Figlio, D., Guryan, J., Karbownik, K., Nielsen, H. S, Roth, J. and Simonsen, M. (2017). Sibling Spillovers. *NBER Working Paper 23062.*

Black, S. E., Devereux, P. J. and Salvanes, K. G. (2007). From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes. *Quarterly Journal of Economics*, Vol. 122(1), 409-439.

Black, S. E., Devereux, P. J. and Salvanes, K. G. (2005). The More the Merrier? The Effects of Family Size and Birth Order on Children's Education. *Quarterly Journal of Economics*, Vol. 120(2), 669-700.

Bound, J. and Johnson, G. (1992). Changes in the Structure of Wages in the 1980's: An Evaluation of Alternative Explanations? *American Economic Review,* Vol. 82(3), 371-392.

Brage, S. and Thune, O. (2015). Ung Uførhet og Psykisk Sykdom. *Arbeid og Velferd*, No. 1, 2015.

Bratsberg, B., Fevang, E., and Røed, K. (2013). Job Loss and Disability Insurance. *Labor Economics*, Vol. 24, 137-150.

Breining, S. (2014). The Presence of ADHD: Spillovers between Siblings. *Economics Letters*, Vol. 124 (3), 469–473.

Case, A. and Paxson, C. (2008). Stature and status: Height, Ability, and Labor Market Outcomes. *Journal of Political Economy*, Vol. 116(3), 499-532.

Case, A., Fertig, E. and Paxson, C. (2005). The Lasting Impact of Childhood Health and Circumstance. *Journal of Health Economics*, Vol. 24, 365-389.

Case, A, Lubotsky, D. and Paxson, C. (2002). Economic Status and Health in Childhood: The origins of the gradient. *American Economic Review,* Vol. 92(5), 1308-1334.

Cnattingius, S, Villamor, E., Lagerros, Y. T, Wikström, A-K. and Granath F. (2012). High Birth Weight and Obesity- a Vicious Circle across Generations. *International Journal of Obesity*, Vol.36, 1320-1324.

Currie, J., Stabile, M., Manivong, P and Roos, L. L. (2010). Child Health and Young Adult Outcomes. *Journal of Human resources,* Vol. 45(3), 517-548.

Currie, J. (2009). Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital development. *Journal of Economic Literature,* Vol. 47(1), 87-122.

Currie, J. and Moretti, E. (2007). Biology as Destiny? Short- and Long-Run Determinants of Intergenerational Transmission of Birth Weight. *Journal of Labor Economics,* Vol. 25(2), 231-264.

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

Currie, J and Hyson, R. (1999a). Is the Impact of Health Shocks Cushioned by Socio-Economic Status? The Case of Low Birth Weight. *American Economic Review, Papers and Proceedings,* Vol. 89(2), 245-250.

Currie, J and Hyson, R. (1999b). Is the Impact of Health Shocks Cushioned by Socio-Economic Status? The Case of Low Birth Weight. *NBER Working Paper No. 6999.*

Currie, J and Madrian, B. C. (1999). Health, health insurance and the labor market. *Handbook of labor economics*, Vol. 3, 3309–3416.

Deming, D. (2017). The Growing Importance of Social Skills in the Labor Market. *Quarterly Journal of Economics,* Vol. 132(4), 1593-164.

Edin, P. E., Fredriksson, P., Nybom, M. and Öckert, B (2017). The Rising Return to Non-Cognitive Skill. *IZA Discussion Paper No. 10914.*

Eide, M., Øyen, N., Skjærven, R. Nilsen, S. T., Bjerkedal, T. and Tell, G. S. (2005). Size at Birth and Gestational Age as Predictors of Adult Height and Weight. *Epidemiology*, Vol.16(2), 175-181.

Elo, I. T. and Preston, S. H. (1992). Effects of Early-Life Conditions an Adult Mortality: A Review. *Population Index,* Vol. 58, 186-212.

European Commision (2017). http://ec.europa.eu/social/main.jsp?catId=1134&langId=en.

Evensen, M., Lyngstad, T. H., Melkevik, O. Reneflot, A. and Mykletun, A. (2017). Adolescent Mental Health and Earnings Inequalities in Adulthood: Evidence from the Young-HUNT Study. *Journal of Epidemiology and Community Health.* Vol. 71, 201-206.

Evensen, M., Lyngstad, T. H., Melkevik, O. and Mykletun, A. (2016). The Role of Internalizing and Externalizing Problems in Adolescence for Adult Educational Attainment: Evidence from Sibling Comparisons using Data from the Young HUNT Study. *European Sociological Review.* Vol. 32(5), 552-566.

Ferrie. J. E. (2001). Is Job Insecurity Harmful to Health? *Journal of the Royal Society of Medicine*, Vol. 94, 71-76.

Figlio, D. N., Guryan, J., Karbownik, K. and Rooth, J. (2014). The Effects of Poor Neonatal Health on Children's Cognitive Development. *American Economic Review,* Vol. 104(12), 3921-3955.

Flegal, K. M., Kit, B. K., Orpana,H. and Graubard, B. I. (2013). Association of All-Cause Mortality with Overweight and Obesity Using Standard Body Mass Index Categories: A Systematic Review And Meta-Analysis. *Journal of the American Medical Association*, Vol. 309(1), 71-82.

Flensborg- Madsen, T. and Mortensen, E. L. (2017). Birth Weight and Intelligence in Young adulthood and Midlife. *Pediatrics,* Vol.139(6): e20163161

Fletcher, J., Hair, N. L. and Wolfe, B. L. (2012). Am I my Brother's Keeper? Sibling Spillover Effects: The Case of Developmental Disabilities and Externalizing Behavior. NBER Working Paper, No. 18279.

Goodman, A., Joyce, R. and Smith, J. P (2011). The Long Shadow Cast by Childhood Physical and Mental Problems on Adult Life. *Proceedings of the National Academy of Sciences,* Vol. 108(15), 6032-6037.

Harder, T., Rodekamp, E., Schellong, K., Dudenhausen, J. W. and Plagemann, A (2007). Birth Weight and Subsequent Risk of Type 2 Diabetes: A Meta-Analysis. *American Journal of Epidemiology.* Vol. 165(8), 849-857.

Hughes, A. and Kumari, M (2017). Unemployment, Underweight and Obesity: Findings from Understanding Society (UKHLS). *Preventive Medicine.* Vol. 97, 19-25.

Johnston, D. W., Schurer and Shields, M. A. (2011). Evidence on the Long Shadow of Poor Mental Health across Three Generations. *IZA Discussion Paper No. 6014.*

Jornayvaz, F. R., Vollenweider, P, Bochud, M., Mooser, V., Waeber, G. and Marques-Vidal, P. (2016). Low Birth Weight Leads to Obesity, Diabetes and Increased Leptin Levels in Adults: The Colaus Study. *Cardiovascular Diabetology*, 15:73.

Kelly, S. J, Lilley, J. M and J. Leonardie-Bee (2010). Associations of Morbidity in the Underweight. *European Journal of Clinical Nutrition,* Vol. 64, 475-482.

Kivimäki, M., Vahtera, J., Ferrie, J. E., Hemingway, H., and Pentti, J. (2001). Organisational Downsizing and Musculoskeletal Problems in Employees: A Prospective Study. *Occupational and Environmental Medicine*, Vol. 58 (12), 811-817.

Linnet, K. M., Wisborg, K., Agerbo, E., Secher, N. J., Thomsen, P. H. and Henrikssen, T. B. (2006). Gestational Age, Birth Weight, and The Risk of Hyperkinetic Disorder. *Archives of Disease in Childhood*, Vol. 91(8): 655-660.

Lundborg, P., Nystedt, P. and Roth, D.-O. (2014). Health and Earnings: The Role of Cognitive and Noncognitive Skills. *Journal of Human Resources,* Vol..49(1), 141-166.

Markussen, S. and Røed, K. (2017). Egalitarism under Pressure: Toward Lower Economic Mobility in the Knowledge Economy? *IZA Discussion Paper No. 10664.*

Minton, J. W., Pickett, K. and Dorling, D. (2012). Health, Employment, and Economic Change, 1973-2009: Repeated Cross Sectional Study. *British Medical Journal* 2012:340:e2316.

Morris, S. (2007). The Impact of Obesity on Employment. *Labor Economics*, Vol. 14(3), 413-433.

Persico, N., Postlewaite, A. and Silverman, D. (2004). The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height. *Journal of Political Economy*, Vol. 112(5), 1019-1053.

Rege, M., Telle, K., and Votruba, M. (2009). The Effect of Plant Downsizing on Disability Pension Utilization. *Journal of the European Economic Association,* Vol. 7(4), 754–785.

Roh, L., Braun, J., Chiolero, A., Bopp, M., Rohrmann S. and Faeh, D. (2014). Mortality Risk Associated with Underweight: A Census-Linked Cohort of 31,578 Individuals with up to 32 Years of Follow-up. *BMC Public Health*, 14:371.

Royer, H. (2009). Separated at Girth: US Twin Estimates of the Effects of Birth Weight. *American Economic Journal: Applied Economics*, Vol. 1(1), 49-85.

Seidman, D. S., Ever-Hadani, P., Stevenson, D. K., Slater, P. E., Harlap, S. and Gale, R. (1988). Birth Order and Birth Weight Reexamined. *Obstetrics Gynecology*; 72 (2): 158-62.

Smith, J. (2009). The Impact of Childhood Health on Adult Labor Market Outcomes. *Review of Economics and Statistics,* Vol. 91(3): 478-489.

Sundet, J. M., Barlaug, D. G. and Torjussen, T. M. (2004). The End Of The Flynn Effect? A Study of Secular Trends in Mean Intelligence Test Scores of Norwegian Conscripts During Half A Century. *Intelligence* Vol. 32(4), 349-362.

Yi, J., Heckman, J. J., Zhang, J. and Conti, G. (2015). Early Health Shocks, Intra-household Resource Allocation and Child Outcomes. *Economic Journal,* Vol. 25(588), 347-371.

# Appendix

***Appendix table 1a). Health indicators and parental education. Males****

| Year of birth | 1967-69 | 1970-72 | 1973-75 | 1976-78 | 1979-81 | 1982-84 |
|---|---|---|---|---|---|---|
| *N* | *91,656* | *88,533* | *80,517* | *70,577* | *68,664* | *67,918* |
| Birth weight categories | | | | | | |
| <2,500 gr | 34.80 | 39.06 | 45.34 | 48.56 | 56.80 | 60.75 |
| 2,500-3,000 gr | 34.78 | 40.07 | 46.04 | 51.99 | 57.75 | 64.30 |
| 3,000-3,500 gr | 37.94 | 42.79 | 49.23 | 55.57 | 61.05 | 67.10 |
| 3,500-4,500 gr | 38.82 | 44.12 | 50.99 | 57.72 | 62.92 | 68.95 |
| >4,500 gr | 36.40 | 42.30 | 48.69 | 54.40 | 61.30 | 67.95 |
| Missing height | 31.73 | 37.18 | 42.14 | 48.12 | 54.07 | 57.37 |
| Height group | | | | | | |
| <170 cm | 28.64 | 34.35 | 41.50 | 46.88 | 52.85 | 61.06 |
| 170-180 cm | 35.26 | 40.71 | 47.84 | 53.57 | 59.50 | 65.94 |
| 180-190 cm | 40.53 | 45.25 | 52.09 | 59.06 | 64.13 | 69.95 |
| >190 cm | 45.62 | 51.53 | 55.77 | 61.70 | 67.52 | 74.07 |
| BMI classification | | | | | | |
| Underweight (<18.5) | 37.81 | 43.34 | 48.40 | 55.43 | 60.44 | 65.72 |
| Normal weight (18.5-25) | 39.25 | 44.69 | 51.73 | 58.08 | 63.65 | 69.94 |
| Overweight (25-30) | 31.33 | 36.91 | 43.60 | 50.69 | 56.83 | 65.25 |
| Obesity (>30) | 25.85 | 29.25 | 35.30 | 41.70 | 48.51 | 57.22 |
| Missing IQ | 28.59 | 33.08 | 43.40 | 47.04 | 55.51 | 62.07 |
| IQ classification | | | | | | |
| Low IQ (1-3) | 19.30 | 22.89 | 29.23 | 34.38 | 41.90 | 49.88 |
| Middle IQ (4-6) | 37.19 | 42.16 | 48.36 | 55.35 | 61.57 | 68.46 |
| High IQ (7-9) | 58.81 | 64.33 | 69.63 | 75.06 | 79.66 | 83.94 |

*Note; Parental education is measured as fraction (in percent) where at least one of the parents has completed high shool*

***Appendix table 1b). Health indicators and parental education. Females****

| Year of birth | 1967-69 | 1970-72 | 1973-75 | 1976-78 | 1979-81 | 1982-84 |
|---|---|---|---|---|---|---|
| *N* | *87,642* | *84,659* | *77,174* | *67,344* | *65,555* | *64,209* |
| Birth weight categories | | | | | | |
| <2,500 gr | 33.16 | 37.09 | 44.53 | 48.11 | 53.17 | 61.65 |
| 2,500-3,000 gr | 35.94 | 40.26 | 45.74 | 52.18 | 57.15 | 63.39 |
| 3,000-3,500 gr | 38.16 | 43.49 | 49.00 | 55.40 | 61.80 | 67.11 |
| 3,500-4,500 gr | 39.12 | 44.74 | 51.18 | 57.12 | 63.30 | 69.09 |
| >4,500 gr | 37.41 | 42.63 | 50.37 | 54.43 | 59.99 | 68.01 |

*Note; Parental education is measured as fraction (in percent) where at least one of the parents has completed high shool*

***Appendix table 2. List of explanatory variables used in the OLS models.***

Birth weight: 5 dummy variables; <2,500 gr, 2,500-3,000 gr, 3,000-3,500 gr, 3,500-4,500 gr, >4,500 gr.
Height: 5 dummy variables; <170 cm, 170-180 cm, 180-190 cm, >190 cm, missing.
Body Mass Index (BMI): 5 dummy variables; underweight (BMI <18.5), normal weight (BMI 18.5-25), overweight (BMI 25-30), obesity (BMI >30), missing.
Cognitive ability: 4 dummy variables; low ability (stanine score 1-3), medium ability (stanine score 4-6), high ability (stanine score 7-9), missing.
Parental earnings' rank: 20 bins.
Multiple birth: 1 dummy variable.
Birth order: 4 dummy variables; $1^{st}$, $2^{nd}$, $3^{rd}$-$4^{th}$, $>4^{th.}$
Siblings: 4 dummy variables; no siblings, 1 sibling, 2-3 siblings, >3 siblings.
Mothers' age at birth: 6 dummy variables; <20, 20-24, 25-29, 30-34, 35-39, 40+.
Month of birth: 12 dummy variables (one for each month).

***Appendix table 3a). Descriptive statistics for brothers (singletons)***

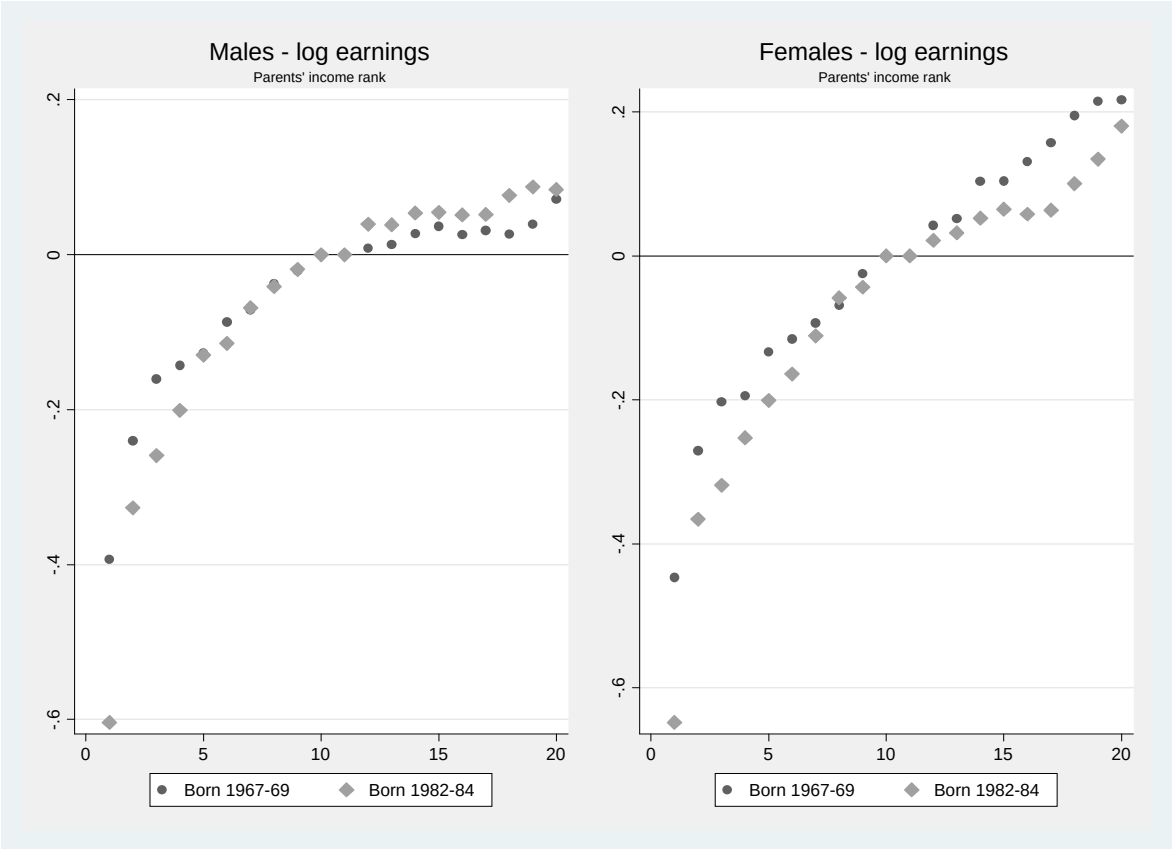
| Year of birth | 1967-72 | 1973-78 | 1979-84 |
|---|---|---|---|
| *N* | *45,329* | *29,494* | *25,628* |
| *Fraction (in percent) of the whole sample* | *25.16* | *19.52* | *18.76* |
| **Individual and background characteristics** | | | |
| Birth weight (in gr) | 3,592 (532) | 3,603 (533) | 3,598 (532) |
| Birth weight categories (percent) | | | |
| <2,500 gr | 2.49 | 2.48 | 2.25 |
| 2,500-3,000 gr | 8.83 | 8.23 | 7.91 |
| 3,000-3,500 gr | 29.65 | 29.61 | 27.99 |
| 3,500-4,500 gr | 54.92 | 55.38 | 57.30 |
| >4,500 gr | 4.11 | 4.30 | 4.56 |
| **Family characteristics** | | | |
| Mother's age at birth | 25.39 (4.67) | 25.61 (4.42) | 25.84 (4.62) |
| Number of siblings | 2.29 (1.34) | 2.07 (1.29) | 2.02 (1.16) |
| **Military data** | | | |
| Height (centimeters) | 179.79 (6.50) | 179.93 (6.54) | 180.13 (6.50) |
| Missing height | 2.15 | 3.19 | 3.34 |
| Height group (percent) | | | |
| <170 cm | 5.31 | 5.16 | 4.85 |
| 170-180 cm | 42.23 | 41.63 | 40.76 |
| 180-190 cm | 45.45 | 45.74 | 46.79 |
| >190 cm | 7.01 | 7.47 | 7.61 |
| BMI | 21.97 (2.89) | 22.23 (3.02) | 22.40 (3.06) |
| BMI classification (percent) | | | |
| Underweight (<18.5) | 6.49 | 5.00 | 6.03 |
| Normal weight (18.5-25) | 80.69 | 79.73 | 75.10 |
| Overweight (25-30) | 10.73 | 12.40 | 14.58 |
| Obesity (>30) | 2.09 | 2.86 | 4.28 |
| IQ (stanines) | 5.06 (1.85) | 5.27 (1.80) | 5.23 (1.73) |
| Missing IQ (percent) | 6.77 | 8.40 | 10.49 |

| *Appendix table 3a). Descriptive statistics for brothers (singletons)* | | | |
|---|---|---|---|
| Year of birth | 1967-72 | 1973-78 | 1979-84 |
| IQ classification (percent) | | | |
| Low IQ (1-3) | 20.21 | 16.18 | 15.84 |
| Middle IQ (4-6) | 57.41 | 58.76 | 60.77 |
| High IQ (7-9) | 22.38 | 25.06 | 23.40 |
| **Outcomes** | | | |
| Percent employed when 27-31 | 90.98 | 89.04 | 87.87 |
| Average annual earnings when 27-31 (NOK, 2017 value) | 497,680 (211,785) | 485,615 (224,951) | 489,554 (223,071) |
| Percent with completed high school | 72.74 | 80.51 | 82.20 |
| Percent receiving permanent disability Insurance when 19 | | 0.47 | 0.57 |

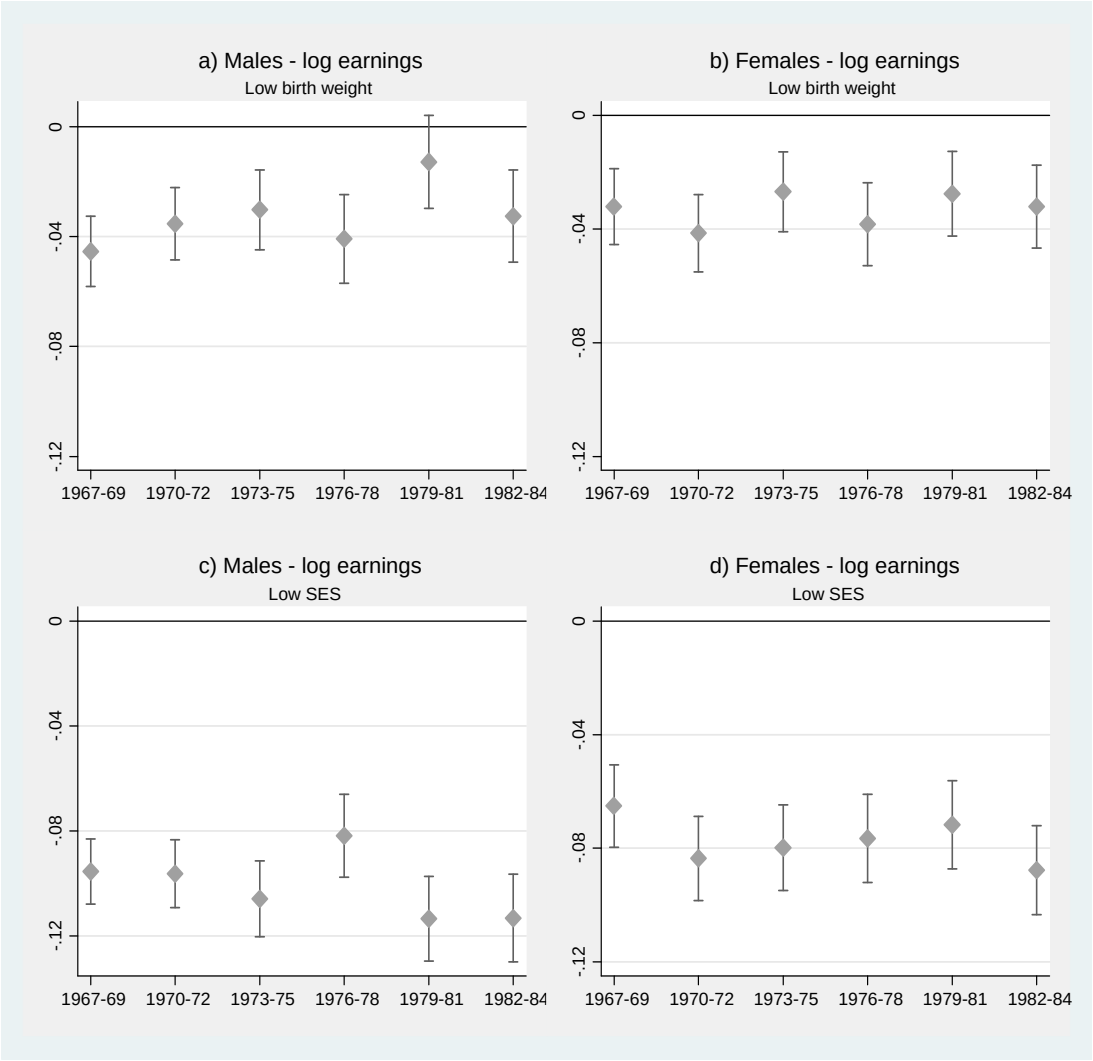
| *Appendix table 3b). Descriptive statistics for sisters (singletons)* | | | |
|---|---|---|---|
| Year of birth | 1967-72 | 1973-78 | 1979-84 |
| *N* | *41,877* | *26,345* | *23,436* |
| *Fraction (in percent) of the whole sample* | *24.30* | *18.23* | *18.06* |
| **Individual and background characteristics** | | | |
| Birth weight (in gr) | 3,462 (505) | 3,481 (507) | 3,517 (501) |
| Birth weight categories (percent) | | | |
| <2,500 gr | 3.07 | 2.81 | 2.35 |
| 2,500-3,000 gr | 12.16 | 12.01 | 10.60 |
| 3,000-3,500 gr | 36.68 | 36.07 | 34.52 |
| 3,500-4,500 gr | 46.06 | 46.92 | 50.17 |
| >4,500 gr | 2.03 | 2.19 | 2.03 |
| **Family characteristics** | | | |
| Mother's age at birth | 25.37 (4.62) | 25.59 (4.42) | 25.84 (4.62) |
| Number of siblings | 2.32 (1.34) | 2.07 (1.28) | 2.02 (1.16) |
| **Outcomes** | | | |
| Percent employed when 27-31 | 74.60 | 79.33 | 87.87 |
| Average annual earnings when 27-31 (NOK, 2017 value) | 320,642 (187,006) | 342,763 (184,542) | 360,302 (180,464) |
| Percent with completed high school | 71.19 | 84.22 | 88.02 |
| Percent receiving permanent disability insurance when 19 | | 0.36 | 0.44 |

| *Appendix table 3c). Descriptive statistics for brothers (twins)* | | | |
|---|---|---|---|
| Year of birth | 1967-72 | 1973-78 | 1979-84 |
| *N* | *1,924* | *1,612* | *1,622* |
| *Fraction (in percent) of the whole sample* | *1.07* | *1.07* | *1.12* |
| **Individual and background characteristics** | | | |
| Birth weight (in gr) | 2,705 (541) | 2,718 (519) | 2,668 (562) |
| Birth weight categories (percent) | | | |
| <2,500 gr | 33.37 | 32.13 | 36.07 |
| 2,500-3,000 gr | 35.14 | 37.53 | 34.71 |
| 3,000-3,500 gr | 24.79 | 24.32 | 23.43 |
| 3,500-4,500 gr | 6.65 | 6.02 | 5.67 |
| >4,500 gr | 0.05 | 0.00 | 0.12 |
| **Family characteristics** | | | |
| Mother's age at birth | 27.52 (5.62) | 27.51 (4.82) | 27.60 (4.87) |
| Number of siblings | 2.66 (1.32) | 2.41 (1.07) | 2.40 (1.00) |
| **Military data** | | | |
| Height (centimeters) | 179.27 (6.47) | 179.28 (6.69) | 179.07 (6.42) |
| Missing height | 2.44 | 3.04 | 5.07 |
| Height group (percent) | | | |
| <170 cm | 6.23 | 6.01 | 6.59 |
| 170-180 cm | 44.01 | 46.32 | 45.06 |
| 180-190 cm | 43.90 | 39.92 | 42.84 |
| >190 cm | 5.86 | 7.74 | 5.51 |
| BMI | 21.56 (2.72) | 21.84 (2.75) | 22.07 (3.11) |
| BMI classification (percent) | | | |
| Underweight (<18.5) | 9.64 | 6.97 | 6.97 |
| Normal weight (18.5-25) | 81.57 | 83.30 | 79.72 |
| Overweight (25-30) | 7.09 | 8.06 | 11.22 |
| Obesity (>30) | 1.70 | 1.66 | 2.09 |
| IQ (stanines) | 4.94 (1.87) | 5.08 (1.80) | 5.07 (1.74) |
| Missing IQ (percent) | 6.50 | 8.31 | 10.97 |
| IQ classification (percent) | | | |
| Low IQ (1-3) | 21.79 | 18.94 | 18.21 |
| Middle IQ (4-6) | 58.42 | 59.27 | 61.77 |
| High IQ (7-9) | 19.79 | 21.79 | 20.01 |
| **Outcomes** | | | |
| Percent employed when 27-31 | 90.70 | 89.70 | 89.70 |
| Average annual earnings when 27-31 (NOK, 2017 value) | 492,172 (210,777) | 490,158 (219,953) | 491,054 (223,713) |
| Percent with completed high school | 73.18 | 82.75 | 84.09 |
| Percent receiving permanent disability insurance when 19 | | 0.06 | 0.55 |

| *Appendix table 3d). Descriptive statistics for sisters (twins)* | | | |
|---|---|---|---|
| Year of birth | 1967-72 | 1973-78 | 1979-84 |
| *N* | *1,920* | *1,762* | *1,584* |
| *Fraction (in percent) of the whole sample* | *1.11* | *1.22* | *1.22* |
| **Individual and background characteristics** | | | |
| Birth weight (in gr) | 2,590 (516) | 2,603 (546) | 2,580 (550) |
| Birth weight categories (percent) | | | |
| <2,500 gr | 41.46 | 39.33 | 40.97 |
| 2,500-3,000 gr | 37.50 | 35.13 | 37.56 |
| 3,000-3,500 gr | 16.98 | 21.91 | 17.61 |
| 3,500-4,500 gr | 4.06 | 3.63 | 3.85 |
| >4,500 gr | 0.00 | 0.00 | 0.00 |
| **Family characteristics** | | | |
| Mother's age at birth | 27.42 (5.62) | 27.09 (5.11) | 27.70 (4.90) |
| Number of siblings | 2.71 (1.32) | 2.48 (1.19) | 2.41 (1.06) |
| **Outcomes** | | | |
| Percent employed when 27-31 | 76.20 | 80.53 | 83.08 |
| Average annual earnings when 27-31 (NOK, 2017 value) | 324,379 (181,217) | 345,606 (174,926) | 369,469 (186,074) |
| Percent with completed high school | 71.51 | 84.68 | 87.21 |
| Percent receiving permanent disability insurance when 19 | | 0.40 | 0.63 |

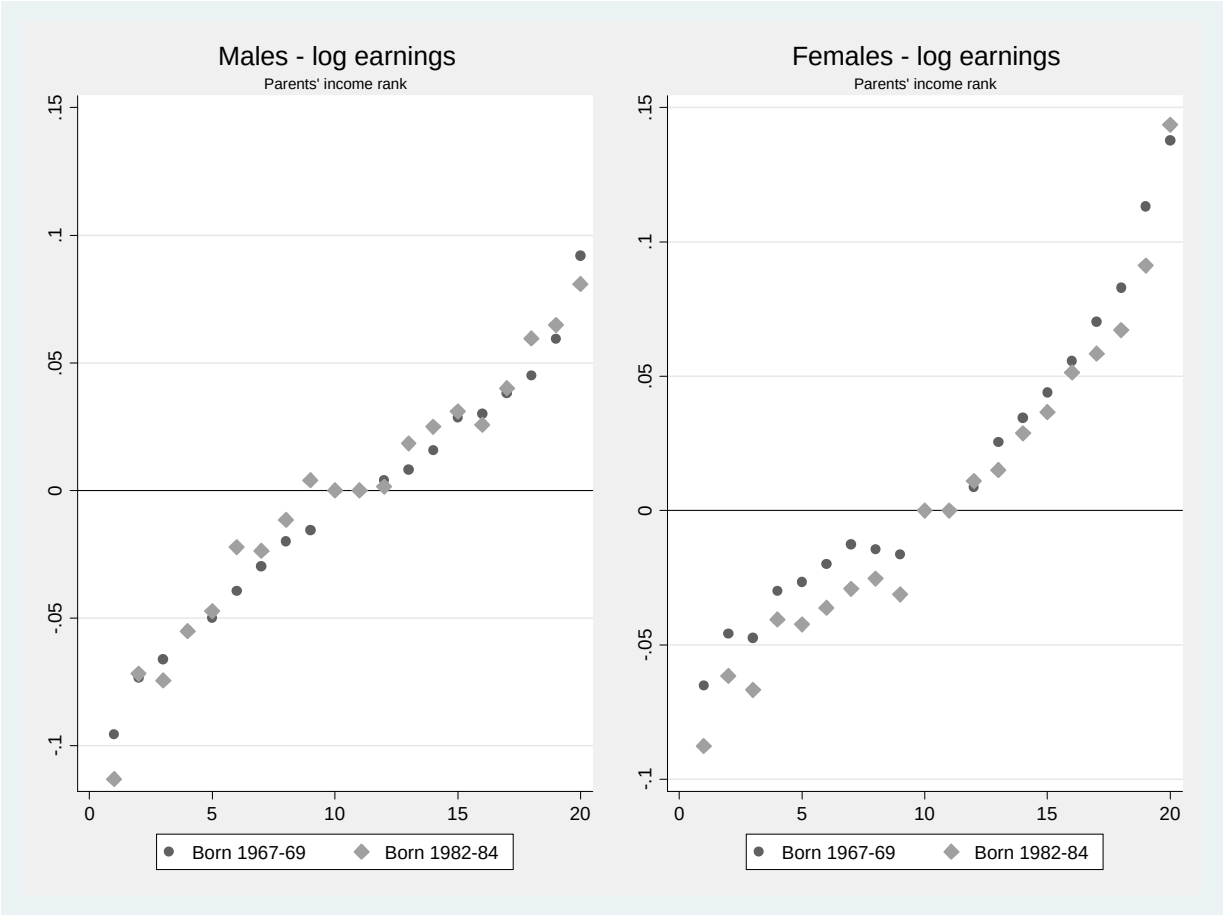



***Appendix figure 1. Estimated effects of parental earnings rank on log earnings at age 27-31.***
*Note; Socioeconomic status (SES) is divived in 20 vigintiles based on parental earnings' rank. Vigintile 10 and 11 is the reference group. The control variables are the same as those reported for table 3.*

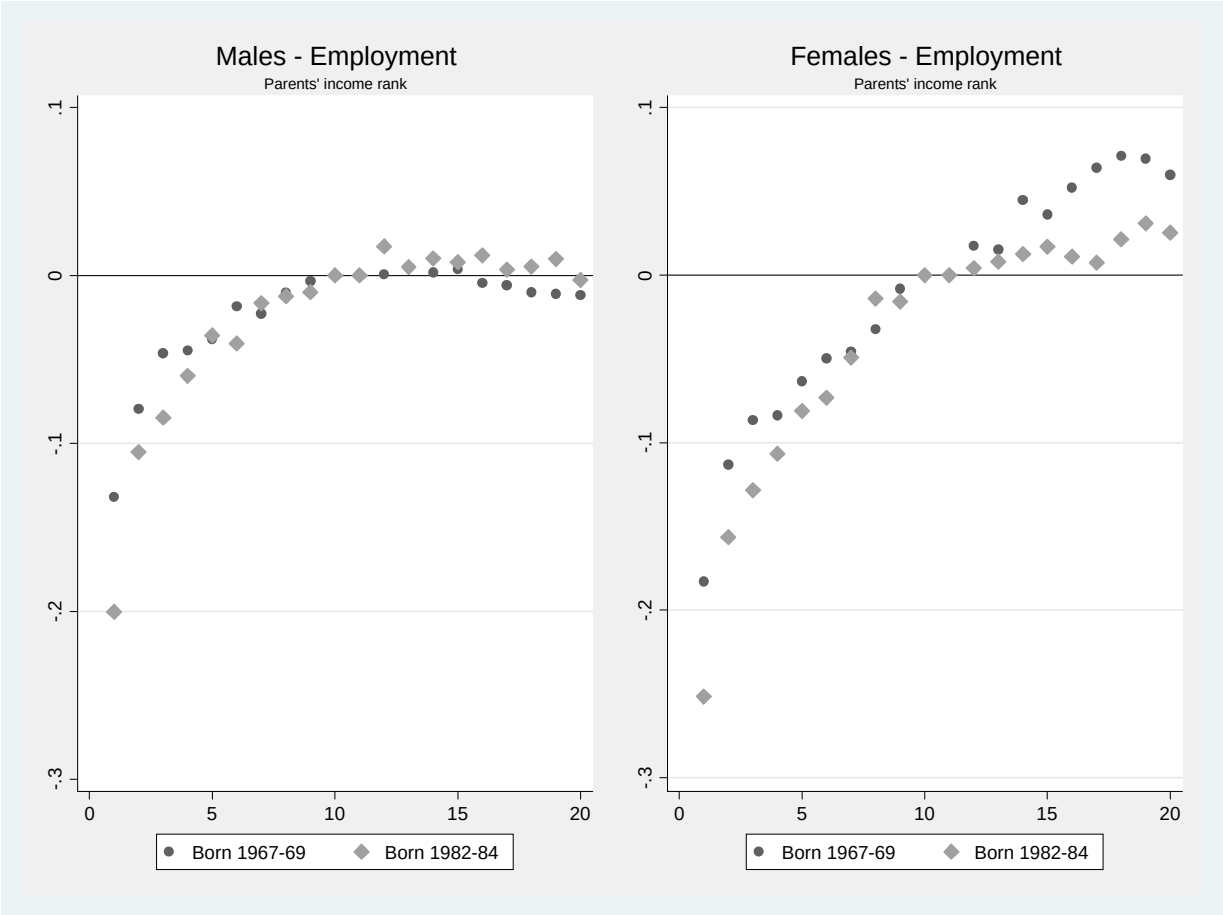



***Appendix figure 2. Estimated effects of birth weight <2,500 gr and low SES. Dependent variable log earnings at age 27-31. Employed individuals, only.***

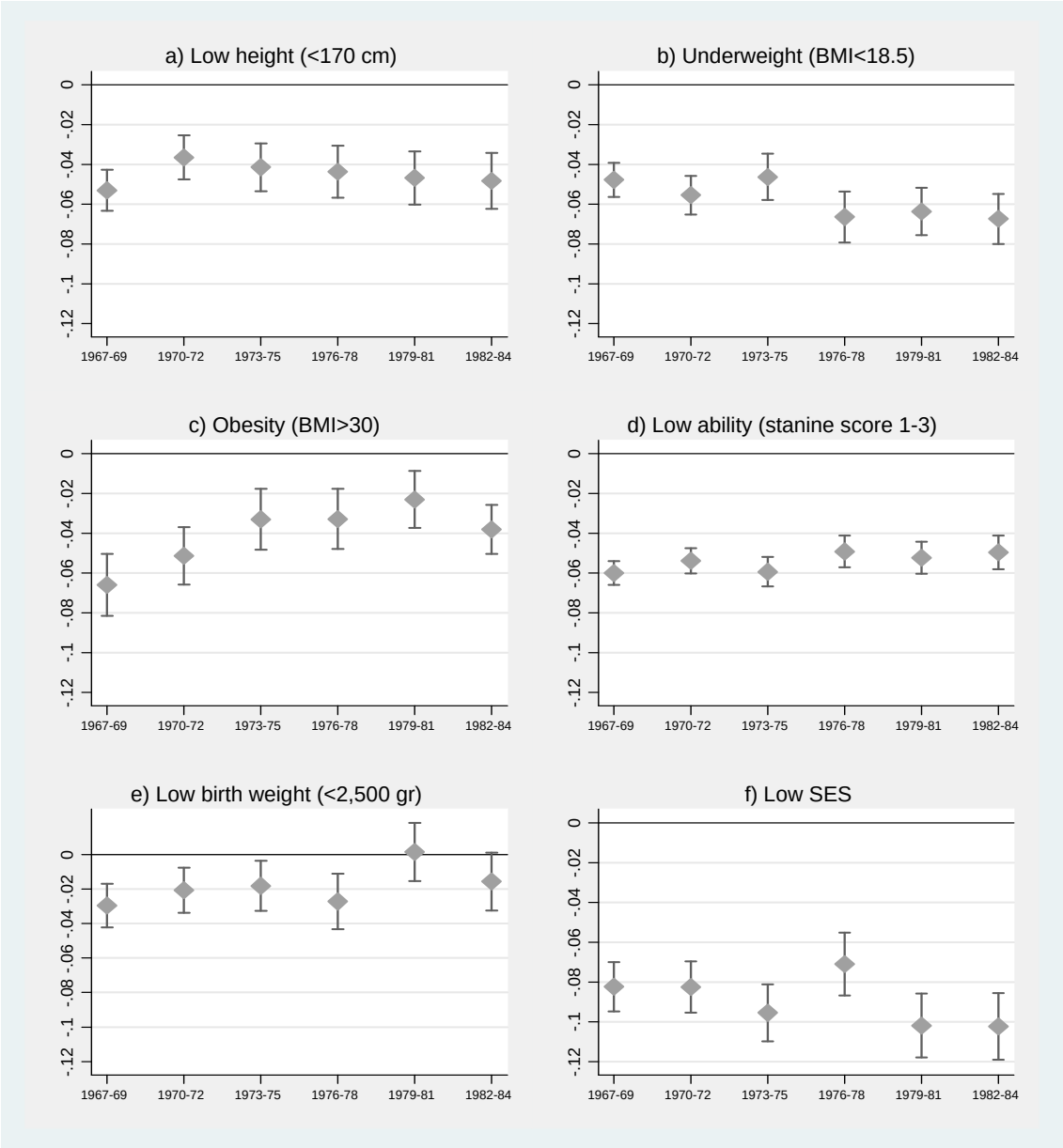
*Note; The reference group for birth weight is 3,500-4,500 gr. Socioeconomic status (SES) is divived in 20 vigintiles based on parental earnings' rank. Vigintile 10 and 11 is the reference group. Low SES is vigintile 1, which is the five percent with lowest parental earnings. Point estimates with confidence interval are shown in the figure. The control variables are the same as those reported for table 3.*




***Appendix figure 3. Estimated effects of parental earnings rank on log earnings at age 27-31. Employed individuals, only.***
*Note; Vigintile 10 and 11 is the reference group. The control variables are the same as those reported for table 3.*



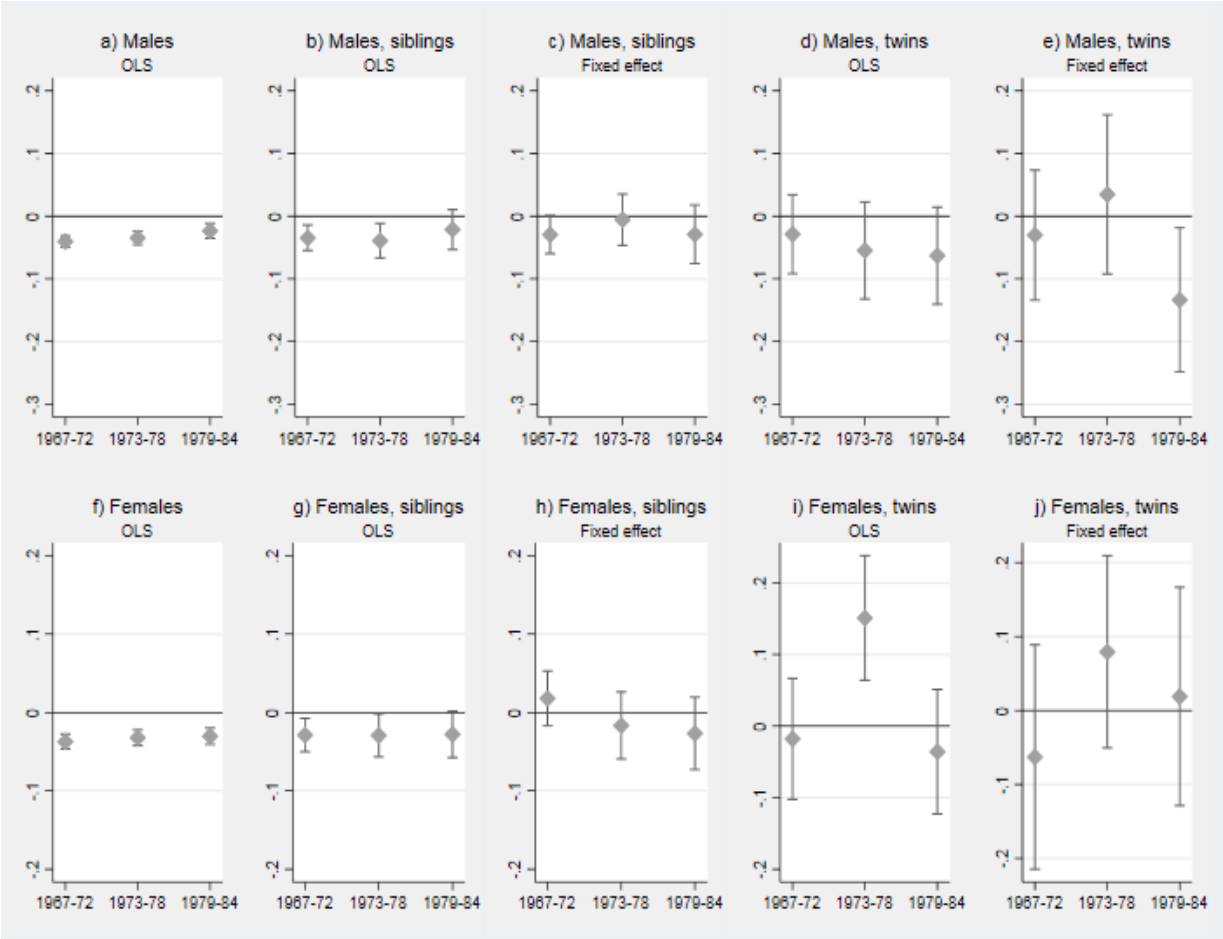

***Appendix figure 4. Estimated effects of parents' earnings rank on employment***
*Note; Vigintile 10 and 11 is the reference group. The control variables are the same as those reported for table 3.*




***Appendix figure 5. Estimated effects of various health-indicators on log earnings at age 27-31. Males. Employed individuals only.***

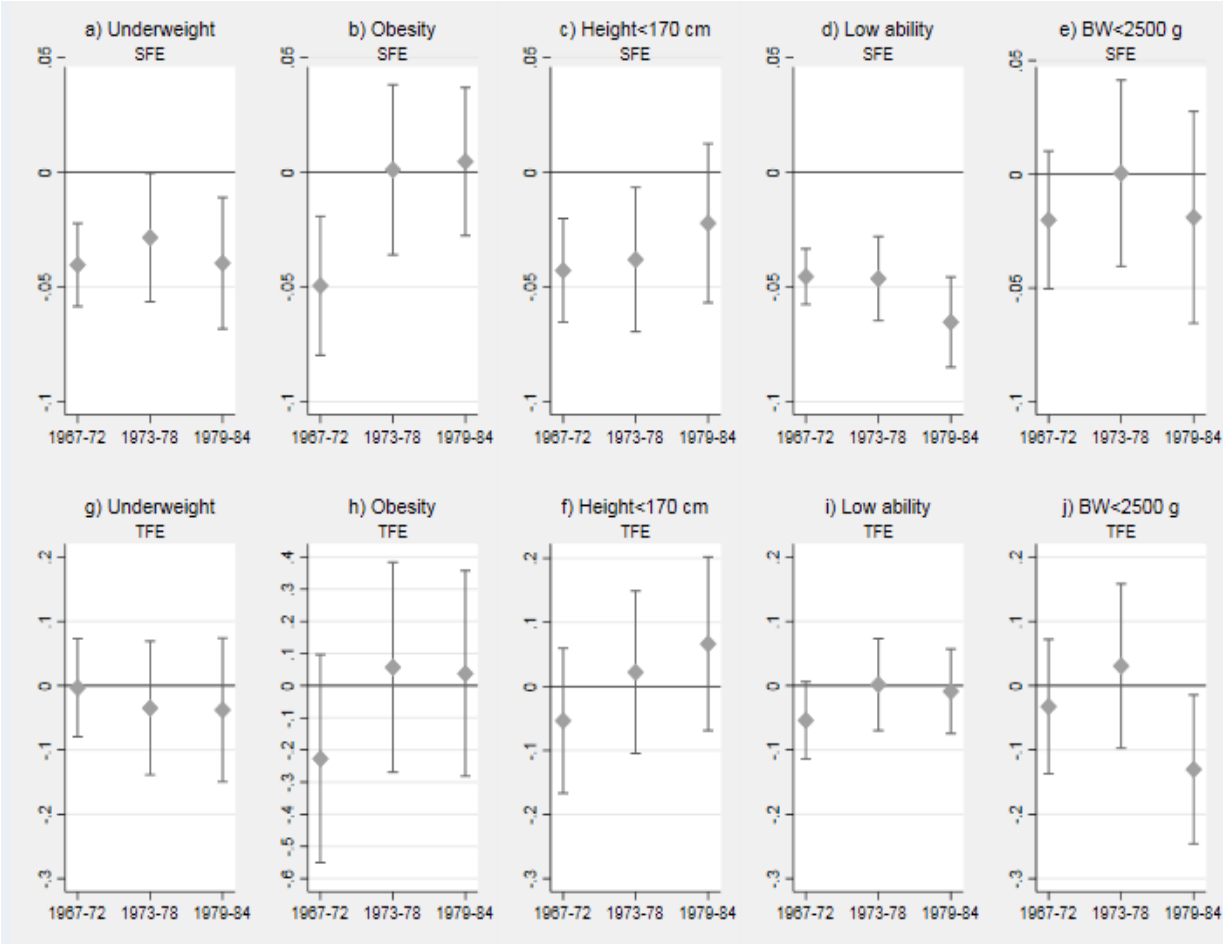
*Note; Reference groups are the following; height:180-190 cm, BMI-group: normal weight (BMI 18.5-25), ability: medium ability (stanine score 4-6), birth weight-category: 3,500-4,500 gr. Socioeconomic status (SES) is divived in 20 vigintiles based on parental earnings' rank. Vigintile 10 and 11 is the reference group. Low SES is vigintile 1, which is the five percent with lowest parental earnings. Point estimates with confidence interval are shown in the figure. The control variables are the same as those reported for table 3.*

.




***Appendix figure 6. Estimated effects of low birth weight (<2,500 gr) on log earnings at age 27-31. Comparison with family fixed effects. Employed individuals only.***

*Note; The reference group is birth weight 3,500-4,500 gr. Results from separate estimations on each cohortgroup. Point estimates with confidence interval are shown in the figure. Control variables for OLS; see note to table 3. For sibling fixed effects; birth order (4 dummy variables), month of birth (12 dummy variables). For twin fixed effects; none.*




***Appendix figure 7. Estimated effects of health- and ability indicators on log earnings at age 27-31. Family fixed effects. Males. Employed individuals only.***

*Note; SFE=sibling fixed effects, TFE=twin fixed effects. Reference groups are the following; height:180-190 cm, BMI-group: normal weight (BMI 18.5-25), ability: medium ability (stanine score 4-6), birth weight: 3,500-4,500 gr. Point estimates with confidence interval are shown in the figure. Control variables for sibling fixed effects; birth order (4 dummy variables), month of birth (12 dummy variables). For twin fixed effects; none.*