Three Economic Essays on Victimization and Social Policies from Childhood to Retirement

Mirjam Wentzel

Acknowledgements

First and foremost, I would like to thank my main supervisor, Katrine V. Løken. For always taking time to guide me in both research and life in general. Your knowledge and advice has been invaluable, and I could not have wished for a better supervisor.

Furthermore, I would like to thank Aline Bütikofer, for helpful comments on my work, and for your overall support and help. I am also grateful to my other colleagues in the CELE group, for contributing with interesting discussions and an inspiring work environment.

I would also like to thank my second supervisor, Matthew Lindquist, for welcoming me to Stockholm University, and for helping me develop my research ideas, especially my single-authored paper. I am grateful for the contributions of my co-authors Prashant Bharadwaj, Manudeep Bhuller and Ola L. Vestad. Working with you has improved the papers in this thesis immensely, and has taught me so much about doing research.

To all my fellow PhD students at the Department of Economics. Thank you for the creating a great social environment, from lunches to lønningspils. A special thanks to Erik's Angles: Charlotte, Kjetil, Mads, and Oda. We have probably gone through all ranges of emotions together, and I am so grateful to have you in my cohort.

To my family and friends, who have supported me through this time, even when I have been quite unreasonable. And to Morten, for endless phone calls, and for being an excellent Corona colleague.

Oslo, August 2021

Mirjam Wentzel

Contents

Introduction

From the cradle to the grave, we are affected by our surroundings such as the welfare system we face and the events that occur around us. These factors have impacts on important parts of our well-being, including our health, working life and criminal behavior. They can also indirectly influence our family, friends, colleagues or classmates, and other people we encounter in life. Understanding how this affects our behavior is of great importance, as it can help us design policies that increase welfare in society. In this thesis, I study how different policies and life-events causally affect the behavior and well-being of individuals.

Abilities are to a large extent determined by the family we grow up in, through genetics, parental investment, and choice of childhood environments (Cunha and Heckman, 2009). These abilities have impacts on long-run outcomes such as income, education and health. However, external factors or interventions can have important implications for skill formation during childhood and adolescence. Some factors, such as high-quality child care, have positive impacts (see e.g., Heckman et al., 2013; Conti et al., 2016; Havnes and Mogstad, 2011), while others, such as childhood trauma, have negative impacts (see e.g., Bindler et al., 2020).

On the other side of working life is retirement. In many countries, public pension is the most important income source for a majority of elderly. In 2019, the average spending on old-age pension in the OECD countries was approximately 7.7 percent of GDP (OECD, 2021). As demographic changes put pressure on the public pension system, several countries have introduced pension reforms to increase labor supply of elderly workers. This can have important implications for both society at large, as well as the affected individuals.

In all three chapters, I use microeconometric techniques, together with detailed register data, to identify causal effects of life-events or policies on different outcomes including criminal behavior, educational attainment, labor supply, and household finances. All the same, each chapter focuses on a different set of research questions. The first two chapters are devoted to childhood events, while the third chapter attends to retirement behavior.

The title of the first chapter is 'Formal Child Care and Later-in-Life Delinquency'. Formal child care has proven to have positive effects on children's cognitive and non-cognitive skills, as well as long-term impacts on education and income. In this chapter, I analyze if enrollment in formal child care affects criminal charges in youth or early adulthood, using variation created by three different reforms in Norway. The first reform introduced universal child care for children aged 3-6. As the reform was implemented by municipalities, the increase in child care coverage varied geographically. I use a difference-in-differences design, comparing cohorts affected by the reform to cohorts that were not affected, between high- and low-expanding municipalities. In the second reform, a compulsory year of child care for six-year-olds was implemented in a school setting, affecting all children born after December 31, 1990. To capture the effects of the reform, I compare children born just after the reform to those born just before, with the same difference in the previous year[1], in a difference-in-regression-discontinuity approach. In the third reform, a cash-for-care benefit was given to parents not enrolling their children (aged 1-2) in formal child care. Different cohorts were affected differently by the reform, and this is used in a difference-in-differences analysis of the effects of the reform. The reforms affected child care enrollment in different ways; one increased access to formal child care, one introduced a compulsory year of child care, and one de-incentivized use of formal child care through a benefit for non-enrollment. The results from all three reforms suggest that formal child care can decrease delinquency, and that the effect is driven by men. Indicative evidence points towards larger impacts on children from families with lower socio-economic status. While the reforms differ in several aspects, including target age groups, incentive schemes, time of implementation, and likely complier groups, the effects all point in the same direction. Potential reasons could be the high quality of formal child care in the Norwegian context. The results from this chapter suggest that child care can be a potential tool in crime prevention policy.

The second chapter, 'Surviving a Mass Shooting', is joint work with Prashant Bharadwaj, Manudeep Bhuller and Katrine V. Løken. Experiencing childhood trauma can have large consequences for long-term outcomes, e.g., in the labor market. In this paper, we study

[1]Too capture age-at-school-start effects, as a consequence of the month of birth.

the impacts of a mass shooting on July 22, 2011 on the island of Utøya, Norway. We focus on school-aged survivors, as well as their family members and school peers. As we have access to administrative data on all registered victims in Norway, we are able to identify all survivors of the shooting. We match the survivors to same-aged control individuals, using coarsened exact matching on a number of observables including gender and parental characteristics. The matched children then form the control group for the affected survivors. Using a difference-in-differences approach, we compare outcomes of survivors before and after the event, to the equivalent for the matched control group. We find that survivors had a substantially lower middle school GPA and on-time completion of high school, as well as increases in health utilization and psychological diagnoses. In the medium-term, we find negative impacts on education and labor market outcomes. The administrative data allows us to link survivors to their family members and school peers, and we compare their outcomes to the equivalent for the children in the matched control group. We find that parents and siblings of survivors increased their GP visits, as well as their likelihood of getting a mental health diagnosis, while we only find limited impacts on school peers. Our results show that children that survive gun violence experience tremendous costs, despite a well-functioning social welfare system.

The third chapter, 'Income, Consumption and Savings around Retirement', is joint work with Ola L. Vestad. In recent years, many countries have implemented pension reforms to increase the work incentives for older individuals. While the labor supply responses to the reforms are well-documented, less is known about other important behavioral responses. In this chapter, we study the evolution of household income, savings and consumption around retirement in the context of a major reform of the Norwegian pension system. The reform induced a large increase in work incentives for workers that were covered by the early retirement scheme AFP, and a lowering of the early retirement age for workers not covered by AFP. It also increased flexibility by disentangling the decision to claim pension from the decision to retire. The reform affected different cohorts of workers differently, mainly those who reached the early retirement age before and after it was implemented. We use a difference-in-differences approach comparing those above the early retirement age before and after the reform, to the equivalent for those below. The findings suggest that the reform had positive impacts on household income, for AFP workers both through labor market income and transfers (mainly pension), and for nonAFP workers only through transfers. The increase in income is reflected in increases both in savings and consumption. Furthermore, we conduct an event study analyses around retirement, indicating that Norwegian households can maintain consumption levels similar to the pre-retirement period. These findings can have important implications in public pension policy design.

While this thesis contributes to the evidence on several topics, many important questions remain. I would therefore like to highlight some potential future areas of research, that can shed light on central issues related to the themes in this thesis. In the first chapter, I find that formal child care can decrease delinquency, but I do not provide answers to which factors that drive these results. Potential mechanisms that would be interesting to explore are e.g., changes in cognitive and non-cognitive skills, education or income. Furthermore, I do not discuss which child care policies that would be most efficient, or compare the effect sizes to other potential interventions for crime prevention. This would of course be important knowledge, especially for policymakers. Finally, it would be essential to explore if the results transfer to other settings, e.g., where the quality of child care is lower.

In the second chapter, we study how one specific mass shooting incident affected survivors, their family, and their peers. This was, as most mass shootings, a unique event, and it is therefore difficult to draw conclusions that are applicable to other settings. Furthermore, our research focuses on a specific type of childhood trauma, and it would be of interest to put this in context of other types of trauma, and events that children experience.

Finally, in the third chapter we focus on how the pension system affects household finances around the early retirement age. In future research, it will be important to examine both how this affects behavior during the full working lives of individuals, as well as to observe if the patterns we discovered are applicable to older pensioners. Furthermore, looking into household dynamics related to these behavioral changes would be of great interest. Finally, as Norway is one of the wealthiest countries in the world, with relatively high income equality, it would be interesting to determine if the evidence from this chapter transfers to other settings, where e.g., credit constraints are more important.

References

Bindler, A., Ketel, N., and Hjalmarsson, R. (2020). *Costs of Victimization*, pages 1–31. Springer International Publishing, Cham.

Conti, G., Heckman, J. J., and Pinto, R. (2016). The Effects of Two Influential Early Childhood Interventions on Health and Healthy Behaviour. *The Economic Journal*, 126(596):F28–F65.

Cunha, F. and Heckman, J. (2009). Human Capital Formation in Childhood and Adolescence. *CESifo DICE Report*, 7:22–28.

Havnes, T. and Mogstad, M. (2011). No Child Left Behind: Subsidized Child Care and Children's Long-Run Outcomes. *American Economic Journal: Economic Policy*, 3(2):97–129.

Heckman, J., Pinto, R., and Savelyev, P. (2013). Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes. *The American Economic Review*, 103(6):2052–2086.

OECD (2021). Pension Spending (Indicator). doi: 10.1787/a041f4ef-en (Accessed on 16 July 2021).

Chapter 1

Formal Child Care and Later-in-Life Delinquency

Formal Child Care and Later-in-Life Delinquency

Mirjam Wentzel*

Abstract

Formal child care has been shown in a wide set of studies to affect long-term outcomes, such as education, income and health, but does it also have an impact on later-in-life delinquency? In this paper, I investigate how formal child care affects criminal charges in youth and early adulthood. I use three reforms connected to formal child care in Norway, together with rich Norwegian register data, identifying differential access or incentives to use formal child care. The first reform introduced formal child care to children aged 3-6, which induced variation in the expansion of child care across municipalities. Using a difference-in-differences approach, I compare affected and non-affected cohorts in low- and high-expanding municipalities. In the second reform, a compulsory year of child care for 6-year-olds was implemented. This affected children born after January 1, 1991, which I use in a difference-in-regression-discontinuity design. Finally, in the third reform, a cash-for-care benefit was introduced for parents of 1- and 2-year-olds not attending formal child care, and hence de-incentivizing child care use. In a difference-in-differences design, I exploit the variation in treatment across cohorts. The results from the three reforms indicate that formal child care decreases criminal charges, and that effects are driven by men. There is no clear pattern in which type of offenses are affected by formal child care.

Keywords: child care, early childhood, crime
JEL codes: J13, I21, I38, K42

Acknowledgments: This research was conducted as part of the Starting Grant project 'Criminality, Victimization and Social Interactions' (CIVICS 757279) funded by the European Research Council. I thank seminar participants at Stockholm University, the Norwegian School of Economics and the Scandinavian PhD Seminar Applied Economics for helpful comments. The paper substantially improved as a result of conversations with Katrine V. Løken, Aline Bütikofer, Matthew Lindquist and Manudeep Bhuller.

*Department of Economics and FAIR, Norwegian School of Economics, and Statistics Norway, E-mail: mirjam.wentzel@gmail.com

1 Introduction

The cost of crime to society is large, including costs of prevention (e.g., police), punishment (e.g., courts and prison) and rehabilitation (e.g., prison and rehabilitation programs). In the US, there were over 1.4 million people in prison in 2019 (U.S. Bureau of Justice Statistics, 2020) and the total yearly cost of crime has been estimated to 1.7 trillion USD (Anderson, 2012). There are also large costs for victims, including decreases in earnings and increases in benefit receipt (Bindler and Ketel, 2020).

Policies directed at prevention of crime have mainly focused on direct measures such as police presence, higher sentences or youth interventions. The evidence suggests that police presence (absence) can reduce (increase) criminal acts in the affected geographic areas (e.g., Cheng and Long, 2018; Mello, 2019; Vollaard and Hamed, 2012; Weisburd, 2021). A number of papers show that increased sentences do not only reduce crime by incapacitation, but also through deterrence (see e.g., Bell et al., 2014; Drago et al., 2009; Abrams, 2012; Kessler and Levitt, 1999). Interventions or events during the teenage years can also be effective in reducing crime in the short run, but there is little evidence on the long-term effects (see e.g., Heller et al., 2017; Eren et al., 2017; Anderson, 2014). There has been much less work on early interventions and crime. One reason is that this requires following cohorts over a long time period as the time between interventions for young children and the time they start committing crime (youth and early adulthood) spans at least 10 years.

A broad literature across different fields provides evidence that non-cognitive skills are strong determinants of criminal behavior (see e.g., Caspi et al., 1994; Agnew et al., 2002; Pratt and Cullen, 2000), and that these skills are determined to a large extent in early childhood (Heckman et al., 2006). Similarly, high-quality subsidized child care has proven to induce positive changes in non-cognitive skills, especially among children from disadvantaged families (Heckman et al., 2013; Felfe et al., 2015). There is also evidence that child care can affect long-term outcomes of children, such as education, income and health (e.g., Campbell et al., 2014; Conti et al., 2016; Garces et al., 2002; Havnes and Mogstad, 2011b, 2015; Heckman et al., 2010, 2013).

It has been proven difficult to identify the effects of attending formal child care on youth or adult delinquency. One reason is data availability. This is because it (ideally) requires a long and representative panel data set with information on individuals' child care enrollment, as well as their criminal behavior. Another problem are the threats to identification caused by endogenous enrollment in child care. Child care enrollment is correlated with other factors such as maternal employment, which may also impact later-in-life outcomes. My paper addresses both the data and identification issues, and provides new evidence on the impacts of formal child care on criminal behavior.

I use three reforms affecting the take-up of universal child care in Norway, together with rich individual level register data, to study the effects on delinquency. The reforms differ in multiple ways; they affect different-aged children, vary in how they affect incentives for take-up of child care, and have different compliers. By using different reforms, we get a more complete picture of the effects of child care on crime within one country where other institutional settings are similar across reforms.

As previous literature has shown that child care has positive effects on long-term outcomes, the expectation is that it will also reduce crime. The mechanisms are potentially both through improvements in cognitive or non-cognitive skills, but also indirectly through the effects on e.g., education or earnings, which we know are factors strongly correlated with criminality. The vast majority of offenders are men, and it is therefore reasonable to expect that male delinquency will drive the results. Since reforms differ by target age groups, compliers, counterfactual care options and time of implementation, effects could differ across reforms. These differences can occur both in effect sizes, but we might also expect different subgroups to respond differently, in terms of, for example, socio-economic background.

The first reform was introduced in 1975, and introduced universal and subsidized child care for children aged 3-6. It regulated the authorization and running of child care centers, including quality measures. The municipalities were given the responsibility for establishing and running the centers. I use the empirical strategy from Havnes and Mogstad (2011b) to estimate the effects of the reform. The idea is that the expansion of child care differed across municipalities, and this is used in a difference-in-differences setting, comparing cohorts in the same municipality, and similar cohorts across municipalities, affected differently by the reform. The results show that the reform decreased the likelihood of being charged as an adult. The decrease is driven by men, and by traffic offenses. Suggestive evidence indicates that children of fathers with low income are impacted to a higher degree.

In the second reform, a compulsory year-long program was introduced for six-year-olds, effectively lowering school starting age from 7 to 6. It was implemented in 1997. The compulsory program was set up as a child care program, but conducted in schools, with the goal of preparing children for school. Children born in 1991 were the first cohort affected by the reform. I use this in a difference-in-regression-discontinuity (DiRD) strategy, using the fact that children born before new year in 1990/91 were not affected, while those after were, and compare this to the equivalent for the previous year. This takes into account that children born before new year may differ from those born after due to seasonality effects of births (Black et al., 2011). I find that the reform decreased the number of charges, driven by men. The effect is driven by decreases in narcotics offenses. The results suggest that effects are larger for children of mothers with only high school education, and fathers with low income.

The third reform introduced a cash-for-care benefit to parents of 1- and 2-year-olds in the end of the 1990s. Parents were eligible for the benefit if their child was not enrolled in subsidized full-time child care, and the monthly payment was equivalent to a state subsidy for a place in formal child care. This implies that the benefit served as an incentive to substitute formal child care with parental care or informal care. Different cohorts were affected differently by the reform, and I use this in a difference-in-differences setting, comparing the outcomes in the years 2010-2018 of 18-year-olds, who were partly or fully treated for later years, and 23- to 27-year-olds, who were never treated. The results suggest that the cash-for-care benefit increased the likelihood of being charged and the number of charges, driven by men. Charges for all types of offenses increased, and effect sizes are larger for children of mothers with a high school diploma or less, and of fathers with low income.

To summarize, the results from all three reforms point in the same direction - enrollment in formal child care seems to decrease the risk of criminal behavior later in life. The effect is driven by men. There is no clear pattern in the type of crime over the three reforms, which could be due to differences in compliers, different ages of outcome measurement and different ages of the children affected by the reforms. Suggestive evidence indicates that effects are larger for children from lower socio-economic backgrounds.

The most prominent literature related to the effect of formal child care on youth and adult outcomes comes from four programs in the US: the Abecedarian, the Perry Preschool, Project CARE, and Head Start.

The Abecedarian, Perry and CARE projects were all three randomized control trials (RCTs), directed at children from disadvantaged families. While the Perry Preschool Program randomly assigned children to high-quality child care, the Abecedarian and CARE projects also included health care and nutritional components (Conti et al., 2016; Campbell et al., 2014). Project CARE also had one treatment group with only home visits (Campbell et al., 2014). All three were small scale projects with 111 (Abecedarian), 123 (Perry) and 66 (CARE) participants in total.

The reduction in youth and adult crime has been high-lighted as the possibly most important benefit of the Perry Preschool Project (Schweinhart et al., 2005; Heckman et al., 2010). The program also had a positive impact on education, income, health and healthy behavior (Schweinhart et al., 2005; Heckman et al., 2010; Conti et al., 2016). The results from Heckman et al. (2013) suggest that changes in personality skills, rather than cognitive skills, can explain (parts) of the adult treatment effects.

The results from the Abecedarian Project, suggest that there is no significant effect on crime at age 21 (Clarke and Campbell, 1998; Campbell et al., 2002). For Project CARE, there are no studies related to delinquency to my knowledge. Both the Abecedarian Project and Project CARE, had positive effects on educational attainment, as well as reducing marijuana use and an increase in the adoption of an active life style (Campbell et al., 2008). The Abecedarian Project also seems to improve health, with the treated having a lower prevalence of risk factors for cardiovascular and metabolic diseases in their mid-30s (Campbell et al., 2014).

Head Start is a larger scale program than those presented above. It provided child care/pre school of a lower quality than, for example, the Perry Preschool (Garces et al., 2002). The evidence suggests that participating in Head Start reduces the risk of being charged in the early 20s (Garces et al., 2002; Carneiro and Ginja, 2014). There are also positive effects on educational attainment, income, health and behavior (Garces et al., 2002; Thompson, 2018; Carneiro and Ginja, 2014).

To the best of my knowledge, there are two previous studies of the effect of large scale universal childcare on delinquency. Baker et al. (2019) study the long-term effects of introduction of universal child care in Quebec, Canada. They use a difference-in-differences estimation, comparing outcomes of pre- and post-reform cohorts in Quebec to the equivalent in other Canadian provinces. They find that the cohort-crime rate in Quebec increased as a result of the reform. Other studies have shown that this reform mainly increased child care use for children of highly educated mothers, and the child care was of lower quality (Haeck et al., 2015) and that the effects were positive for children from single-parent households (Kottelenberg and Lehrer, 2017). This is very different from the Norwegian setting where the quality of child care is good. Brutti and Montolio (2019) study the effect of a reform expanding formal child care access for 3-year-olds in Spain. They use the difference in roll-out rate across regions, in a difference-in-differences framework, together with cohort and region level crime. The findings suggest that a 1 percentage point increase in preschool access for 0-3-year-olds decreases the number of reported crime actions by 1.6 percent. They find larger effects on impulsive-crime categories. My main contribution relative to this paper is that I study multiple reforms that affected child care use in different ways and affected different age groups. Contrary to

Brutti and Montolio (2019), I use detailed individual level register data, with the possibility to look at heterogeneous effects with regards to e.g., gender and socio-economic background, as well as in the type of offense committed.

There are several papers on the effect of universal child care on other outcomes (Berlinski et al., 2009; Havnes and Mogstad, 2011b, 2015; Datta Gupta and Simonsen, 2010; Dumas and Lefranc, 2012; Cornelissen et al., 2018; Felfe et al., 2015; Felfe and Lalive, 2018; Magnuson et al., 2007; Black et al., 2014; Loeb et al., 2007). The main takeaways from these papers are that universal childcare on average has long-term effects on education and income and that effects are larger for children from low-socio economic backgrounds. The evidence on short- to medium-term outcomes, mainly related to school outcomes, suggests that effects are positive, but possibly fading. There are a few exceptions, finding no or even negative effect (Baker et al., 2008, 2019; Haeck et al., 2015; Carta and Rizzica, 2018; Fort et al., 2020), and this seems to be related to the programs being of lower quality and directed or used by high SES families.

Finally the paper is related to studies on effects of other types of early life conditions and education on crime. There is evidence that conditions in early life, such as exposure to lead (Aizer and Currie, 2018; Billings and Schnepel, 2018), neighborhood crime (Damm and Dustmann, 2014), and illegal labor markets (Sviatschi, 2018), as well as family background (Eriksson et al., 2016) and childhood maltreatment (Currie and Tekin, 2012), affects youth and adult delinquency. Studies also suggest that for children on the margin, being placed in foster care increases crime (Doyle, 2007; Doyle Jr., 2008), while a recent study finds an imprecise decrease in delinquency (Gross and Baron, 2021). There is a consensus in the literature that both the quality and quantity of education affects crime (Anderson, 2014; Bell et al., 2016; Bennett, 2018; Berthelon and Kruger, 2011; Deming, 2011; Fella and Gallipoli, 2014; Machin et al., 2011). The effect is driven both by incarceration and behavioral changes, and the effect is heterogeneous across individual observables and education levels.

The remainder of the paper proceeds as follows. In section 2, I discuss the expected effects of universal child care on delinquency. Section 3 provides a background of the criminal justice system in Norway. Section 4 describes the data. In section 5, I present the institutional setting, identification and results for the introduction reform. Sections 6 and 7 provide the equivalent for the reforms of compulsory child care year and cash-for-care. The final section includes a discussion and conclusion.

2 Expected Effects

Previous literature has provided evidence that formal child care affects cognitive and non-cognitive skills (e.g., Heckman et al., 2010, 2013; Berlinski et al., 2009). Cognitive skills consists of factors promoting e.g., information acquisition, processing and problem-solving. Non-cognitive skills can include personality traits, persistence, motivation and charm (Heckman et al., 2006). The effects of child care on cognitive skills is seemingly fading (e.g., Magnuson et al., 2007; Blanden et al., 2016) but the effects on non-cognitive skills has proven to be persistent (e.g., Heckman et al., 2013). The improvement in skills is driven by children from disadvantaged families (e.g., Felfe et al., 2015; Cornelissen et al., 2018), and there is even some evidence of negative effects for children from advantaged families (e.g., Fort et al., 2020). The mechanisms behind both the positive and negative effects could be e.g., interaction with trained staff, less one-to-one interaction with adults, and interaction with other or different children. The results from previous research also suggests that the quality of care and the alternative to formal child care is important in interpreting effects (e.g., Haeck et al., 2015). Both cognitive and non-cognitive skills can affect delinquency both directly and indirectly, through e.g., education, income and (mental) health.

Heckman et al. (2006) study the predictive power of cognitive and non-cognitive skills on labor market outcomes and social behavior. They find that while both types of skills are of importance to the risk of delinquency, non-cognitive skills hold a larger predictive power for males.[1] If a man in the lowest decile of cognitive ability is moved from the lowest to the highest decile of non-cognitive ability, the risk of incarceration decreases significantly, while doing the equivalent for cognitive ability only decreases the risk slightly. This suggest that the link between child care and delinquency operates through effects on non-cognitive skills.

Formal child care generally has a positive effect on educational attainment (e.g., Havnes and Mogstad, 2011b; Heckman et al., 2010; Dumas and Lefranc, 2012). For income, the evidence suggests that there is an equalizing effect, increasing the income in the lower part of the distribution, while the higher part of the distribution may suffer from a decrease in income (e.g., Havnes and Mogstad, 2011b, 2015). The literature suggests that both higher educational attainment and income equality leads to lower delinquency (e.g., Hjalmarsson and Lochner, 2012; Choe, 2008). There is also evidence that formal child care can improve mental health in the long-run (e.g., Breivik et al., 2019). While I have found little evidence of the causal impact of mental health on crime, the literature suggests

[1] Incarceration is not an empirically important phenomenon for females(Heckman et al., 2006).

that mental illness is more common among offenders than in the general population (e.g., Gottfried and Christopher, 2017). Child care also seems to decrease marijuana use (e.g., Campbell et al., 2008). Since this is an offense in Norway, and might also be connected to use of heavier substances, this might be a mechanism for the results as well.

Finally, since children in formal child care are more likely to come in contact with other/more adults, including trained staff, problems in the home might be easier to detect, and contacts with social services could increase as a result. Two studies from the US have shown that foster care can increase criminal behavior for marginal individuals (Doyle, 2007; Doyle Jr., 2008). A recent study by Gross and Baron (2021), finds an imprecise decrease in delinquency, and positive impacts of foster care on educational outcomes. Foster care might have different impacts on different populations, and is also not the only outcome of contact with social services. It is therefore difficult to conclude how this may affect adult delinquency.

Most of the factors discussed above suggest that formal child care should reduce crime, and that it is most likely driven by children from disadvantaged families. One should note that most of this evidence comes from interventions that were successful in the short run, i.e., with promoting children's cognitive or non-cognitive skills. If child care is of lower quality than the alternative, the effect might be reversed, especially for children with highly educated parents. However, these are children with low risk of committing crime at the outset.

3 Crime in Norway

In this section I provide an overview of what happens when a report of criminal offense is made to the Norwegian Police. I also shortly discuss the development of crime in Norway over time, as well as present descriptive statistics.

When there is suspicion that a crime has been committed, a report is generally filed with the Police authority. Reports can be filed by the public or by the police. One example could be that an individual is stopped by the police for driving over the speed limit. The police would then file the report.

An illustration of the process after an offense is reported is presented in figure 1. The first thing that happens after reporting is that the police opens an investigation. In 2019, approximately 288 000 police investigations were conducted (Statistics Norway, 2020a).

When the investigation is completed, a decision is made on whether to press charges or not. Cases can be dismissed because they are unresolved. These cases stand for approximately half of the investigations. The most common reasons for unresolved cases are missing information about the perpetrator or a lack of evidence. The police and prosecuting authority can also dismiss a case because no punishable offense was committed. These cases are not a part of the statistics over investigated offenses.

Even if the case is solved, there are other reasons why charges are not made. If the perpetrator is younger than 15 years, he or she is not criminally liable, and will therefore not be charged. This stands for approximately 3 percent of all investigated cases. In other cases, the perpetrator is exempt from charges or there are other reasons why charges might not be filed. This could, for example, be cases of minor offenses committed by individuals without previous criminal records. These cases stand for approximately 8 percent of the total number of investigations.

In almost 40 percent of the cases, the prosecution will decide to press charges. Over half of these cases will go to court proceedings. This includes all serious offenses, including those where a prison sentence is a possible outcome. It also includes all cases where the perpetrator does not plead guilty.

In the vast majority of cases where the charge does not lead to trial, a fine is issued. This could be the case for our example with speeding, e.g., if the driver pleads guilty to driving only slightly above the speed limit and does not have a criminal record. These cases are not a part of the statistics on criminal charges. A small number of cases go to mediation.

In 2019, almost 280 000 punishments were set, of which a vast majority (78 percent) were on-the-spot fines (Statistics Norway, 2020b). Over 8600 prison sentences were set in 2019. There are also other possible punishments such as community service. In our example with speeding, the punishment of the driver will be dependent on factors such as how far above the speed limit he or she was driving and his or her criminal record. Punishments could then range from prison to a revoked driver's license to a single on-the-spot fine, depending on these factors.

Figure 2a displays the development in the number of reported offenses per 1000 inhabitants in Norway between 2003 and 2019. There is a clear downward trend in the number of charges. In 2003, the number of reported offenses per 1000 inhabitants was 92.4, while the same number in 2019 was under 58.2. The main driver is a large decrease in the reported offenses of property crime (Statistics Norway, 2021).

In figure 2b, I present the age and gender profiles of charged individuals in Norway in 2018. From the figure, we can see that the age of the charged individuals is concentrated among youth and young adults, especially around ages 18-25. The share is then quite steadily decreasing with age. Over 80 percent of the charged individuals are men.

4 Data

In this paper, I link several Norwegian administrative data sources using unique personal identifiers. I am also able to link the individuals to their family members.

From Statistics Norway, I have access to population panel data with individual demographics, including gender, immigrant status, education and family members. Furthermore, I use tax registers for information on income and social insurance registers including information on payments. Through the municipality database from the Norwegian Centre for Research Data (NSD), I have gathered information on child care coverage[2] on the municipality level.

From the Norwegian Police Directorate, I have access to data on all criminal charges between 1992 and 2018. The age of criminal responsibility in Norway is 15. The combination of the available years, the age of criminal responsibility and the timing of the reforms, imply that different ages are studied for different reforms. This will be explained further in the identification section of each reform, and is important for interpretation of the results.

The main outcomes are an indicator for being charged and the number of charges during the studied period. Furthermore, I have access to the type of offenses connected to the charges. The standard in Norway is to divide the type of offense into ten groups; economic offenses, other offenses for profit, violence offenses, sexual offenses, narcotics offenses, damage to property, environment offenses, work environment offenses, traffic offenses, and other offenses. As some of these groups of crimes have few observations in the data, I have taken outset in this grouping to limit the type of offense into 5 groups: offenses for profit (includes economic offenses and other offenses for profit), violent or sexual offenses, narcotics offenses, traffic offenses and other offenses (includes damage to property, environment offenses, work environment offenses, and other offenses). I use an indicator for being charged with an offense in each of these groups as outcomes.

Finally I have information on the set punishments. These are grouped into five categories; prison, probation, community service, fines and other.

[2]Child care coverage is the share of children within a certain age span (e.g., 3-6) enrolled in child care.

5 Introduction of Universal Child Care

Institutional Setting

In 1975, the Kindergarten Act was passed by the Norwegian Parliament, introducing universal and subsidized child care for children aged 3-6. Through the act, the municipalities were made responsible for the provision of child care. Furthermore, the act regulated prices, group size, educational content, physical environment and staff skill composition.

Women's entry in to the labor market, caused an increase in the demand for out-of-home child care in the 1950s and 1960s. A survey conducted in 1968 provided evidence of an unmet demand for formal child care (Norwegian Ministry of Administration and Consumer Affairs, 1972). Only 14 percent of children of respondents were in formal child care, while 32 percent of respondents expressed a demand for it.

The aim of the reform was to increase the number of child care spots up to 100 000 by 1981. In figure 3, the average municipal coverage rate is presented. In 1975, the average coverage rate was 4.3 percent. In the years following the reform, the number of children in child care increased substantially. In 1979, the average coverage rate was 25.2 percent. Government funding of child care also increased, from USD 34.9 million in 1975 to 85.8 million in 1977 (Havnes and Mogstad, 2015).

Havnes and Mogstad (2011a) provide evidence that the reform did not lead to an increase in maternal labor supply. This suggests that the reform induced a move from informal to formal child care, which is something we should have in mind when interpreting the results in this paper.

Previous studies of the effects of the reform also suggests that it lead to an increase in the educational attainment and labor market participation (Havnes and Mogstad, 2011b). The reform had an equalizing effect on income (Havnes and Mogstad, 2015). Finally, Breivik et al. (2019) find that it had a positive effect on long-term health.

For a more detailed description of the institutional setting, see Havnes and Mogstad (2011b) or Breivik et al. (2019).

Identification

To analyze the effect of the 1975 reform, I use the empirical strategy from Havnes and Mogstad (2011b). The main idea is that the roll-out rate differed across municipalities in the years following the reform. I focus on the expansion period between 1976 to 1979, and divide municipalities into treatment (control) if they were above (below) the median increase in coverage rate during the selected years.

Figure 4 displays the average municipal child care coverage rate for children aged 3-6, in treatment and control municipalities in the years 1973-1985. In the years leading up to the reform, the coverage rate was low, and similar in treated and control municipalities. Between 1976 and 1979, the coverage rate grew substantially in the treated municipalities, up to almost 40 percent in 1979. The growth in the control municipalities was much lower in these years. They caught up slightly in the early 1980s, but remained at a much lower level than the treated municipalities.

The sample consists of individuals born in 1967-1976. Individuals born before 1969, henceforth referred to as the pre-reform cohorts, where never affected by the reform, since they had already started compulsory schooling in 1976. Individuals born between 1970 and 1972, were between 3 and 6 years old when the expansion started, and therefore were partly treated. I refer to these individuals as belonging to the phase-in cohorts. Finally, individuals born in 1973-1976, were 3 or younger when the expansion started and therefore fully treated. They are referred to as the post-reform cohorts.[3]

I implement a difference-in-differences (DiD) strategy, using the differential treatment of municipalities and cohorts, with the following specification:

$$Y_{ijt} = \beta_1 + \beta_2(Phase - in_t \times Treat_j) + \beta_3(Post_t \times Treat_j) + \beta_4 X_i + \theta_j + \gamma_t + \varepsilon_{ijt} \quad (1)$$

where Y_{ijt} is the outcome for individual i in municipality j born in year t. $Phase - i_t$ ($Post_t$) is an indicator for being born in the phase-in (post) cohorts. $Treat_j$ is in indicator for individuals living in a treated municipality in 1976. X_i is a set of individual demographics including gender, number of older siblings, and parental education, birth year and age at birth. I also include municipality (θ_j) and cohort (γ_t) fixed effects.

The parameters of interest are β_2 and β_3, measuring the DiD coefficients for phase-in and post-reform cohorts respectively. Coefficients should be interpreted as intention to treat (ITT), as we cannot measure which children attended child care and not. Standard errors are clustered on the municipality level. Outcomes are measured in 1992-2006. This implies that criminal behavior is measured when the oldest cohort (born in 1967) is between 25 and 39, and the youngest cohort (born in 1976) is between 16 and 30 years old.

[3] I exclude individuals that have moved between a treated and control municipality between 1976 and 1979 from the estimation sample. This is approximately 5 percent of the sample. This exclusion has no substantial impact on the estimates.

The main identifying assumption is that the expansion of child care does not coincide with other municipality level changes affecting pre-reform, phase-in and post-reform cohorts differentially. In figures A.1-A.3 in appendix A, I provide evidence that municipality child care expansion does not correlate with other confounding factors such as changes in family structure, unemployment, total expenditure, expenditure on primary school and health expenditure.

Results

Table 1 displays the intention-to-treat (ITT) effects of the child care coverage change induced by the 1975 reform, on the likelihood of being charged and the number of criminal charges between 1992 and 2006. The estimating equation is described in the previous section, and the coefficient DiD is equivalent to β_3 in the equation. In panel A, the estimates for the full sample are presented, while the estimates for men (women) are displayed in panel B (C). For phase-in cohorts, there were no significant effects for the full sample or for men (see table A.6 in appendix A). For women, there was an increase in the likelihood of being charged (0.5 pp) for the phase-in cohorts.

In the full sample, the reform decreased the likelihood of being charged by around 0.8 percentage points. In the pre-reform cohorts, approximately 18 percent of the sample had been charged between 1992 and 2006. While the coefficient is not small for an intention to treat estimate, it is only significant at the 10 percent level when controls are included. The number of charges did not decrease significantly.[4]

The decrease in the likelihood of being charged is driven by men. When we include controls, the share of men being charged decreases by 1.5 percentage points, significant on the 10 percent level. For women, the likelihood of being charged is not significantly changed by the reform. There is no significant change in the number of charges for either women, or men.

In table 2, I present the ITT effects of the reform on the likelihood of being charged with the different types of offenses described in section 5, i.e., offenses for profit, violent or sexual offenses, narcotics offenses, traffic offenses, and other offenses. In panel A, the results for the full sample are shown. The reform induced a decrease in the likelihood of being charged with traffic offenses (0.8 percentage points) and other offenses (0.5 percentage points), significant on the 10 percent level. The likelihood of being charged with offenses for profit, violent offenses or narcotics offenses does not seem to be affected

[4]While we do not have data on individual take-up, we could scale by the average difference in change in coverage between treatment and control municipalities. This is about 17.62 percentage points and if we scale the ITT by this, the effect size is 4.3 pp.

by the access to child care.

For men, the decrease in the share being charged with traffic offenses increases even more, by approximately 1.6. For women there is no significant effect on any type of offense.

In table 3, I present the ITT effects of the reform on the likelihood of being charged and the number of charges by the mother's education level. Including controls, the decrease in the likelihood of being charged for children with a mother with a lower education level than high school is 0.8 percentage points, while the equivalent from high school or higher is 0.6 percentage points. For the number of charges, children with low-educated mothers have an slight increase in the number of charges (0.03), and children of high-educated mothers a decrease in the number of charges (0.06). None of the results are significantly different from zero, suggesting no clear pattern in the effect by mothers' education level.

Table 4 displays the ITT estimates of the effects by fathers' income quartile. The highest point estimates are found in the first quartile, with a decrease in the likelihood of being charged by 1.4 percentage points, and a decrease in the number of charges by 0.2. There is no clear pattern between the other income quartiles. While none of the results by fathers' income level are statistically significant, the pattern is consistent with previous research, showing that children from low SES families are the drivers of positive effects of formal child care.

A number of robustness checks are presented in appendix A. These include estimates by treatment intensity and tertile treatment groups. For treatment intensity, estimates point in the same direction as the main results, but the only significant estimate is for the likelihood of being charged for men. The tertile treatment groups suggest that it is the municipalities in the highest tertile treatment that are driving the effects. Furthermore, I have estimated the effects on the likelihood of being charged with more than 2, 5 and 10 charges, and find a significant effect on the likelihood of being charged at least 2 times, but no significant effect for 5 and 10 charges. I have also studied the effect on charges at age 25-30, and find similar, but not significant, results as in the main analysis. Finally, I have estimated the effects on the likelihood of being sentenced to different punishments, and find a decrease in the likelihood of getting a fine for men, but no other statistically significant effects. This is in line with traffic offenses being the driver of the decrease in delinquency. All in all, while I am lacking precision, the point estimates from the robustness checks support the results from the main analysis.

To summarize, the reform decreased the likelihood of being charged, driven by men. It did not affect the average number of charges significantly. The share charged with traffic offenses and other offenses decreased, but other type of offenses are not significantly affected by the reform. Suggestive evidence indicates that children of fathers with low income are affected to a larger extent.

6 Compulsory Child Care for Six-Year-Olds

Institutional Setting

In 1997, a reform that lowered the school age from 7 to 6 years was implemented in Norway. As a result, individuals born in 1991 or later also got an additional year of compulsory schooling.

Before the reform, 6-year-olds (and younger children) had access to high quality subsidized child care. 89 percent of non-immigrant 6-year-olds were enrolled in formal child care in 1996 (Drange et al., 2016).

The background for the reform was that children enrolled in formal child care could take part in school preparation, which was not available for others. The compulsory program for 6-year-olds, was to be child care like, preparing children for school. There was a social gradient in child care participation prior to the reform (Drange et al., 2016), and the hope was that the compulsory program would even out differences in learning outcomes.

The government proposed the reform in a White Paper in 1993 (Norwegian Ministry of Education, 1993) and the bill was passed in 1994 (Norwegian Ministry of Education, 1994). It was implemented in 1997, and the first cohort starting school at age 6 were those born in 1991.

Drange et al. (2016) uses a difference-in-differences approach, comparing the difference in schooling outcomes for children not attending formal child care at age 5 born before and after 1991, to the equivalent for children attending formal child care at age 5. They find no effect of the reform on schooling outcomes.

For a more detailed description of the institutional setting, see Drange et al. (2016).

Identification

In Norway, the calendar year serves as the cut-off for school cohorts. This implies that if a child is born on December 31 in one year, or January 1 in the following year, this will generally determine what school cohort the child belongs to. While it is possible to enroll one year ahead of time, or postpone the enrollment one year, this requires that parents formally apply for an exception, and approval from health and school specialists (see e.g., Norwegian Directorate for Education and Training, 2017). While there is no register data on children's age of enrollment in Norway, studies have found that only around 0.5 to 1.2 percent postpone enrollment, and that this is somewhat more frequent amongst boys (1.5 percent) (Cools et al., 2017; Gabrielsen and Lundetræ, 2017).

Normally, this rule implies that those born in December start when they are (almost) one year younger than those born in January. Black et al. (2011) analyze the effect of this difference in starting age, and find that starting at a younger age decreases IQ test scores and earnings until age 30. For the cut-off between children born in 1996 and 1997, the reform implied that children born in December 1996 started school at the same age as those born in January 1997, but one grade above.

With this as a background, I use a difference-in-regression-discontinuity (DiRD) design to analyze the effect of the compulsory program for 6-year-olds. The main idea is to compare the outcomes of children born around the cut-off December 31, 1990/January 1, 1991. The results from Black et al. (2011) suggest that using a simple RD to identify the effects of the reform will be problematic, as we will expect to find discontinuities around the cut-off regardless of the reform due to the age at school start effect.[5] I therefore take the difference between the discontinuity in the treated year and compare it to the same discontinuity in the year prior to the reform. It is estimated using the following baseline specification:

$$Y_{it} = \beta_1 + \beta_2(Z_{it} - c_t) + \beta_3 E_i + \beta_4 T_t + \beta_5(E_i \times T_t) + \beta_6(Z_i t - c_t)E_i + X_i + \epsilon_i \tag{2}$$

where Y_{it} is the outcome for individual i born around cut-off t. Z_{it} is the date of birth and c_t is the cut-off date. E_i is an indicator equal to one if the individual is born after the cut-off and T_t is an indicator equal to one if the individual is born around the treatment cut-off. X_i is a set of control variables including gender, immigrant background, birth municipality, parents' education, year of birth and age at birth. The interaction term $E_i \times T_t$ is then equal to one if the individual is born around and after the reform cut-off, that is, after December 31, 1990, and therefore β_5 is the coefficient of interest.

This approach differs from previous work by Drange et al. (2016), who use a differences-in-differences design, comparing children not enrolled in child care at age 5 to those enrolled, in cohorts starting school before and after the implementation of the reform. The reason I could use this new identification strategy is that the outcome variable is a monthly panel while their school outcomes are one time measures. Their measure captures the difference of attending the compulsory program for those who would likely not have been enrolled in

[5]This strategy is also consistent with the results from (Landersø et al., 2017), suggesting that the age at school start effect is also present for patterns of criminal behavior.

formal child care otherwise, to those who would have. The measure in this paper instead captures the difference between the compulsory program and all alternatives for care of 6-year-olds.

In the baseline estimations, I include linear trends that are separate for each side of the cut-offs, but equal for individuals born in different years on the same side of the cut-offs[6]. I use 180 day bandwidth and triangular weights, and cluster standard errors on the municipality level. The outcomes are measured in 2008-2018. This implies that the criminal behavior is measured at ages 19-29 for the oldest individuals in the sample (born in 1989), and at ages 17-27 for the youngest individuals in the sample (born in 1991).

Results

In figure 5, the regression discontinuity figures for the share being charged and the average number of charges are presented. In each figure, I show the binned 7 day average (the scatter plot), and the linear trend with 95 percent confidence interval for each side of the cut-offs. The bandwidth is 180 days. In panels (A) and (C), the RD plots for the treatment cut-off (January 1, 1991) are presented, and in panels (B) and (D) the equivalent is presented for the control cut-off (January 1, 1990).

In panels (A) and (C), we can see that the share being charged and the average number of charges drops around the cut-off, giving us a first indication that the reform affected criminal behavior. But as discussed in the identification section above, we need to take into account the potential age of school start effect that occurs every year. Since there is no clear drop around the cut-off for the control group (see panels (B) and (D)), it is likely that the reform is causing the drop in delinquency.

Table 5 displays the DiRD estimates of the effect of the reform on the likelihood of being charged and the number of charges. In panel A, I present the estimates for the full sample, while panel B and panel C show the results for men and women respectively.

In the full sample, the estimate for the likelihood of being charged is negative (0.5 percentage points), but not significant, while the number of charges seem to decrease by approximately 0.1 as a result of the reform, from an average of 0.9. For men, the likelihood of being charged decreases, but not significantly, while the number of charges decreases by 0.3 and is significant on the 5 percent level when controls are included. For women, the effect on both the likelihood of being charged is small and positive, but not significant.

[6]Another option would be using local linear regression. However, results from local linear regression are imprecise and sensitive to choice of bandwidth, see figure B.3 in appendix B.

In table 6, I present the estimates of the compulsory year of pre-school on the likelihood of being charged with different types of offenses. In panel A, we can see that in the full sample the reform seemingly decreased the likelihood of being charged with narcotics offenses and other offenses, but did not seem to affect offenses for profit, violent or sexual offenses, or traffic offenses significantly. For men (see panel B), the reform significantly decreased the likelihood of being charged with all types offenses, except for traffic offenses, that did not significantly change. For women (see panel C), the likelihood of being charged with other offenses decreased slightly. This effect is significant on the 10 percent level. For the other types of offenses, the effect is small, positive but not significant for women.

Table 7 displays the estimates by mothers' education divided in to three categories; less than high school (low), high school and some higher education (university or college). There are no statistically significant results on the likelihood of being charged. From the estimates on the number of charges, while children of mothers with low or higher education have no significant effect on the number of charges, children of high school (only) educated mothers displays a decrease of approximately 0.2, suggesting that they are driving the results in our main analysis.

In table 8, the results are divided by fathers' income quartile. Again, we find no significant effect on the likelihood of being charged. The largest decrease in the number of charges is in the 1st quartile, by approximately 0.3, but this decrease is only marginally significant. For quartiles 2 and 3, there are no significant effects, but for the last quartile we see a significant decrease in the number of charges by approximately 0.1.

In appendix B, I provide a number of robustness checks. First, I study the effects on indicators for more than 2, 5 and 10 charges. In total, as well as for men, I find significant negative effects on the likelihood of being charged at least two times, and negative but non-significant effects for five and 10 charges. I also study the effects on different types of punishments, and find significant decreases in the likelihood of getting community service for the full sample and for men. Furthermore, I study the effects on charges at age 18-25, and find significant decreases in the likelihood of being charged as well as the number of charges, for both the full sample and for men. I run a number of specifications testing different bandwidths and trends. While the precision and point estimates vary somewhat, the results from the specification support the results from the main analysis.

To summarize, the compulsory child care year for 6-year-olds significantly decreases the number of charges, driven by men. The effects seem to come mainly from a decrease in narcotics offenses and other offenses. Effect sizes are larger for children of mothers with high school only, and children of fathers with low income.

7 Cash-for-Care

Institutional Setting

In August 1998, the cash-for-care benefit was introduced for 1-year-olds, and in January 1999 for 2-year-olds. It gave parents of 1- and 2-year-olds a tax-free benefit, given that they did not use subsidized full-time child care (more than 32 hours per week). The benefit was equivalent to a state subsidy for a place in formal child care when it was introduced (Rønsen, 2009). In the first years after the introduction, around 80 percent of parents of 1- and 2-year-olds received the benefit (Statistics Norway, 2019).

In 2012, the benefit was removed for 2-year-olds (Statistics Norway, 2019). As per 2020, the full monthly benefit is 7500 NOK ($\approx$ 835 USD) per child (Norwegian Labour and Welfare Administration, 2020). It is possible to receive a share of the benefit if the child is in part-time subsidized child care.

The cash-for care benefit gives incentive to substitute formal child care for either parental care or informal child care. For never-takers of formal child care, the benefit only serves as an increase in income. In addition, since supply of child care slots for 1-2-year-olds is still limited in this period, it serves as an extra parental leave benefit for those who need to wait for formal child care.

Rønsen (2001) provided evidence that the reform lead to increases in both parental care, and informal care. The evidence from previous studies suggests that maternal labor supply decreased both in the short- and long-run (Rønsen, 2001, 2009; Drange and Rege, 2013). The results from Drange and Rege (2013) indicate that mothers with low education or low pre-reform earnings are driving the effects, and that effects fade out when the child is around age 6. Previous research also suggests that older siblings' 10th grade GPA increased as a result of the reform, driven by a decrease in maternal labor supply (Bettinger et al., 2014).

For a more detailed decription of the institutional setting, see Rønsen (2009) and Bettinger et al. (2014).

Identification

Different cohorts were affected differently by the reform. Parents of individuals born in 1995 or earlier never had access to the cash-for-care benefit, and were therefore not treated. Parents of individuals born in 1996-1997, could have as much as 24 months eligibility or as little as 1 month, and I therefore refer to them as partly treated. Parents of individuals born in 1998 or later were eligible for 24 months of cash-for-care benefit, and are referred to as fully treated.

In 2010-2018, there are 18-year-olds belonging to all three treatment groups, while individuals 23 years or older in the same time period were only born in the pre-reform cohorts. I use this in a difference-in-differences (DiD) framework, which is illustrated in figure 6. This figure displays the share of 18-year-olds and 23-year-olds being charged in each year between 2010-2018. In years 2010-2013, neither the 18-year-olds, nor the 23-year-olds belong to cohorts affected by the cash-for-care reform. In 2014 and 2015, the 18-year-olds were in the partly treated cohorts, while the 23-year-olds were still in the non-treated cohorts. In the years 2016-2018, the 18-year-olds were in the fully treated cohorts, while the 23-year-olds were still not treated.

The estimation sample consists of 18-year-olds (treated), and individuals ages 23-27 (control group). The choice of 18-year-olds as the treatment group is based on two factors. The first factor is the number of post-reform years observed given the data limitations.[7] For 18-year-olds, three post-reform years are observed, while for e.g., 19-year-olds we only observe two. Second, 18-year-olds, are amongst the age groups with the highest prevalence of delinquency (see figure 2b in section 3). As outcomes are measured yearly in my estimation, using one of the peak ages of criminal behavior is likely to more accurately reflect the overall likelihood of being charged in early adulthood.

Two aspects have been the focus in the choice of control group. First, it is likely that age groups that are closer together have more similar criminal patterns, and it is therefore preferable to include ages as close to 18 as possible. Second, there may be birth cohort specific shocks that could affect the outcomes.[8] To reduce this risk, I therefore choose to include five ages/cohorts per year in the control group.

To formalize, I use the following difference-in-difference specification as my baseline estimation:

$$Y_{it} = \beta_1 + \beta_2 Treat_i + \beta_3(Partly_t \times Treat_i) + \beta_4(Fully_t \times Treat_i) + \beta_5 X_i + \theta_t + \epsilon_{it} \quad (3)$$

where Y_{it} is the outcome for individual i in year t. $Treat_i$ is equal to one if the observation is for an 18-year-old. $Partly$ ($Fully$) is an indicator for years 2014-2015 (2016-2018). X_i is a set of control variables including gender, immigrant background, birth municipality, parents' education, year of birth and age at birth. θ_t are year fixed effects.

[7] The last observed year is 2018.

[8] An example: The compulsory year of child care discussed in section 6 affected cohorts born from 1991. This implies that for 23-year-olds, outcomes may be affected from 2014, for 24-year-olds from 2015 etc. By including multiple cohorts yearly, I hope to reduce the risk of this bias.

The coefficients of interest are β_3 and β_4, which is the mean difference in outcomes between 18- and 23- to 27-year-olds in 2014-2015 and 2016-2018, minus the equivalent for years 2010-2013. Since take-up of the benefit was not compulsory, estimates presented will be intention-to-treat (ITT).[9]

For this analysis to be credible, the common trend assumption needs to hold, implying the the trends in criminal behavior for 18- and 23- to 27-year-olds would have been similar if the cash-for-care reform had not been implemented. Since the outcomes are measured yearly, the likelihood of being charged will be lower then in the analysis of the previous two reforms, where we measure the outcomes over multiple years and ages. To assess the common trend assumption, I provide yearly DiD estimates in figure 7, which are further discussed in the following section.

My strategy is somewhat similar to Drange and Rege (2013), who compares outcomes for mothers having same-aged children that were or were not affected by the reform, to mothers of older children who were never affected. My analysis differs from the other previous research of this reform, where Rønsen (2001, 2009) compares the outcomes of mothers giving birth before the reform to mothers giving birth after as a main approach. Bettinger et al. (2014) compares the outcomes of children with siblings born after the reform, to those with siblings born before.

Results

In table 9, the difference-in-differences (DiD) estimates for the effect of cash-for-care on the likelihood of being charged and the number of criminal charges are presented, for the post-reform cohorts. The equivalent for the phase-in cohorts is presented in table C.3 in appendix C.

Panel A displays the results for the full sample. The likelihood of being charged in a given year increases by 0.1 percentage points without controls, and 0.4 percentage points with controls, as a result of the reform, from a pre-reform level of 4 percent. The results with controls, indicate an increase of approximately 10.5 percent from the pre-reform level of 18-year-olds. The number of charges is estimated to increase by 0.01 (0.02) without (with) controls, from a pre-reform mean of 0.07. The results are significant on the 1 percent level when including controls.

For men (panel B), the increases in the likelihood of being charged and the number of charges is even larger, with an increase of 1 percentage points in the likelihood of being

[9] For cohorts born in 1998-2000, the average number of months with take-up was 18.2, which corresponds to 76 percent of the total possible months with take-up (24 months).

charged and of 0.04 number of charges, when controls are included. For women, we see a small decrease in the likelihood of being charged by approximately 0.02 percentage points. Results for phase-in cohorts show an increase in the the number of charges in the full sample, but no significant effect on the likelihood of being charged. The increase is smaller than for the post cohorts.

In figure 7, I present the yearly DiD estimates for the likelihood of being charged and the number and charges for the full sample (panels A and B), men (panels C and D) and women (panels E and F). First, there is little evidence of pre-trends. In the full sample, the outcomes for the non-treated years are significantly different from the base year 2013.[10] The overall pattern for the full sample and for men, is a small increase for phase-in/partly treated years as well as for the first fully treated year, and then an increasing effect for 2017 and 2018. While this may be a result of increasing compliance, we need to have some concern that other factors might give rise to this pattern, and should therefore interpret results with some caution. For women, there are no clear effects to read out from these graphs.

In table 10, the DiD estimates for the effect of cash-for-care on the likelihood of being charged with different types of offenses are presented. In panel A, I show the results for the full sample, while the estimates for men (women) are presented in panel B (C).

In the full sample, we see an increase in offenses for profit, violent or sexual offenses, narcotics offenses, and traffic offenses, as a possible result of the reform. There is no significant effect on other offenses. For men, the results are similar, but also with a significant increase in other offenses. For women, we see marginally significant increases in offenses for profit and narcotics, as well as traffic offenses, while we see decreases in violent or sexual offenses and other offenses. These results for women should be interpreted with special caution, as very few women are charged with the different types of offenses.

Table 11 displays the effects of cash-for-care by mothers' education divided in to three groups; lower than high school (low), high school and higher education (some university or college). Children of mothers' with low or high school education only display a significant increase in the likelihood of being charged by 0.3 and 0.5 percentage points (including controls), as well as an increase in the number of charges of approximately 0.02. For children of mothers with higher education, there are no significant effects on the likelihood of being charged, and an increase in the number of charges by approximately 0.007.

[10]For the number of charges for men, 2010 is marginally significantly different from 2013, and for women the coefficient for likelihood of being charged in 2011 is significanttly different. While this of course is unwanted, there are no clear patterns in the pre-trends overall.

In table 12, I display the effects by fathers' income quartile. There is a significant increase in the likelihood of being charged for children of fathers' in income quartiles 1, 2 and 4 (when we include controls), and no significant effect for quartile 3. The effect on the number of charges is positive in all quartiles. Point estimates are larger in quartiles 1 and 2, which is consistent with the previous literature that children from lower SES backgrounds have larger benefits from attending formal child care.

In appendix C, I provide a couple of additional robustness checks. First, the effects of the reform on the likelihood of being charged at least 2, 5 and 10 times are presented, suggesting an increase in all levels for both the full sample and men, and an increase in the likelihood for more then 5 and 10 charges for women. Second, I provide estimates on the likelihood of being sentenced to different punishments, and find increases in prison sentences, probation and fines.

To summarize, the introduction of cash-for-care benefit increased delinquent behavior. While men are driving these results, there are some indications that the likelihood of being charged decreased for women. The increase in delinquency is visible in all types of offenses, and effects are larger for children from low SES backgrounds.

8 Conclusion

In this paper, I investigate whether formal child care can affect long-term criminal behavior. To overcome identification and data challenges, I use variation in formal child care use created by three different reforms, together with detailed register data including all criminal charges in Norway between 1992 and 2018. I find that child care enrollment can decrease youth or adult delinquency, an effect driven by men. There is no clear pattern in what type of offenses that are driving the results. Indicative results suggest that children from lower socio-economic backgrounds are affected to a larger extent, which is in line with the previous literature.

While the three reforms all affected enrollment in formal child care, they also had substantial differences. First, the target age groups vary from 1 to 6 years old. Second, the reforms affected child care enrollment differently; one through increased access, one by compulsory enrollment and one by de-incentivizing use of formal child care. They were therefore likely to have different complier groups. Finally, they were implemented at different points in time. Despite these differences, the effects of the reforms point in the same direction. Common factors such as the surrounding institutional setting, as well as the high quality of formal child care in Norway, are likely important contributors to the outcomes.

These findings can have important implications for policies focusing on crime prevention. My results suggest that early childhood interventions such as child care not only increases education and equalizes income, but can also be a tool in decreasing youth or adult delinquency. While policies aimed at young children might not have an immediate affect on crime, they can have long-term benefits which are important to account for in policy-makers' decisions.

My paper is the first to use multiple reforms connected to child care use together with the possibility to follow individuals over an extended period of time. This contributes to the previous literature by showing how different early childhood interventions can impact criminal behavior in the long-run. However, several questions remain. What mechanisms are driving the results? What child care policies are most efficient for prevention? Do the results from this paper transfer to other settings, e.g., where quality of child care is lower? These are interesting questions for future research.

References

Abrams, D. S. (2012). Estimating the Deterrent Effect of Incarceration Using Sentencing Enhancements. *American Economic Journal: Applied Economics*, 4(4):32–56.

Agnew, R., Brezina, T., Wright, J., and Cullen, F. (2002). Strain, Personality Traits, and Delinquency: Extending General Strain Theory. *Criminology*, 40:43–72.

Aizer, A. and Currie, J. (2018). Lead and Juvenile Delinquency: New Evidence from Linked Birth, School and Juvenile Detention Records. *The Review of Economics and Statistics*, 101.

Anderson, D. A. (2012). *The Cost of Crime, Foundation and Trends in Microeconomics*, volume 7. Citeseer. pp 209-265.

Anderson, D. M. (2014). In School and Out of Trouble? The Minimum Dropout Age and Juvenile Crime. *The Review of Economics and Statistics*, 96(2):318–331.

Baker, M., Gruber, J., and Milligan, K. (2008). Universal Child Care, Maternal Labor Supply, and Family Well-Being. *Journal of Political Economy*, 116(4):709–745.

Baker, M., Gruber, J., and Milligan, K. (2019). The Long-Run Impacts of a Universal Child Care Program. *American Economic Journal: Economic Policy*, 11(3):1–26.

Bell, B., Costa, R., and Machin, S. (2016). Crime, Compulsory Schooling Laws and Education. *Economics of Education Review*, 54:214–226.

Bell, B., Jaitman, L., and Machin, S. (2014). Crime Deterrence: Evidence from the London 2011 Riots. *The Economic Journal*, 124(576):480–506.

Bennett, P. (2018). The Heterogeneous Effects of Education on Crime: Evidence from Danish Administrative Twin Data. *Labour Economics*, 52:160–177.

Berlinski, S., Galiani, S., and Gertler, P. (2009). The Effect of Pre-Primary Education on Primary School Performance. *Journal of Public Economics*, 93(1):219–234.

Berthelon, M. E. and Kruger, D. I. (2011). Risky Behavior among Youth: Incapacitation Effects of School on Adolescent Motherhood and Crime in Chile. *Journal of Public Economics*, 95(1):41–53.

Bettinger, E., Hægeland, T., and Rege, M. (2014). Home with Mom: The Effects of Stay-at-Home Parents on Children's Long-Run Educational Outcomes. *Journal of Labor Economics*, 32(3):443–467.

Billings, S. B. and Schnepel, K. T. (2018). Life after Lead: Effects of Early Interventions for Children Exposed to Lead. *American Economic Journal: Applied Economics*, 10(3):315–44.

Bindler, A. and Ketel, N. (2020). Scaring or Scarring? Labour Market Effects of Criminal Victimisation. Technical report, ECONtribute Discussion Paper no. 030.

Black, S. E., Devereux, P. J., Løken, K. V., and Salvanes, K. G. (2014). Care or Cash? The Effect of Child Care Subsidies on Student Performance. *The Review of Economics and Statistics*, 96(5):824–837.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2011). Too Young to Leave the Nest? The Effects of School Starting Age. *The Review of Economics and Statistics*, 93(2):455–467.

Blanden, J., Del Bono, E., McNally, S., and Rabe, B. (2016). Universal Pre-school Education: The Case of Public Funding with Private Provision. *The Economic Journal*, 126(592):682–723.

Breivik, A.-L., Del Bono, E., and Riise, J. (2019). Effects of Universal Childcare on Long-Run Health. Working Paper.

Brutti, Z. and Montolio, D. (2019). Preventing Criminal Minds: Early Education Access and Adult Offending Behavior. IEB Working Paper.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica*, 82(6):2295–2326.

Campbell, F., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E., and Pan, Y. (2014). Early Childhood Investments Substantially Boost Adult Health. *Science*, 343(6178):1478–1485.

Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J., and Miller-Johnson, S. (2002). Early Childhood Education: Young Adult Outcomes From the Abecedarian Project. *Applied Developmental Science*, 6(1):42–57.

Campbell, F. A., Wasik, B. H., Pungello, E., Burchinal, M., Barbarin, O., Kainz, K., Sparling, J. J., and Ramey, C. T. (2008). Young adult outcomes of the Abecedarian and CARE early childhood educational interventions. *Early Childhood Research Quarterly*, 23(4):452–466.

Carneiro, P. and Ginja, R. (2014). Long-Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start. *American Economic Journal: Economic Policy*, 6(4):135–173.

Carta, F. and Rizzica, L. (2018). Early Kindergarten, Maternal Labor Supply and Children's Outcomes: Evidence from Italy. *Journal of Public Economics*, 158:79–102.

Caspi, A., Mofitt, T. E., Silva, P. A., Stouthamer-Loeber, M., Krueger, R. F., and Schmutte, P. S. (1994). Are Some People Crime-Prone? Replications of the Personality-Crime Relationship Across Countries, Genders, Races, and Methods. *Criminology*, 32(2):163–196.

Cheng, C. and Long, W. (2018). Improving police services: Evidence from the French Quarter Task Force. *Journal of Public Economics*, 164(C):1–18.

Choe, J. (2008). Income inequality and crime in the United States. *Economics Letters*, 101(1):31–33.

Clarke, S. H. and Campbell, F. A. (1998). Can Intervention Early Prevent Crime Later? The Abecedarian Project Compared with Other Programs. *Early Childhood Research Quarterly*, 13(2):319–343.

Conti, G., Heckman, J. J., and Pinto, R. (2016). The Effects of Two Influential Early Childhood Interventions on Health and Healthy Behaviour. *The Economic Journal*, 126(596):F28–F65.

Cools, S., Schøne, P., and Strøm, M. (2017). Forskyvninger i skolestart: Hvilken rolle spiller kjønn og sosial bakgrunn? *Søkelys pa arbeidslivet*, 34(4).

Cornelissen, T., Dustmann, C., Raute, A., and Schönberg, U. (2018). Who Benefits from Universal Child Care? Estimating Marginal Returns to Early Child Care Attendance. *Journal of Political Economy*, 126(6):2356–2409.

Currie, J. and Tekin, E. (2012). Understanding the Cycle: Childhood Maltreatment and Future Crime. *The Journal of Human Resources*, 47(2):509–549.

Damm, A. P. and Dustmann, C. (2014). Does Growing Up in a High Crime Neighborhood Affect Youth Criminal Behavior? *The American Economic Review*, 104(6):1806–1832.

Datta Gupta, N. and Simonsen, M. (2010). Non-Cognitive Child Outcomes and Universal High Quality Child Care. *Journal of Public Economics*, 94(1):30–43.

Deming, D. J. (2011). Better Schools, Less Crime? *The Quarterly Journal of Economics*, 126(4):2063–2115.

Doyle, J. J. (2007). Child Protection and Child Outcomes: Measuring the Effects of Foster Care. *The American Economic Review*, 97(5):1583–1610.

Doyle Jr., J. (2008). Child Protection and Adult Crime: Using Investigator Assignment to Estimate Causal Effects of Foster Care. *Journal of Political Economy*, 116(4):746–770.

Drago, F., Galbiati, R., and Vertova, P. (2009). The Deterrent Effects of Prison: Evidence from a Natural Experiment. *Journal of Political Economy*, 117(2):257–280.

Drange, N., Havnes, T., and Sandsør, A. M. (2016). Kindergarten for All: Long Run Effects of a Universal Intervention. *Economics of Education Review*, 53:164–181.

Drange, N. and Rege, M. (2013). Trapped at Home: The Effect of Mothers' Temporary Labor Market Exits on Their Subsequent Work Career. *Labour Economics*, 24:125–136.

Dumas, C. and Lefranc, A. (2012). Early Schooling and Later Outcomes: Evidence from Pre-School Extension in France. In *From Parents to Children : The Intergenerational Transmission of Advantage*. Ed. by Ermisch, J., Jäntti, M., and Smeeding, T. Russell Sage Foundation. New York.

Eren, O., Depew, B., and Barnes, S. (2017). Test-Based Promotion Policies, Dropping Out, and Juvenile Crime. *Journal of Public Economics*, 153:9–31.

Eriksson, K. H., Hjalmarsson, R., Lindquist, M. J., and Sandberg, A. (2016). The Importance of Family Background and Neighborhood Effects as Determinants of Crime. *Journal of Population Economics*, 29(1):219–262.

Felfe, C. and Lalive, R. (2018). Does Early Child Care Affect Children's Development? *Journal of Public Economics*, 159:33–53.

Felfe, C., Nollenberger, N., and Rodríguez-planas, N. (2015). Can't Buy Mommy's Love? Universal Childcare and Children's Long-Term Cognitive Development. *Journal of Population Economics*, 28(2):393–422.

Fella, G. and Gallipoli, G. (2014). Education and Crime over the Life Cycle. *The Review of Economic Studies*, 81(4 (289)):1484–1517.

Fort, M., Ichino, A., and Zanella, G. (2020). Cognitive and Noncognitive Costs of Day Care at Age 0–2 for Children in Advantaged Families. *Journal of Political Economy*, 128(1):158–205.

Gabrielsen, E. and Lundetræ, K. (2017). 11 Indikerer de norske PIRLS-resultatene et behov for a justere retningslinjene for skolestartsalder? In Klar framgang! p.p. 204–221.

Garces, E., Thomas, D., and Currie, J. (2002). Longer-Term Effects of Head Start. *The American Economic Review*, 92(4):999–1012.

Gottfried, E. D. and Christopher, S. C. (2017). Mental Disorders Among Criminal Offenders: A Review of the Literature. *Journal of Correctional Health Care*, 23(3):336–346. PMID: 28715985.

Gross, M. and Baron, J. (2021). Temporary Stays and Persistent Gains: The Causal Effects of Foster Care. *American Economic Journal: Applied Economics*. Forthcoming.

Haeck, C., Lefebvre, P., and Merrigan, P. (2015). Canadian Evidence on Ten Years of Universal Preschool Policies: The Good and the Bad. *Labour Economics*, 36.

Havnes, T. and Mogstad, M. (2011a). Money for Nothing? Universal Child Care and Maternal Employment. *Journal of Public Economics*, 95(11):1455–1465. Special Issue: International Seminar for Public Economics on Normative Tax Theory.

Havnes, T. and Mogstad, M. (2011b). No Child Left Behind: Subsidized Child Care and Children's Long-Run Outcomes. *American Economic Journal: Economic Policy*, 3(2):97–129.

Havnes, T. and Mogstad, M. (2015). Is Universal Child Care Leveling the Playing Field? *Journal of Public Economics*, 127:100–114. The Nordic Model.

Heckman, J., Moon, S. H., Pinto, R., Savelyev, P. A., and Yavitz, A. (2010). The Rate of Return to the HighScope Perry Preschool Program. *Journal of Public Economics*, 94(1):114–128.

Heckman, J., Pinto, R., and Savelyev, P. (2013). Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes. *The American Economic Review*, 103(6):2052–2086.

Heckman, J., Stixrud, J., and Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3):411–482.

Heller, S. B., Shah, A. K., Guryan, J., Ludwig, J., Mullainathan, S., and Pollack, H. A. (2017). Think Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago. *Quarterly Journal of Economics*, 132(1):1–54.

Hjalmarsson, R. and Lochner, L. (2012). The Impact of Education on Crime: International Evidence. *CESifo DICE Report*, 10(2):49–55.

Kessler, D. and Levitt, S. (1999). Using Sentence Enhancements to Distinguish between Deterrence and Incapacitation. *The Journal of Law and Economics*, 42(S1):343–364.

Kottelenberg, M. J. and Lehrer, S. F. (2017). Targeted or Universal Coverage? Assessing

Heterogeneity in the Effects of Universal Child Care. *Journal of Labor Economics*, 35(3):609–653.

Landersø, R., Nielsen, H. S., and Simonsen, M. (2017). School Starting Age and the Crime-age Profile. *The Economic Journal*, 127(602):1096–1118.

Loeb, S., Bridges, M., Bassok, D., Fuller, B., and Rumberger, R. W. (2007). How Much is Too Much? The Influence of Preschool Centers on Children's Social and Cognitive Development. *Economics of Education Review*, 26(1):52–66. The Economics of Early Childhood Education.

Machin, S., Marie, O., and Vujić, S. (2011). The Crime Reducing Effect of Education. *The Economic Journal*, 121(552):463–484.

Magnuson, K. A., Ruhm, C., and Waldfogel, J. (2007). Does Prekindergarten Improve School Preparation and Performance? *Economics of Education Review*, 26(1):33–51. The Economics of Early Childhood Education.

Mello, S. (2019). More COPS, Less Crime. *Journal of Public Economics*, 172:174–200.

Norwegian Directorate for Education and Training (2017). Hva gjør PP-tjenesten? https://www.udir.no/kvalitet-og-kompetanse/samarbeid/pp-tjenesten/hva-gjor-pp-tjenesten/#124909. Read 10 June 2021.

Norwegian Labour and Welfare Administration (2020). Kontantstøtte. https://www.nav.no/no/person/familie/barnetrygd-og-kontantstotte/kontantstotte2. Read 15 October 2020.

Norwegian Ministry of Administration and Consumer Affairs (1972). NOU 1972:39 Preschools. Technical Report.

Norwegian Ministry of Education (1992-1993). St.meld. nr. 40:..vi smaa, en alen lange; om 6-åringer i skolen - konsekvenser for skoleløpet og retningslinjer for dets innhold. White Paper.

Norwegian Ministry of Education (1993-1994). Innst. o. nr. 36: Innstilling fra kirke-, utdannings- og forskningskomiteen om lov om endringer i lov av 13. juni 1969 nr. 24 om grunnskolen. Committee Recommendation.

Pratt, T. and Cullen, F. (2000). The Empirical Status of Gottfredson and Hirschi's General Theory of Crime: A Meta-Analysis. *Criminology*, 38:931–964.

Rønsen, M. (2001). Market Work, Child Care and the Division of Household Labour. R.

Rønsen, M. (2009). Long-term Effects of Cash for Childcare on Mothers' Labour Supply. *LABOUR*, 23:507–533.

Schweinhart, L., Montie, J., Xiang, Z., Barnett, W., Belfield, C., Nores, M., and Ypsilanti, M. (2005). Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40. Summary, Conclusions, and Frequently Asked Questions.

Statistics Norway (2009). Kriminalitet og rettsvesen. Technical Report.

Statistics Norway (2019). Laveste andel mottakere på 20 år. Tech Report.

Statistics Norway (2020a). Etterforskede lovbrudd. https://www.ssb.no/lovbrudde. Read 22 January 2021.

Statistics Norway (2020b). Straffereaksjoner. https://www.ssb.no/straff. Read 25 January 2021.

Statistics Norway (2021). Anmeldte lovbrudd og ofre. https://www.ssb.no/lovbrudda. Read 25 January 2021.

Sviatschi, M. M. (2018). Making a Narco: Childhood Exposure to Illegal Labor Markets and Criminal Life Paths. Working Papers 2018-03, Princeton University, Woodrow Wilson School of Public and International Affairs, Research Program in Development Studies.

Thompson, O. (2018). Head Start's Long-Run Impact: Evidence from the Program's Introduction. *Journal of Human Resources*, 53(4):1100–1139.

U.S. Bureau of Justice Statistics (2020). Prisoners in 2019. Technical Report.

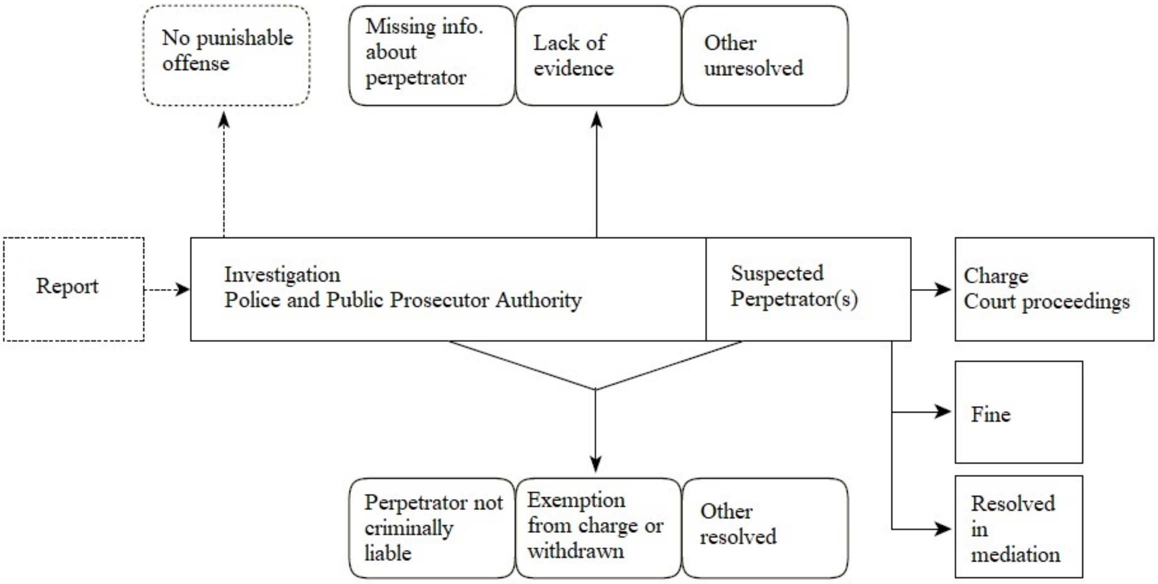
Vollaard, B. and Hamed, J. (2012). Why the Police Have an Effect on Violent Crime After All: Evidence from the British Crime Survey. *The Journal of Law and Economics*, 55(4):901–924.

Weisburd, S. (2021). Police Presence, Rapid Response Rates, and Crime Prevention. *The Review of Economics and Statistics*, (103 (2)):280–293.

Figures

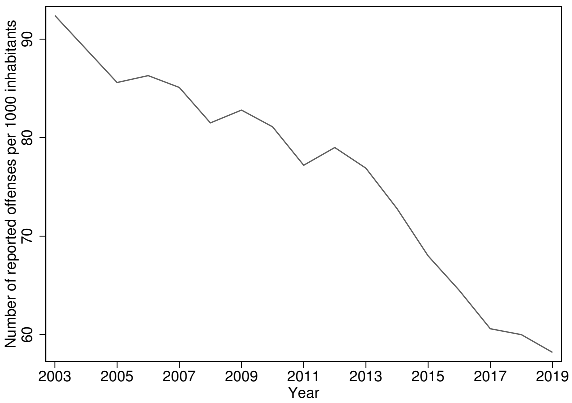
Crime in Norway

Figure 1: Criminal Offenses and Prosecution Decisions

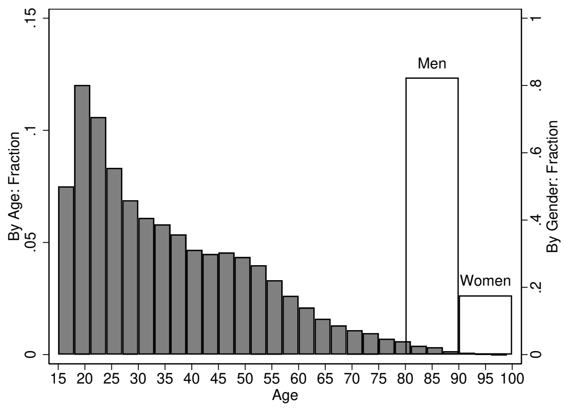


Source: Statistics Norway (2009)
Notes: This figure is an illustration of the criminal offenses and prosecution decisions in the Norwegian context.

Figure 2: Descriptive Figures, Crime in Norway



(A) Number of Reported Offenses per 1000 inhabitants

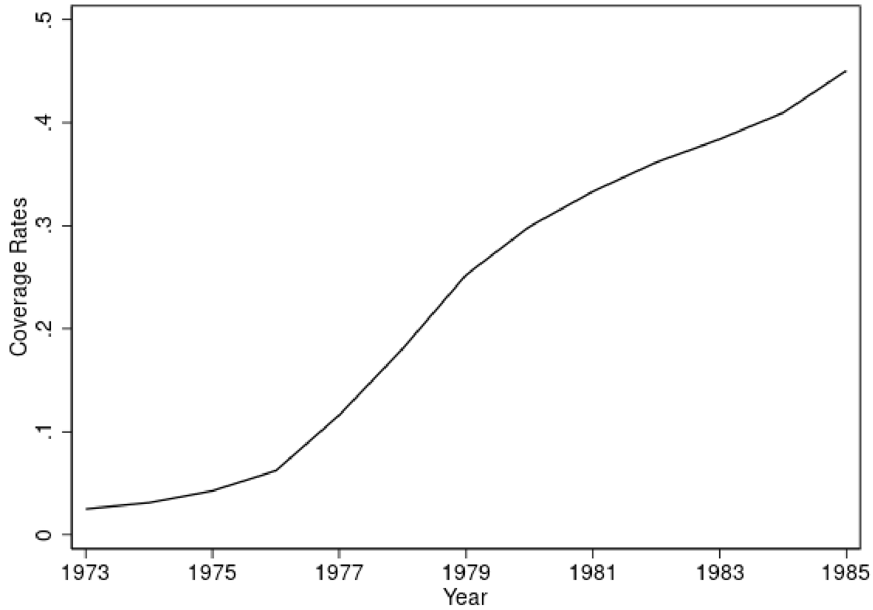


(B) Charged Individuals by Age and Gender

Notes: Panel (A) displays the number of reported offenses per 1000 inhabitants between 2003 and 2019. The data comes from official statistics from Statistics Norway. Panel B displays the age distribution and gender of charged individuals in 2018, based on own calculations of register data.

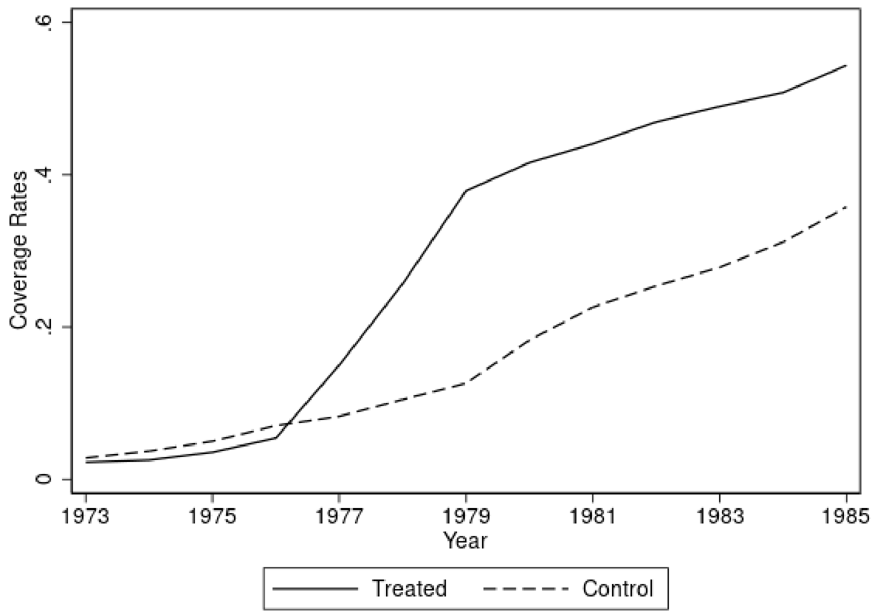
Introduction of Universal Child Care

Figure 3: Average Child Care Coverage



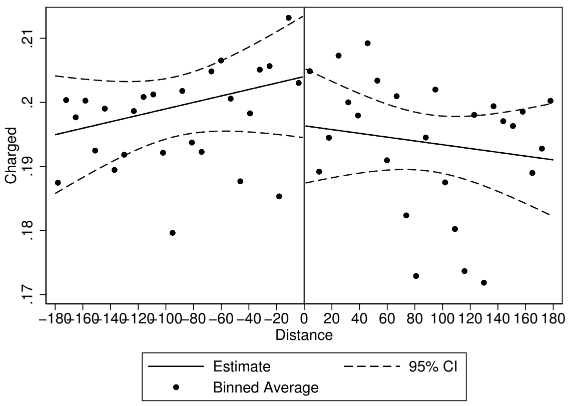
Notes: This figure displays the average municipal child care coverage rate for children ages 3-6 between 1973 and 1985.

Figure 4: Average Child Care Coverage, Treatment and Control Municipalities, 1973-1985

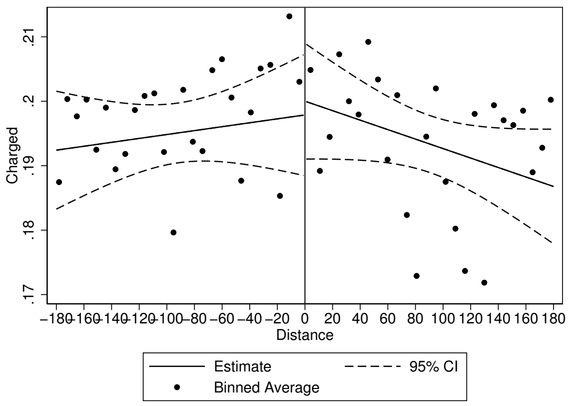


Notes: This figure displays the average municipal child care coverage rate for children ages 3-6 between 1973 and 1985 in treatment and control municipalities. Treatment (control) municipalities are defined as having an above (below) median increase in the child care coverage rate between 1976 and 1979.

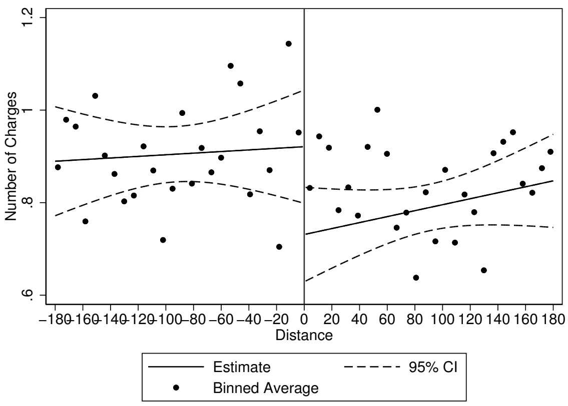
Figure 5: Regression Discontinuity Figures, Criminal Charges



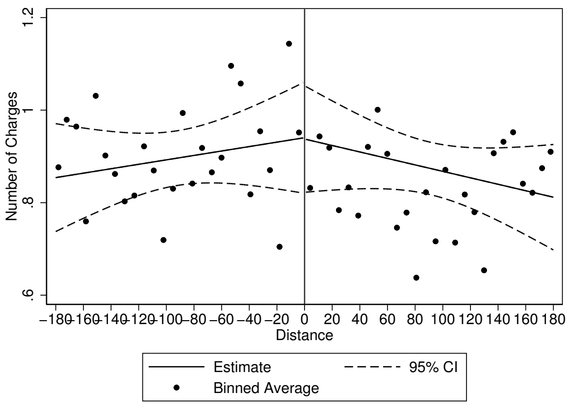
(A) Charged,
Cutoff: January 1, 1991



(B) Charged,
Cutoff: January 1, 1990



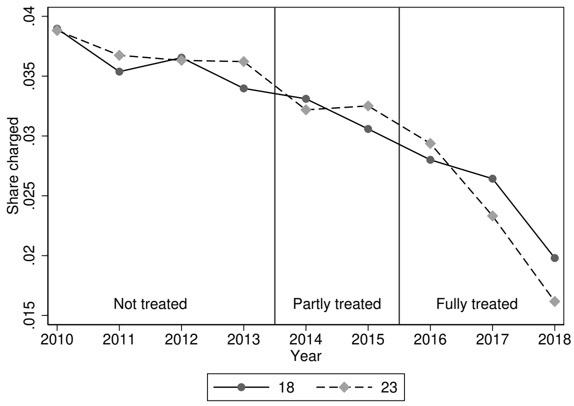
(C) Number of Charges,
Cutoff: January 1, 1991



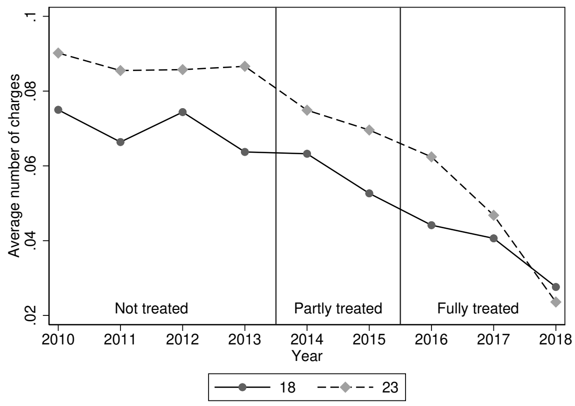
(D) Number of Charges,
Cutoff: January 1, 1990

Notes: This figure displays the regression discontinuity figures for the treatment group (born around January 1, 1991) for an indicator of being charged (panel (A)) and number of charges(panel (C)), and the equivalent for control group (born around January 1, 1990) in panels (b) and (d). I use linear trends on each side of the cutoffs, and a bandwidth of 180 days. Outcomes are measured between 2008 and 2018.

Figure 6: Criminal Charges, 18- vs 23-Year-Olds



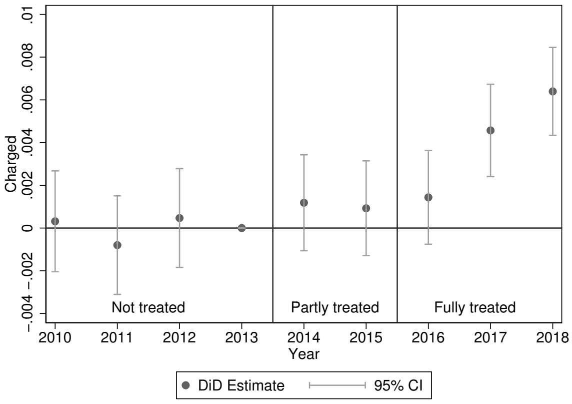
(A) Share Charged



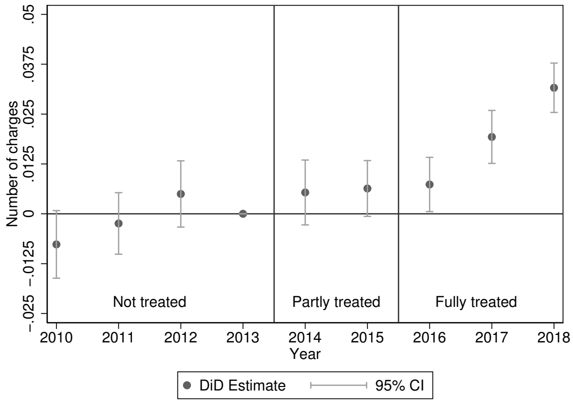
(B) Number of Charges

Notes: This figure is an illustration of the difference-in-differences strategy discussed in section 6. I display the share charged yearly in panel (A), and average yearly number of charges in panel (B), for 18- and 23-year-olds between 2010 and 2018.

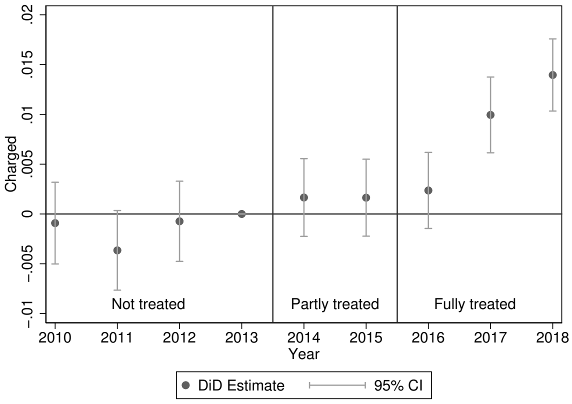
Figure 7: The Effects of Cash for Care on Criminal Charges, Yearly Difference-in-Differences Estimates, ITT, 2010-2018



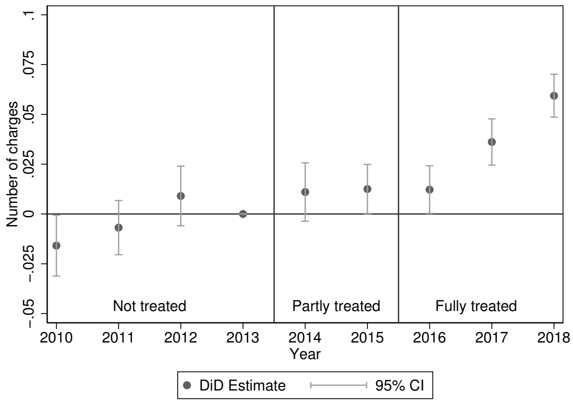
(A) All: Charged



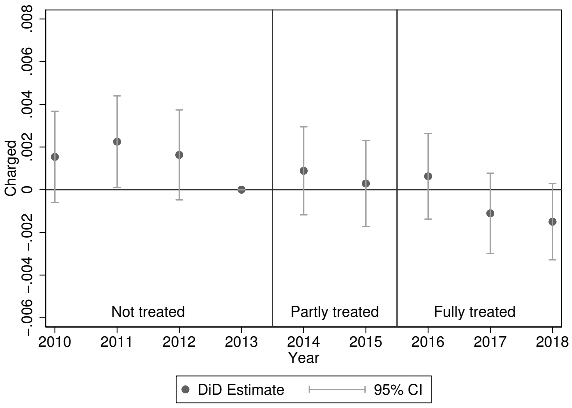
(B) All: Number of charges



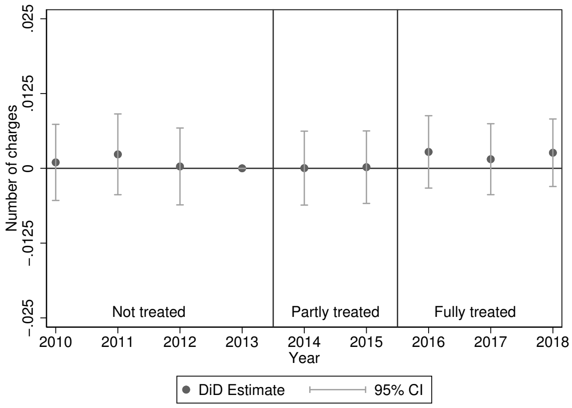
(C) Males: Charged



(D) Males: Number of charges



(E) Females: Charged



(F) Females: Number of charges

Notes: This figure displays the β_k coefficients along with 95% confidence intervals from the specification $Y_{it} = \beta_1 + \beta_2 Treat_i + \sum_{k=2010}^{2018} \beta_k (Treat_i \times Year_k) + \beta_4 X_i + \theta_t + \epsilon_{it}$, with 2012 as the base year (see section 7 for more information). The DiD estimator is the difference in mean between 18-year-olds and 23-27-year-olds in the given year, relative to the same difference in 2013. Estimates are intention-to-treat (ITT). The outcomes are an indicator for being charged in a specific year, and the yearly total number of charges. Robust standard errors are used to calculate confidence intervals.

Tables

Introduction of Subsidized Child Care

Table 1: Effects of Child Care Coverage on Criminal Charges, ITT

	Charged		Number of charges	
	(1)	(2)	(3)	(4)
Panel A: All				
DiD	-0.0085	-0.0075*	-0.0482	-0.0382
	(0.0052)	(0.0045)	(0.0747)	(0.0698)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.18	0.18	0.96	0.96
Observations	561039	561039	561039	561039
Panel B: Males				
DiD	-0.0167**	-0.0148*	-0.0982	-0.0855
	(0.0083)	(0.0076)	(0.1450)	(0.1430)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.28	0.28	1.65	1.65
Observations	285694	285694	285694	285694
Panel C: Females				
DiD	-0.0014	-0.0004	-0.0073	-0.0006
	(0.0029)	(0.0028)	(0.0297)	(0.0303)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.07	0.07	0.24	0.24
Observations	275345	275345	275345	275345

Notes: This table displays the β_3 coefficient from equation (1). The DiD estimate is the difference in mean outcomes between post-reform (born in 1973-1976) and pre-reform cohorts (born in 1967-1969) in treated municipalities, relative to the same difference in control municipalities. Results for phase-in cohorts (born in 1970-1973) from the same estimations are showed in table A.6. Estimates are intention-to-treat. Outcomes are measured between 1992 and 2006. Columns (1) and (2) shows the likelihood of being charged during this time period, and columns (3) and (4) the total number of charges in this time period. Standard errors are clustered on the municipality level and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table 2: Effects of Child Care Coverage on Criminal Charges, Type of Offenses, ITT

	Offenses for profit	Violent or sex. offenses	Narcotics offenses	Traffic offenses	Other offenses
	(1)	(2)	(3)	(4)	(5)
Panel A: All					
DiD	0.0006	0.0007	0.0009	-0.0084*	-0.0052*
	(0.0017)	(0.0013)	(0.0013)	(0.0044)	(0.0029)
Controls	Yes	Yes	Yes	Yes	Yes
Pre-reform mean, control	0.05	0.03	0.02	0.11	0.07
Observations	561039	561039	561039	561039	561039
Panel B: Males					
DiD	-0.0008	0.0015	0.0019	-0.0161**	-0.0086
	(0.0031)	(0.0026)	(0.0023)	(0.0074)	(0.0053)
Controls	Yes	Yes	Yes	Yes	Yes
Pre-reform mean, control	0.07	0.05	0.04	0.18	0.12
Observations	285694	285694	285694	285694	285694
Panel C: Females					
DiD	0.0020	-0.0004	-0.0003	-0.0009	-0.0019
	(0.0013)	(0.0007)	(0.0010)	(0.0022)	(0.0012)
Controls	Yes	Yes	Yes	Yes	Yes
Pre-reform mean, control	0.02	0.00	0.01	0.04	0.02
Observations	275345	275345	275345	275345	275345

Notes: This table displays the β_3 coefficient from equation (1). The DiD estimate is the difference in mean outcomes between post-reform (born in 1973-1976) and pre-reform cohorts (born in 1967-1969) in treated municipalities, relative to the same difference in control municipalities. Estimates are intention-to-treat. The outcomes are indicators for being charged with a specific type of offense at any time between 1992 and 2006. Standard errors are clustered on the municipality level and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table 3: Effects of Child Care Coverage on Criminal Charges, ITT, Heterogeneity by Mothers' Education

	Charged		Number of charges	
	(1)	(2)	(3)	(4)
Panel A: Less than High School				
DiD	-0.0073	-0.0079	0.0350	0.0313
	(0.0067)	(0.0060)	(0.1543)	(0.1490)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.20	0.20	1.25	1.25
Observations	207847	207847	207847	207847
Panel B: High School or More				
DiD	-0.0077	-0.0061	-0.0621	-0.0594
	(0.0051)	(0.0042)	(0.0602)	(0.0556)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.15	0.15	0.73	0.73
Observations	353192	353192	353192	353192

Notes: This table displays the β_3 coefficient from equation (1), separately for each level of maternal education. The DiD estimate is the difference in mean outcomes between post-reform (born in 1973-1976) and pre-reform cohorts (born in 1967-1969) in treated municipalities, relative to the same difference in control municipalities. Estimates are intention-to-treat. Outcomes are measured between 1992 and 2006. Columns (1) and (2) show the likelihood of being charged during this time period, and columns (3) and (4) the total number of charges in this time period. Standard errors are clustered on the municipality level and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table 4: Effects of Child Care Coverage on Criminal Charges, ITT, Heterogeneity by Fathers' Income

	Charged		Number of charges	
	(1)	(2)	(3)	(4)
Panel A: 1st Quartile				
DiD	-0.0158	-0.0135	-0.2679	-0.2328
	(0.0106)	(0.0083)	(0.1969)	(0.1737)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.20	0.20	1.23	1.23
Observations	135807	135807	135807	135807
Panel B: 2nd Quartile				
DiD	-0.0043	-0.0033	-0.0490	-0.0367
	(0.0063)	(0.0062)	(0.0896)	(0.0931)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.17	0.17	0.88	0.88
Observations	138139	138139	138139	138139
Panel C: 3rd Quartile				
DiD	-0.0116	-0.0107	0.0476	0.0520
	(0.0079)	(0.0066)	(0.0987)	(0.0927)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.16	0.16	0.88	0.88
Observations	136542	136542	136542	136542
Panel D: 4th Quartile				
DiD	-0.0074	-0.0058	-0.0473	-0.0445
	(0.0058)	(0.0055)	(0.0851)	(0.0880)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.15	0.15	0.67	0.67
Observations	135509	135509	135509	135509

Notes: This table displays the β_3 coefficient from equation (1), estimated separately for each paternal income quartile. The DiD estimate is the difference in mean outcomes between post-reform (born in 1973-1976) and pre-reform cohorts (born in 1967-1969) in treated municipalities, relative to the same difference in control municipalities. Estimates are intention-to-treat. Outcomes are measured between 1992 and 2006. Columns (1) and (2) show the likelihood of being charged during this time period, and columns (3) and (4) the total number of charges in this time period. Standard errors are clustered on the municipality level and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Compulsory Child Care for Six-Year-Olds

Table 5: Effects of Compulsory Child Care for Six-Year-Olds on Criminal Charges

	Charged		Number of charges	
	(1)	(2)	(3)	(4)
Panel A: All				
DiRD	-0.0059	-0.0053	-0.1233**	-0.1140*
	(0.0055)	(0.0052)	(0.0625)	(0.0606)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.19	0.19	0.89	0.89
Observations	117152	117152	117152	117152
Panel B: Males				
DiRD	-0.0138	-0.0110	-0.2886**	-0.2561**
	(0.0089)	(0.0086)	(0.1150)	(0.1129)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.30	0.30	1.49	1.49
Observations	60152	60152	60152	60152
Panel C: Females				
DiRD	0.0013	0.0012	0.0448	0.0510
	(0.0052)	(0.0051)	(0.0427)	(0.0470)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.08	0.08	0.25	0.25
Observations	57000	57000	57000	57000

Notes: This table displays the β_5 coefficient from equation (2), using linear trends, 180 days bandwidth and triangular weights (see section 6 for more information). The DiRD estimate is the difference in mean outcomes between those born just after December 31, 1990, relative to the same difference in the previous year. Outcomes are measured between 2008 and 2018. Columns (1) and (2) show the likelihood of being charged during this time period, and columns (3) and (4) the total number of charges in this time period. Standard errors are clustered on the municipality level and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

TABLE 6: Effects of Compulsory Child Care for Six-Year-Olds on Criminal Charges, Type of Offense

	Offenses for profit	Violent or sex. offenses	Narcotics offenses	Traffic offenses	Other offenses
	(1)	(2)	(3)	(4)	(5)
Panel A: All					
DiRD	-0.0031	-0.0040	-0.0075**	0.0036	-0.0120***
	(0.0021)	(0.0025)	(0.0032)	(0.0037)	(0.0038)
Controls	Yes	Yes	Yes	Yes	Yes
Pre-reform mean, control	0.04	0.04	0.06	0.10	0.08
Observations	117152	117152	117152	117152	117152
Panel B: Males					
DiRD	-0.0074**	-0.0086**	-0.0160***	0.0061	-0.0174**
	(0.0035)	(0.0044)	(0.0057)	(0.0064)	(0.0067)
Controls	Yes	Yes	Yes	Yes	Yes
Pre-reform mean, control	0.05	0.06	0.09	0.16	0.14
Observations	60152	60152	60152	60152	60152
Panel C: Females					
DiRD	0.0017	0.0015	0.0018	0.0010	-0.0061*
	(0.0026)	(0.0019)	(0.0024)	(0.0031)	(0.0032)
Controls	Yes	Yes	Yes	Yes	Yes
Pre-reform mean, control	0.02	0.01	0.02	0.03	0.03
Observations	57000	57000	57000	57000	57000

Notes: This table displays the β_5 coefficient from equation (2), using linear trends, 180 days bandwidth and triangular weights (see section 6 for more information). The DiRD estimate is the difference in mean outcomes between those born just after December 31, 1990, relative to the same difference in the previous year. Outcomes are measured between 2008 and 2018. The outcomes are indicators for being charged with a specific type of offense at any time between 2008 and 2018. Standard errors are clustered on the municipality level and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table 7: Effects of Compulsory Child Care for Six-Year-Olds on Criminal Charges, Heterogeneity by Mothers' Education

	Charged		Number of charges	
	(1)	(2)	(3)	(4)
Panel A: Less than High School				
DiRD	-0.0102	-0.0105	-0.0250	-0.0531
	(0.0118)	(0.0104)	(0.1412)	(0.1405)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.25	0.25	1.44	1.44
Observations	36936	36936	36936	36936
Panel B: High School				
DiRD	-0.0048	-0.0032	-0.2349***	-0.2177***
	(0.0078)	(0.0077)	(0.0763)	(0.0746)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.19	0.19	0.73	0.73
Observations	48230	48230	48230	48230
Panel C: Higher Education				
DiRD	0.0055	0.0033	-0.0340	-0.0387
	(0.0096)	(0.0093)	(0.0860)	(0.0833)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.12	0.12	0.38	0.38
Observations	30510	30510	30510	30510

Notes: This table displays the β_5 coefficient from equation (2), separately for each level of mothers' education. I use linear trends, 180 days bandwidth and triangular weights (see section 6 for more information). The DiRD estimate is the difference in mean outcomes between those born just after December 31, 1990, relative to the same difference in the previous year. Outcomes are measured between 2008 and 2018. Columns (1) and (2) show the likelihood of being charged during this time period, and columns (3) and (4) the total number of charges in this time period. Standard errors are clustered on the municipality level and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table 8: Effects of Compulsory Child Care for Six-Year-Olds on Criminal Charges, Heterogeneity by Fathers' Income

	Charged		Number of charges	
	(1)	(2)	(3)	(4)
Panel A: Quartile 1				
DiRD	-0.0173	-0.0120	-0.3533*	-0.2985
	(0.0115)	(0.0113)	(0.1890)	(0.1929)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.25	0.25	1.43	1.43
Observations	28815	28815	28815	28815
Panel B: Quartile 2				
DiRD	-0.0023	-0.0038	-0.0787	-0.1193
	(0.0102)	(0.0099)	(0.1207)	(0.1156)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.20	0.20	0.87	0.87
Observations	28630	28630	28630	28630
Panel C: Quartile 3				
DiRD	0.0022	0.0006	0.1339	0.1623
	(0.0109)	(0.0106)	(0.1063)	(0.1059)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.17	0.17	0.64	0.64
Observations	28777	28777	28777	28777
Panel D: Quartile 4				
DiRD	-0.0054	-0.0055	-0.1717**	-0.1411**
	(0.0096)	(0.0095)	(0.0734)	(0.0694)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.15	0.15	0.48	0.48
Observations	28819	28819	28819	28819

Notes: This table displays the β_5 coefficient from equation (2), separately for each quartile of fathers' income. I use linear trends, 180 days bandwidth and triangular weights (see section 6 for more information). The DiRD estimate is the difference in mean outcomes between those born just after December 31, 1990, relative to the same difference in the previous year. Outcomes are measured between 2008 and 2018. Columns (1) and (2) show the likelihood of being charged during this time period, and columns (3) and (4) the total number of charges in this time period. Standard errors are clustered on the municipality level and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table 9: Effects of Cash-for-Care on Criminal Charges, ITT

	Charged		Number of charges	
	(1)	(2)	(3)	(4)
Panel A: All				
DiD	0.0010*	0.0042***	0.0082***	0.0205***
	(0.0006)	(0.0006)	(0.0018)	(0.0020)
Controls	No	Yes	No	Yes
Pre-reform mean, 18-year-olds	0.04	0.04	0.07	0.07
Observations	2943581	2943581	2943581	2943581
Panel B: Males				
DiD	0.0049***	0.0100***	0.0172***	0.0388***
	(0.0010)	(0.0011)	(0.0032)	(0.0036)
Controls	No	Yes	No	Yes
Pre-reform mean, 18-year-olds	0.06	0.06	0.11	0.11
Observations	1514156	1514156	1514156	1514156
Panel C: Females				
DiD	-0.0028***	-0.0019***	-0.0007	0.0015
	(0.0005)	(0.0006)	(0.0013)	(0.0014)
Controls	No	Yes	No	Yes
Pre-reform mean, 18-year-olds	0.02	0.02	0.02	0.02
Observations	1429425	1429425	1429425	1429425

Notes: This table displays the β_4 coefficient from equation (3). The DiD estimate is the difference in mean outcomes in between 18-year-olds and 23-27-year-olds in 2016-2018, relative to the same difference in 2010-2013. Results for phase-in years 2014-2015 from the same estimations are showed in table C.3. Estimates are intention-to-treat. Outcomes are measured yearly. Columns (1) and (2) show the likelihood of being charged in given year, and columns (3) and (4) the yearly total number of charges. Standard errors are robust and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table 10: Effects of Cash-for-Care on Criminal Charges, ITT, Type of Offense

	Offenses for profit	Violent or sex. offenses	Narcotics offenses	Traffic offenses	Other offenses
	(1)	(2)	(3)	(4)	(5)
Panel A: All					
DiD	0.0014***	0.0005**	0.0038***	0.0021***	0.0002
	(0.0002)	(0.0002)	(0.0003)	(0.0004)	(0.0003)
Controls	Yes	Yes	Yes	Yes	Yes
Pre-reform mean, 18-year-olds	0.01	0.01	0.01	0.01	0.01
Observations	2943581	2943581	2943581	2943581	2943581
Panel B: Males					
DiD	0.0022***	0.0014***	0.0070***	0.0036***	0.0029***
	(0.0004)	(0.0004)	(0.0006)	(0.0007)	(0.0006)
Controls	Yes	Yes	Yes	Yes	Yes
Pre-reform mean, 18-year-olds	0.01	0.01	0.01	0.02	0.02
Observations	1514156	1514156	1514156	1514156	1514156
Panel C: Females					
DiD	0.0005*	-0.0004*	0.0005*	0.0006**	-0.0026***
	(0.0003)	(0.0002)	(0.0003)	(0.0003)	(0.0003)
Controls	Yes	Yes	Yes	Yes	Yes
Pre-reform mean, 18-year-olds	0.00	0.00	0.00	0.00	0.01
Observations	1429425	1429425	1429425	1429425	1429425

Notes: This table displays the β_4 coefficient from equation (3). The DiD estimate is the difference in mean outcomes in between 18-year-olds and 23-27-year-olds in 2016-2018, relative to the same difference in 2010-2013. Estimates are intention-to-treat. The outcomes are indicators for being charged with a specific type of offense in a given year. Standard errors are robust and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table 11: Effects of Cash-for-Care on Criminal Charges, ITT, Heterogeneity by Mothers' Education

	Charged		Number of charges	
	(1)	(2)	(3)	(4)
Panel A: Lower than High School				
DiD	0.0010	0.0030**	0.0073	0.0196***
	(0.0014)	(0.0015)	(0.0046)	(0.0048)
Controls	No	Yes	No	Yes
Pre-reform mean, 18-year-olds	0.05	0.05	0.12	0.12
Observations	935894	935894	935894	935894
Panel B: High School				
DiD	0.0029***	0.0049***	0.0087***	0.0160***
	(0.0009)	(0.0009)	(0.0025)	(0.0030)
Controls	No	Yes	No	Yes
Pre-reform mean, 18-year-olds	0.03	0.03	0.06	0.06
Observations	1213490	1213490	1213490	1213490
Panel C: Higher Education				
DiD	-0.0001	0.0009	0.0052***	0.0066***
	(0.0008)	(0.0009)	(0.0017)	(0.0020)
Controls	No	Yes	No	Yes
Pre-reform mean, 18-year-olds	0.02	0.02	0.03	0.03
Observations	758044	758044	758044	758044

Notes: This table displays the β_4 coefficient from equation (3), estimated separately for each level of mothers' education. The DiD estimate is the difference in mean outcomes in between 18-year-olds and 23-27-year-olds in 2016-2018, relative to the same difference in 2010-2013. Estimates are intention-to-treat. Outcomes are measured yearly. Columns (1) and (2) show the likelihood of being charged in given year, and columns (3) and (4) the yearly total number of charges. Standard errors are robust and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table 12: Effects of Cash-for-Care on Criminal Charges, ITT, Heterogeneity by Fathers' Income

	Charged		Number of charges	
	(1)	(2)	(3)	(4)
Panel A: Quartile 1				
DiD	-0.0003	0.0056***	0.0046	0.0272***
	(0.0014)	(0.0015)	(0.0050)	(0.0051)
Controls	No	Yes	No	Yes
Pre-reform mean, 18-year-olds	0.05	0.05	0.11	0.11
Observations	720458	720458	720458	720458
Panel B: Quartile 2				
DiD	0.0011	0.0050***	0.0111***	0.0233***
	(0.0012)	(0.0013)	(0.0035)	(0.0039)
Controls	No	Yes	No	Yes
Pre-reform mean, 18-year-olds	0.04	0.04	0.07	0.07
Observations	733062	733062	733062	733062
Panel C: Quartile 3				
DiD	0.0004	0.0004	0.0071**	0.0114***
	(0.0011)	(0.0012)	(0.0029)	(0.0038)
Controls	No	Yes	No	Yes
Pre-reform mean, 18-year-olds	0.03	0.03	0.05	0.05
Observations	731816	731816	731816	731816
Panel D: Quartile 4				
DiD	0.0026***	0.0039***	0.0085***	0.0107***
	(0.0010)	(0.0011)	(0.0024)	(0.0029)
Controls	No	Yes	No	Yes
Pre-reform mean, 18-year-olds	0.02	0.02	0.04	0.04
Observations	725632	725632	725632	725632

Notes: This table displays the β_4 coefficient from equation (3), estimated separately for each quartile of fathers' income. The DiD estimate is the difference in mean outcomes in between 18-year-olds and 23-27-year-olds in 2016-2018, relative to the same difference in 2010-2013. Estimates are intention-to-treat. Outcomes are measured yearly. Columns (1) and (2) show the likelihood of being charged in given year, and columns (3) and (4) the yearly total number of charges. Standard errors are robust and reported in parentheses.

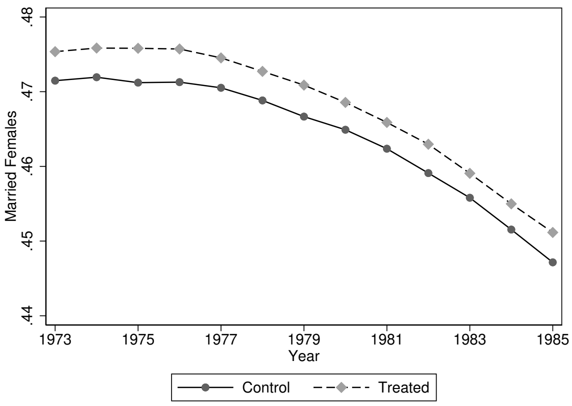
$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Appendices

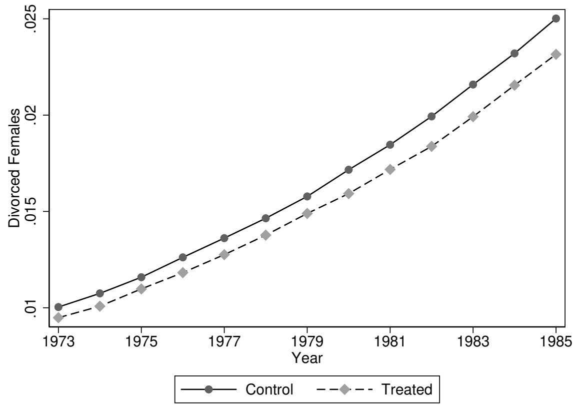
A Introduction of Formal Child Care

A.1 Figures

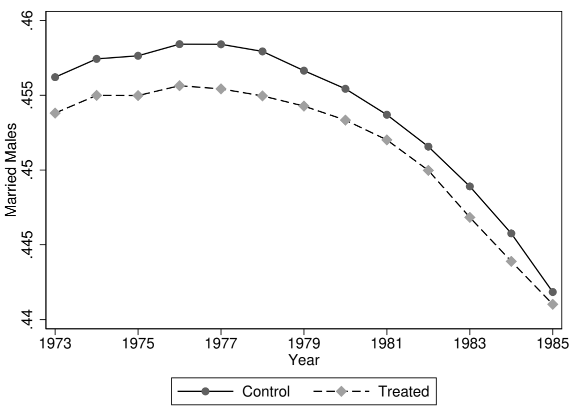
Figure A.1: Municipality Characteristics 1973-1985 1



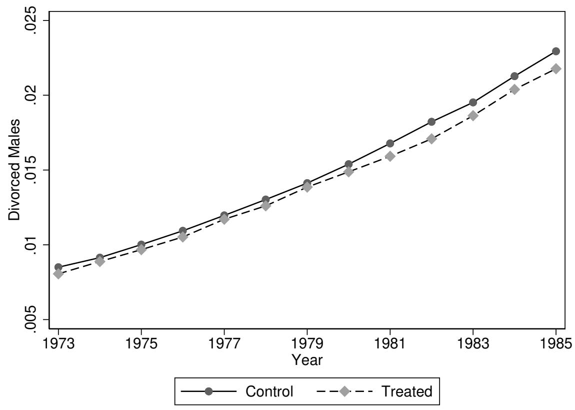
(a) Married Females



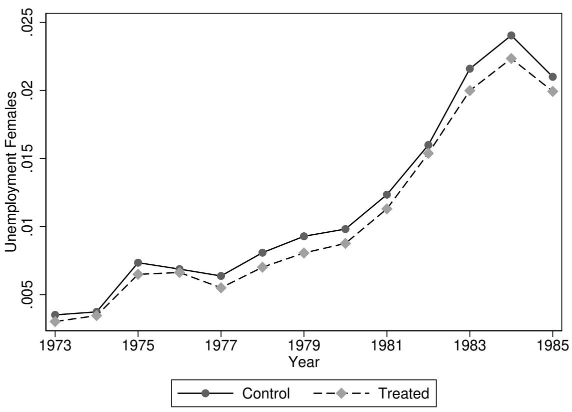
(b) Divorced Females



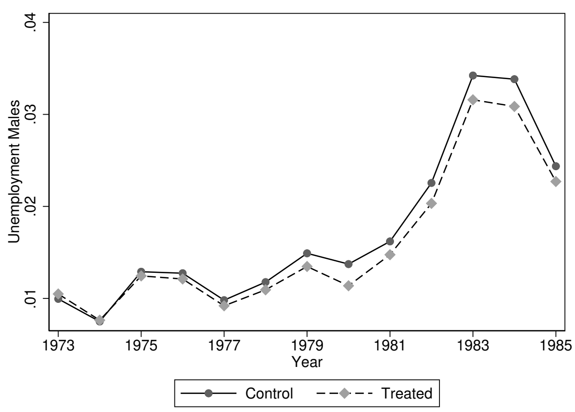
(c) Married Males



(d) Divorced Males



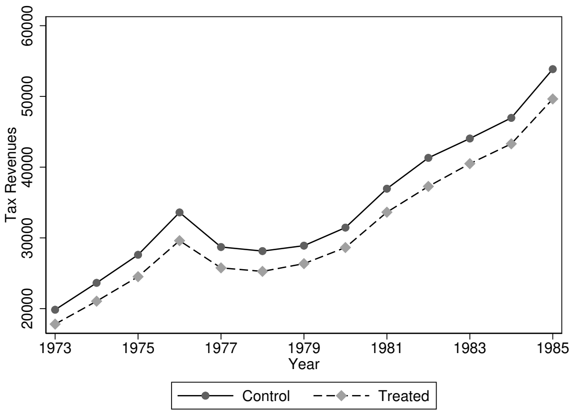
(e) Unemployed Females



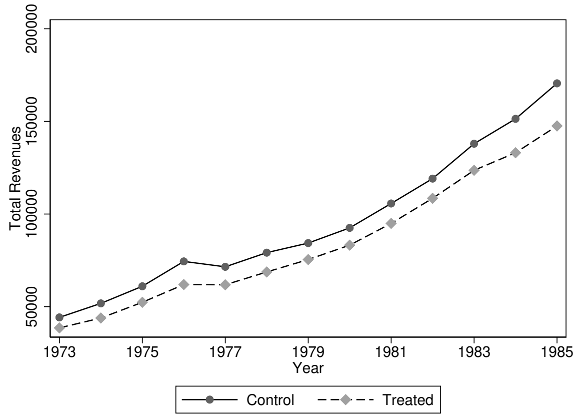
(f) Unemployed Males

Notes: These figures display trends in municipality characteristics in 1973-1975 for treatment and control municipalities (see section 5 for definition). Data source: Norwegian Centre for Research Data (NSD).

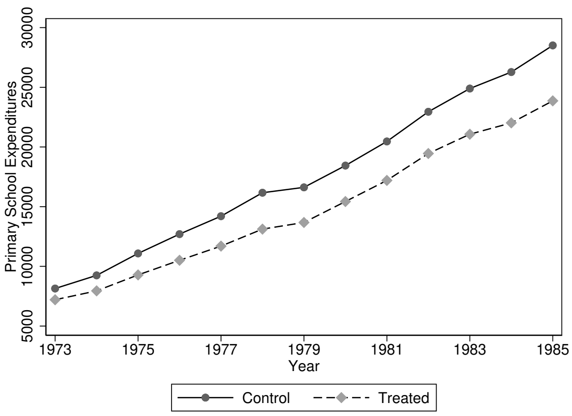
Figure A.2: Municipality Characteristics 1973-1985 2



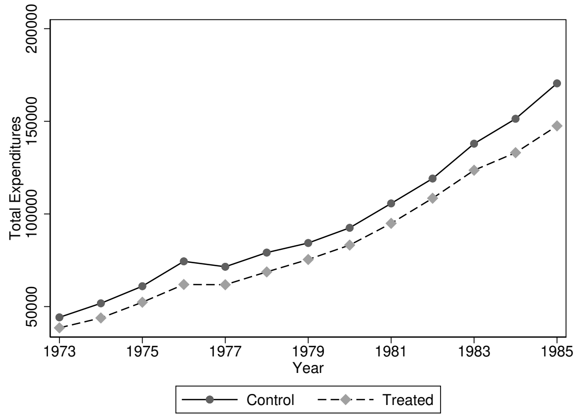
(A) Tax Revenues



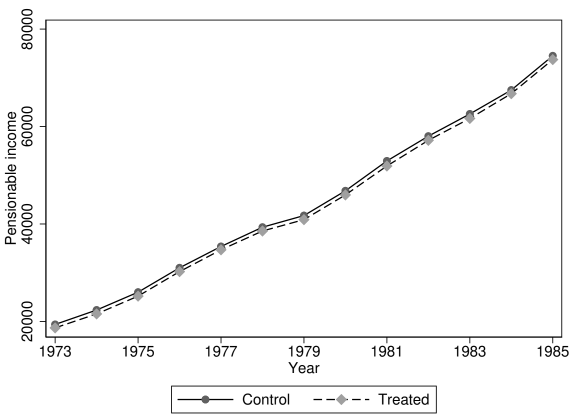
(B) Total Revenues



(C) Primary School Expenditures



(D) Total Expenditures

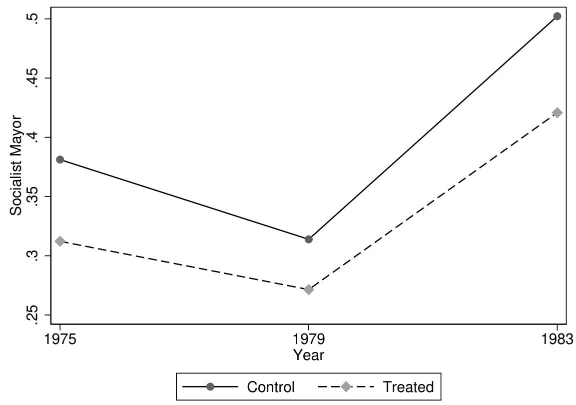


(E) Pensionable Income

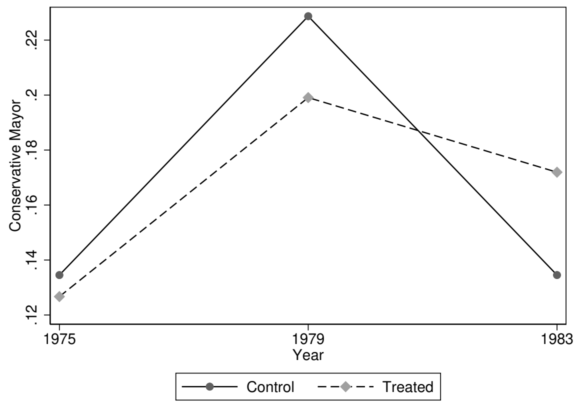
(F) Higher Education

Notes: These figures display trends in municipality characteristics in 1973-1975 for treatment and control municipalities (see section 5 for definition). Data source: Norwegian Centre for Research Data (NSD) (Panels A-D) and Norwegian Administrative Registers (Panels E-F).

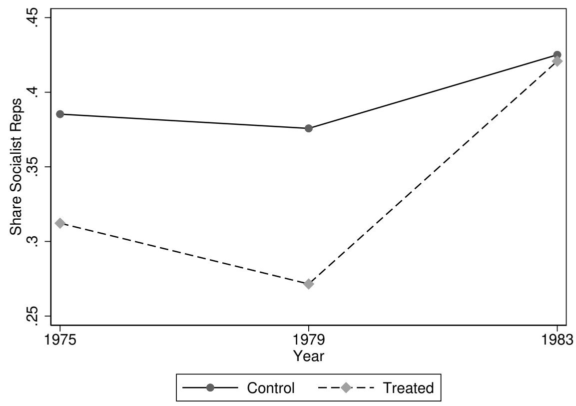
Figure A.3: Municipality Characteristics 1973-1985 3



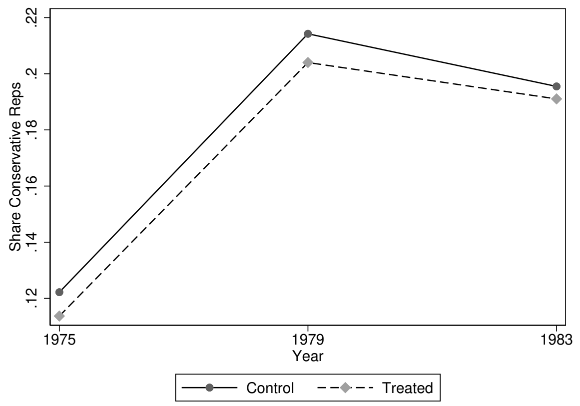
(A) Socialist Mayor



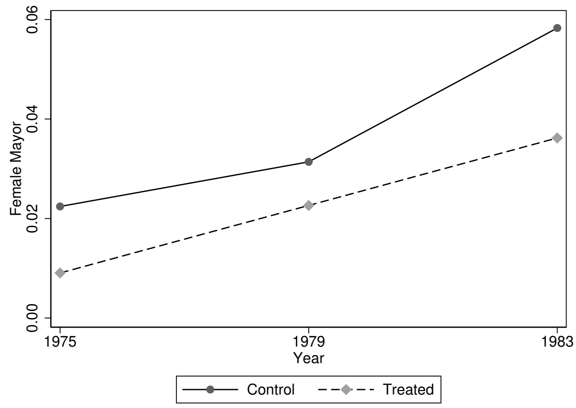
(B) Conservative Mayor



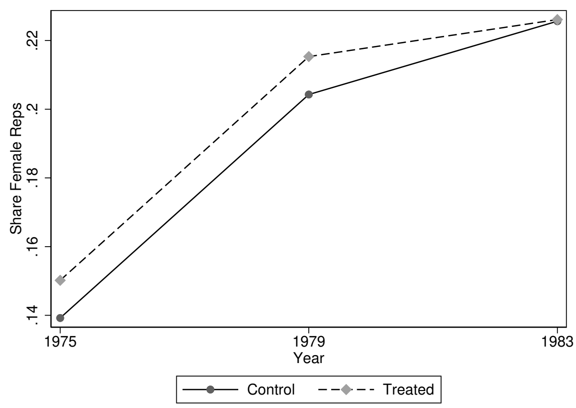
(C) Socialist Share of Elected Representatives



(D) Conservative Share of Elected Representatives



(E) Female Mayor



(F) Female Share of Elected Representatives

Notes: These figures display trends in municipality characteristics in 1973-1975 for treatment and control municipalities (see section 5 for definition). Data source: Norwegian Centre for Research Data (NSD).

A.2 Tables

Table A.1: Effects of Child Care Coverage on Criminal Charges, ITT, Treatment Intensity

	Charged		Number of charges	
	(1)	(2)	(3)	(4)
Panel A: All				
DiD	-0.0002	-0.0002	-0.0020	-0.0015
	(0.0001)	(0.0001)	(0.0018)	(0.0018)
Controls	No	Yes	No	Yes
Pre-reform mean	0.17	0.17	0.96	0.96
Observations	561039	561039	561039	561039
Panel B: Males				
DiD	-0.0004*	-0.0004*	-0.0042	-0.0038
	(0.0002)	(0.0002)	(0.0035)	(0.0036)
Controls	No	Yes	No	Yes
Pre-reform mean	0.28	0.28	1.63	1.63
Observations	285694	285694	285694	285694
Panel C: Females				
DiD	0.0000	0.0000	0.0001	0.0003
	(0.0001)	(0.0001)	(0.0010)	(0.0010)
Controls	No	Yes	No	Yes
Pre-reform mean	0.07	0.07	0.26	0.26
Observations	275345	275345	275345	275345

Notes: This table displays the β_3 coefficient from equation $Y_{ijt} = \beta_1 + \beta_2(Phase{-}in_t \times Treat_j) + \beta_3(Post_t \times Treat_j) + \beta_4 X_i + \theta_j + \gamma_t + \varepsilon_{ijt}$, where $Treat_j$ is the increase in the municipal child care coverage rate, and the other factors are explain in section (5). The DiD estimate is the difference in mean outcomes between post-reform (born in 1973-1976) and pre-reform cohorts (born in 1967-1969) in the same municipality, relative to the equivalent in municipalities with other treatment intensities. Estimates are intention-to-treat. Outcomes are measured between 1992 and 2006. Columns (1) and (2) show the likelihood of being charged during this time period, and columns (3) and (4) the total number of charges in this time period. Standard errors are clustered on the municipality level and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table A.2: Effects of Child Care Coverage on Criminal Charges, ITT, Tertile Treatment

	Charged		Number of charges	
	(1)	(2)	(3)	(4)
Panel A: All				
DiD Medium	-0.0000	0.0009	0.0033	0.0142
	(0.0064)	(0.0056)	(0.0794)	(0.0733)
DiD High	-0.0039	-0.0028	-0.0343	-0.0228
	(0.0052)	(0.0048)	(0.0799)	(0.0804)
Controls	No	Yes	No	Yes
Pre-reform mean control	0.18	0.18	0.98	0.98
Observations	561039	561039	561039	561039
Panel B: Males				
DiD Medium	-0.0004	0.0014	0.0296	0.0424
	(0.0103)	(0.0094)	(0.1530)	(0.1476)
DiD High	-0.0094	-0.0093	-0.1000	-0.1011
	(0.0080)	(0.0077)	(0.1505)	(0.1534)
Controls	No	Yes	No	Yes
Pre-reform mean control	0.28	0.28	1.67	1.67
Observations	285694	285694	285694	285694
Panel C: Females				
DiD Medium	-0.0004	0.0005	-0.0286	-0.0217
	(0.0032)	(0.0031)	(0.0283)	(0.0291)
DiD High	0.0013	0.0025	0.0329	0.0394
	(0.0036)	(0.0036)	(0.0440)	(0.0445)
Controls	No	Yes	No	Yes
Pre-reform mean control	0.07	0.07	0.26	0.26
Observations	275345	275345	275345	275345

Notes: This table displays the $\beta_3 1$ and $\beta_3 2$ coefficients from equation $Y_{ijt} = \beta_1 + \beta_2 1(Phase - in_t \times TreatMedium_j) + \beta_2 2(Phase - in_t \times TreatHigh_j) + \beta_3 1(Post_t \times TreatMedium_j) + \beta_3 1(Post_t \times TreatHigh_j) + \beta_4 X_i + \theta_j + \gamma_t + \varepsilon_{ijt}$, where $TreatMedium_j$ ($TreatHigh_j$) is in the middle (highest) tertile of the increase in the municipal child care coverage rate, and the other factors are explain in section (5). The DiD Medium (High) estimate is the difference in mean outcomes between post-reform (born in 1973-1976) and pre-reform cohorts (born in 1967-1969) in the middle (highest) treatment tertile, relative to the equivalent in municipalities in the lowest treatment tertile. Estimates are intention-to-treat. Outcomes are measured between 1992 and 2006. Columns (1) and (2) show the likelihood of being charged during this time period, and columns (3) and (4) the total number of charges in this time period. Standard errors are clustered on the municipality level and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table A.3: Effects of Child Care Coverage on Number of Charges, ITT

	≥ 2	≥ 5	≥ 10
	(1)	(2)	(3)
Panel A: All			
DiD	-0.0051*	-0.0017	0.0012
	(0.0028)	(0.0017)	(0.0011)
Controls	Yes	Yes	Yes
Pre-reform mean, control	0.08	0.03	0.02
Observations	561039	561039	561039
Panel B: Males			
DiD	-0.0106**	-0.0038	0.0016
	(0.0052)	(0.0033)	(0.0020)
Controls	Yes	Yes	Yes
Pre-reform mean, control	0.14	0.06	0.03
Observations	285694	285694	285694
Panel C: Females			
DiD	0.0002	0.0004	0.0007
	(0.0015)	(0.0008)	(0.0006)
Controls	Yes	Yes	Yes
Pre-reform mean, control	0.02	0.01	0.00
Observations	275345	275345	275345

Notes: This table displays the β_3 coefficient from equation (1). The DiD estimate is the difference in mean outcomes between post-reform (born in 1973-1976) and pre-reform cohorts (born in 1967-1969) in treated municipalities, relative to the same difference in control municipalities. Estimates are intention-to-treat. The outcomes are indicators for being charged at least 2, 5 and 10 times at any time between 1992 and 2006. Standard errors are clustered on the municipality level and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table A.4: Effects of Child Care Coverage on Criminal Charges, ITT, Ages 25-30

	Charged		Number of charges	
	(1)	(2)	(3)	(4)
Panel A: All				
DiD	-0.0020	-0.0014	-0.0175	-0.0123
	(0.0030)	(0.0026)	(0.0344)	(0.0322)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.10	0.10	0.40	0.40
Observations	561039	561039	561039	561039
Panel B: Males				
DiD	-0.0049	-0.0039	-0.0244	-0.0176
	(0.0054)	(0.0050)	(0.0666)	(0.0649)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.17	0.17	0.69	0.69
Observations	285694	285694	285694	285694
Panel C: Females				
DiD	0.0003	0.0008	-0.0155	-0.0123
	(0.0016)	(0.0016)	(0.0156)	(0.0159)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.03	0.03	0.09	0.09
Observations	275345	275345	275345	275345

Notes: This table displays the β_3 coefficient from equation (1). The DiD estimate is the difference in mean outcomes between post-reform (born in 1973-1976) and pre-reform cohorts (born in 1967-1969) in treated municipalities, relative to the same difference in control municipalities. Estimates are intention-to-treat. Outcomes are measured at ages 25-30. Columns (1) and (2) show the likelihood of being charged during this time period, and columns (3) and (4) the total number of charges in this time period. Standard errors are clustered on the municipality level and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table A.5: Effects of Child Care Coverage on Punishments, ITT

	Prison	Probation	Community service	Fine	Other
	(1)	(2)	(3)	(4)	(5)
Panel A: All					
DiD	-0.0009	-0.0000	-0.0000	-0.0103	0.0000
	(0.0013)	(0.0013)	(0.0005)	(0.0067)	(0.0001)
Controls	Yes	Yes	Yes	Yes	Yes
Pre-reform mean, control	0.02	0.02	0.00	0.41	0.00
Observations	561039	561039	561039	561039	561039
Panel B: Males					
DiD	-0.0022	-0.0004	0.0002	-0.0204**	0.0001
	(0.0024)	(0.0026)	(0.0009)	(0.0081)	(0.0001)
Controls	Yes	Yes	Yes	Yes	Yes
Pre-reform mean, control	0.04	0.04	0.01	0.56	0.00
Observations	285694	285694	285694	285694	285694
Panel C: Females					
DiD	0.0004	0.0003	-0.0003	-0.0002	0.0000
	(0.0006)	(0.0008)	(0.0004)	(0.0065)	(0.0000)
Controls	Yes	Yes	Yes	Yes	Yes
Pre-reform mean, control	0.01	0.01	0.00	0.25	0.00
Observations	275345	275345	275345	275345	275345

Notes: This table displays the β_3 coefficient from equation (1). The DiD estimate is the difference in mean outcomes between post-reform (born in 1973-1976) and pre-reform cohorts (born in 1967-1969) in treated municipalities, relative to the same difference in control municipalities. Estimates are intention-to-treat. The outcomes are indicators for being convicted to a certain type punishment at any time between 1992 and 2006. Standard errors are clustered on the municipality level and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table A.6: Effects of Child Care Coverage on Criminal Charges, ITT, Phase-In Cohorts

	Charged		Number of charges	
	(1)	(2)	(3)	(4)
Panel A: All				
DiD phase-in	0.0006	-0.0002	0.1128	0.1051
	(0.0036)	(0.0034)	(0.0731)	(0.0707)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.18	0.18	0.96	0.96
Observations	561039	561039	561039	561039
Panel B: Males				
DiD phase-in	-0.0059	-0.0059	0.1770	0.1663
	(0.0061)	(0.0059)	(0.1392)	(0.1353)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.28	0.28	1.65	1.65
Observations	285694	285694	285694	285694
Panel C: Females				
DiD phase-in	0.0054**	0.0056**	0.0266	0.0284
	(0.0026)	(0.0026)	(0.0291)	(0.0295)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.07	0.07	0.24	0.24
Observations	275345	275345	275345	275345

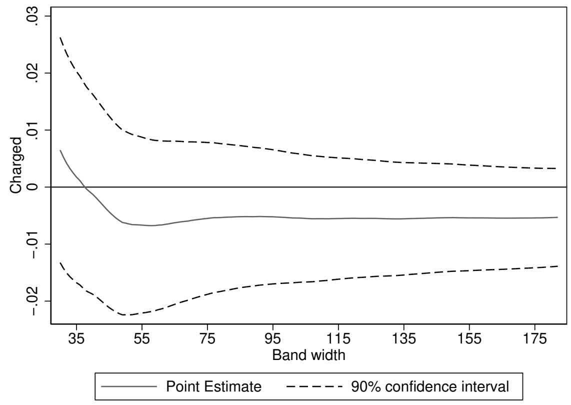
Notes: This table displays the β_2 coefficient from equation (1). The DiD estimate is the difference in mean outcomes between phase-in (born in 1973-1976) and pre-reform cohorts (born in 1967-1969) in treated municipalities, relative to the same difference in control municipalities. Results for post-reform cohorts (born in 1970-1973) from the same estimations are showed in table 1. Estimates are intention-to-treat. Outcomes are measured between 1992 and 2006. Columns (1) and (2) show the likelihood of being charged during this time period, and columns (3) and (4) the total number of charges in this time period. Standard errors are clustered on the municipality level and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

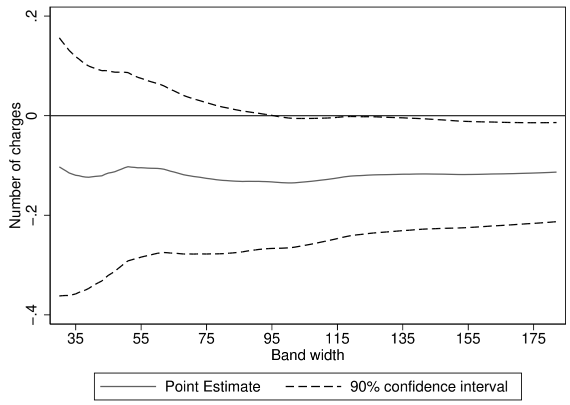
B Compulsory year of Child Care for Six-Year-Olds

B.1 Figures

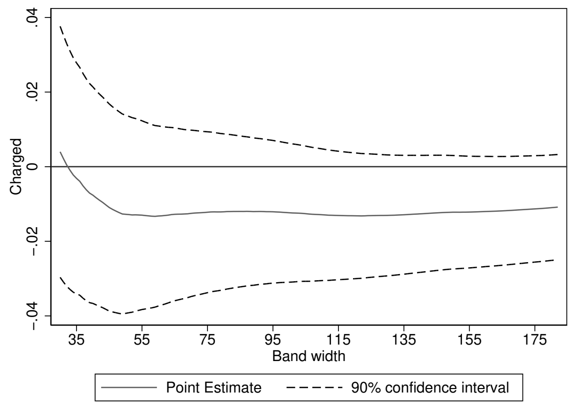
Figure B.1: Effects of Compulsory Year of Child Care on Criminal charges, Main Specification with Different Bandwidths



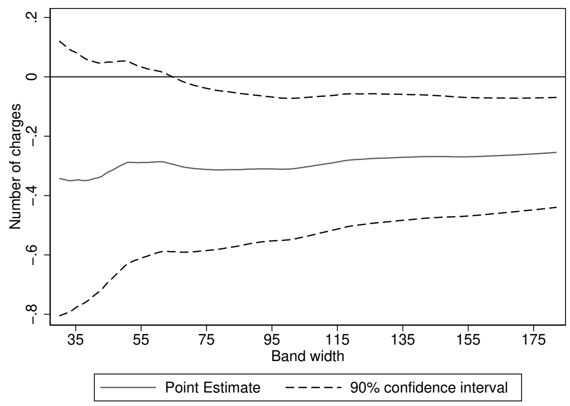
(A) All: Charged



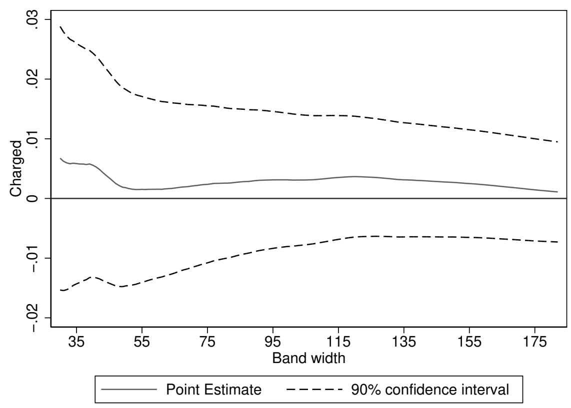
(B) All: Number of charges



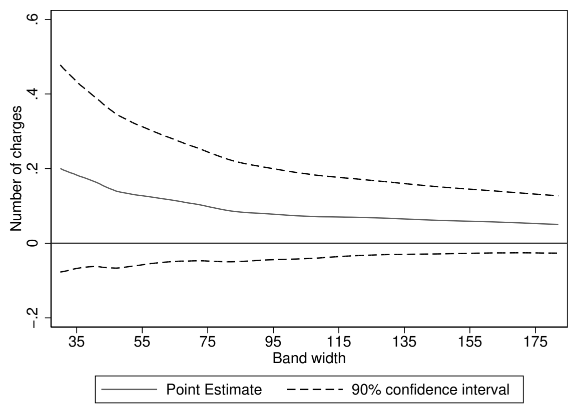
(C) Males: Charged



(D) Males: Number of charges



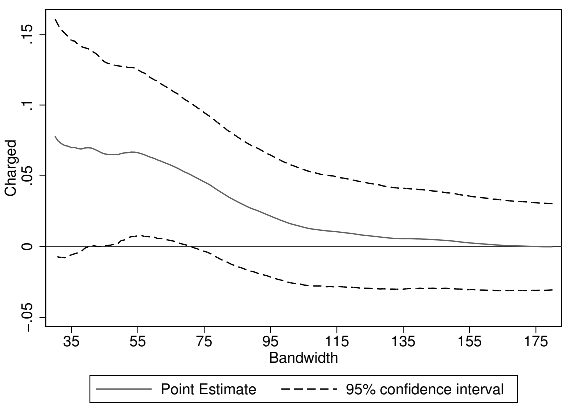
(E) Females: Charged



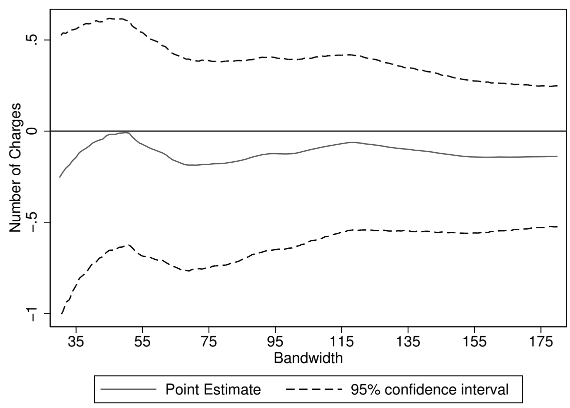
(F) Females: Number of charges

Notes: These figures display the β_5 coefficients from equation 2 along with 95% confidence intervals, using bandwidths between 30 and 180 days, together with linear trends, separate on each side of the cut-off but equal for individuals born in different year on the same side of the cutoff, and triangular weights. The difference-in-regression-discontinuity (DiRD) estimates is the estimated difference in mean outcomes between those born just after December 31, 1990 and those born just before, relative to the same difference in the previous year. Outcomes are measured between 2008 and 2018. Panels (A), (C) and (E) show the likelihood of being charged during this time period, and panels (B), (D) and (E) the total number of charges in this time period. Standard errors are clustered on the municipality level.

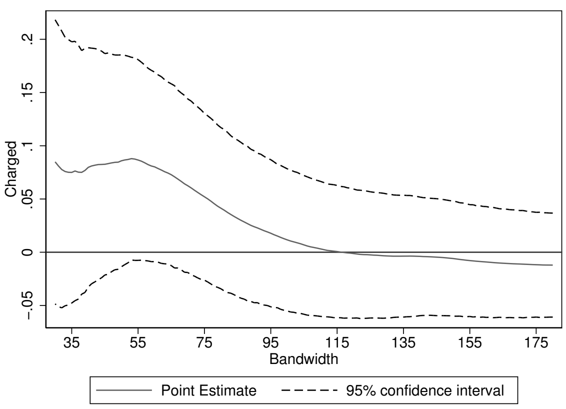
Figure B.3: Effects of Compulsory Year of Child Care on Criminal Charges, Local Linear Regression, Different Bandwidths



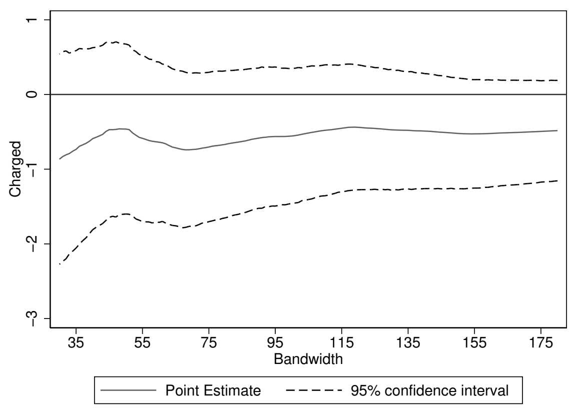
(A) All: Charged



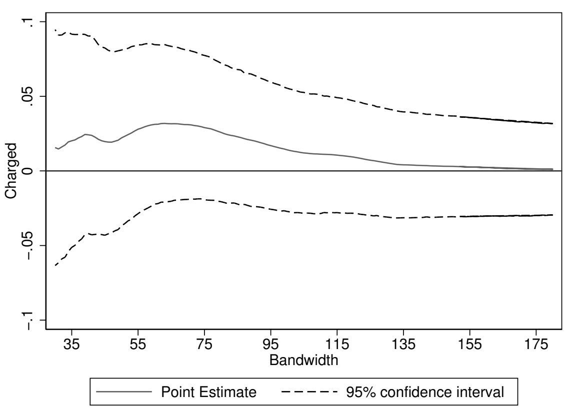
(B) All: Number of charges



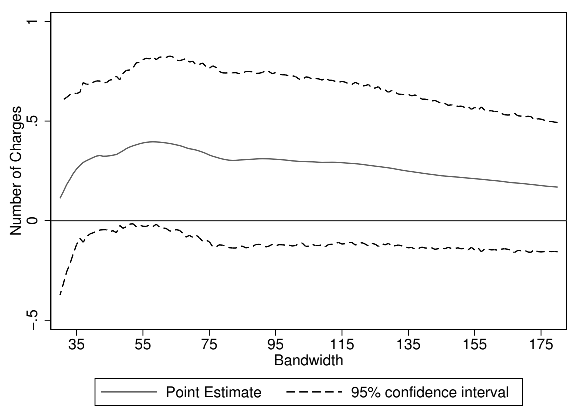
(C) Males: Charged



(D) Males: Number of charges



(E) Females: Charged



(F) Females: Number of charges

Notes: These figures display the β_5 coefficients from equation 2 along with 95% confidence intervals, using bandwidths between 30 and 180 days and local linear regression with triangular weights. The difference-in-regression-discontinuity (DiRD) estimates are the estimated difference in mean outcomes between those born just after December 31, 1990 and those born just before, relative to the same difference in the previous year. Outcomes are measured between 2008 and 2018. Panel (A) shows the likelihood of being charged during this time period, and panel (B) the total number of charges in this time period. Standard errors are robust and calculated in line with Calonico et al. (2014).

B.2 Tables

Table B.1: Effects of Compulsory Child Care for Six-Year-Olds on Number of Charges

	≥ 2	≥ 5	≥ 10
	(1)	(2)	(3)
Panel A: All			
DiRD	-0.0077**	-0.0040	-0.0011
	(0.0039)	(0.0025)	(0.0019)
Controls	Yes	Yes	Yes
Pre-reform mean, control	0.10	0.04	0.02
Observations	117152	117152	117152
Panel B: Males			
DiRD	-0.0120*	-0.0075	-0.0034
	(0.0068)	(0.0047)	(0.0032)
Controls	Yes	Yes	Yes
Pre-reform mean, control	0.17	0.07	0.03
Observations	60152	60152	60152
Panel C: Females			
DiRD	-0.0027	0.0002	0.0020*
	(0.0032)	(0.0017)	(0.0012)
Controls	Yes	Yes	Yes
Pre-reform mean, control	0.03	0.01	0.00
Observations	57000	57000	57000

Notes: This table displays the β_5 coefficient from equation (2), using linear trends, 180 days bandwidth and triangular weights (see section 6 for more information). The DiRD estimate is the difference in mean outcomes between those born just after December 31, 1990, relative to the same difference in the previous year. The outcomes are indicators for being charged at least 2, 5 and 10 times at any time between 2008 and 2018. Standard errors are clustered on the municipality level and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table B.2: Effects of Compulsory Child Care for Six-Year-Olds on Punishments

	Prison	Probation	Community service	Fine	Other
	(1)	(2)	(3)	(4)	(5)
Panel A: All					
DiRD	-0.0001	-0.0022	-0.0030**	-0.0081	0.0002
	(0.0025)	(0.0025)	(0.0014)	(0.0058)	(0.0001)
Controls	Yes	Yes	Yes	Yes	Yes
Pre-reform mean, control	0.03	0.03	0.02	0.39	0.00
Observations	117152	117152	117152	117152	117152
Panel B: Males					
DiRD	-0.0019	-0.0046	-0.0052*	-0.0157*	0.0001
	(0.0045)	(0.0043)	(0.0027)	(0.0081)	(0.0002)
Controls	Yes	Yes	Yes	Yes	Yes
Pre-reform mean, control	0.06	0.05	0.02	0.52	0.00
Observations	60152	60152	60152	60152	60152
Panel C: Females					
DiRD	0.0025	0.0002	-0.0004	-0.0011	0.0002*
	(0.0015)	(0.0018)	(0.0014)	(0.0078)	(0.0001)
Controls	Yes	Yes	Yes	Yes	Yes
Pre-reform mean, control	0.01	0.01	0.00	0.24	0.00
Observations	57000	57000	57000	57000	57000

Notes: This table displays the β_5 coefficient from equation (2), using linear trends, 180 days bandwidth and triangular weights (see section 6 for more information). The DiRD estimate is the difference in mean outcomes between those born just after December 31, 1990, relative to the same difference in the previous year. Outcomes are measured between 2008 and 2018. The outcomes are indicators for being convicted to a specific type of punishment at any time between 2008 and 2018. Standard errors are clustered on the municipality level and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table B.3: Effects of Compulsory Child Care for Six-Year-Olds on Criminal Charges, Different Trends, All

	Linear	Squared		Cubed	
	(1)	(2)	(3)	(4)	(5)
Panel A: Charged					
DiRD	-0.0060	-0.0053	-0.0065	-0.0053	-0.0068
	(0.0095)	(0.0052)	(0.0072)	(0.0052)	(0.0093)
Differencing trends	Yes	No	Yes	No	Yes
Observations	117152	117152	117152	117152	117152
Panel B: Number of charges					
DiRD	-0.1439	-0.1140*	-0.1332	-0.1138*	-0.1357
	(0.1099)	(0.0606)	(0.0836)	(0.0607)	(0.1091)
Differencing trends	Yes	No	Yes	No	Yes
Observations	117152	117152	117152	117152	117152

Notes: This table displays the β_5 coefficient from equation (2) for the full sample, using 180 days bandwidth and triangular weights (see section 6 for more information). In column (1) linear trends are included, in columns (2) and (3) quadratic trends and in columns (4) and (5) cubic trends. In columns (2) and (4), trends are equal for both the treatment and control group on each side of the cutoff, while in columns (1), (3) and (5) trends are allowed to differ between the treatment and control group. The DiRD estimate is the difference in mean outcomes between those born just after December 31, 1990, relative to the same difference in the previous year. Outcomes are measured between 2008 and 2018. Columns (1) and (2) show the likelihood of being charged during this time period, and columns (3) and (4) the total number of charges in this time period. Standard errors are clustered on the municipality level and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table B.4: Effects of Compulsory Child Care for Six-Year-Olds on Criminal Charges, Different Trends, Males

	Linear	Squared		Cubed	
	(1)	(2)	(3)	(4)	(5)
Panel A: Charged					
DiRD	-0.0176	-0.0110	-0.0169	-0.0110	-0.0150
	(0.0156)	(0.0086)	(0.0117)	(0.0086)	(0.0154)
Differencing trends	Yes	No	Yes	No	Yes
Observations	60152	60152	60152	60152	60152
Panel B: Number of charges					
DiRD	-0.3679*	-0.2562**	-0.3127**	-0.2558**	-0.3407*
	(0.2006)	(0.1129)	(0.1561)	(0.1131)	(0.1970)
Differencing trends	Yes	No	Yes	No	Yes
Observations	60152	60152	60152	60152	60152

Notes: This table displays the β_5 coefficient from the equation (2) for males, using 180 days bandwidth and triangular weights (see section 6 for more information). In column (1) linear trends are included, in columns (2) and (3) quadratic trends and in columns (4) and (5) cubic trends. In columns (2) and (4), trends are equal for both the treatment and control group on each side of the cutoff, while in columns (1), (3) and (5) trends are allowed to differ between the treatment and control group. The DiRD estimate is the difference in mean outcomes between those born just after 31 December 1990, relative to the same difference in the previous year. Outcomes are measured between 2008 and 2018. Columns (1) and (2) show the likelihood of being charged during this time period, and columns (3) and (4) the total number of charges in this time period. Standard errors are clustered on the municipality level and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table B.5: Effects of Compulsory Child Care for Six-Year-Olds on Criminal Charges, Different Trends, Females

	Linear	Squared		Cubed	
	(1)	(2)	(3)	(4)	(5)
Panel A: Charged					
DiRD	0.0067	0.0012	0.0052	0.0012	0.0017
	(0.0096)	(0.0051)	(0.0071)	(0.0051)	(0.0092)
Differencing Trends	Yes	No	Yes	No	Yes
Observations	57000	57000	57000	57000	57000
Panel B: Number of charges					
DiRD	0.1108	0.0511	0.0757	0.0511	0.0932
	(0.0995)	(0.0470)	(0.0692)	(0.0469)	(0.0984)
Differencing trends	Yes	No	Yes	No	Yes
Observations	57000	57000	57000	57000	57000

Notes: This table displays the β_5 coefficient from the equation (2) for females, using 180 days bandwidth and triangular weights (see section 6 for more information). In column (1) linear trends are included, in columns (2) and (3) quadratic trends and in columns (4) and (5) cubic trends. In columns (2) and (4), trends are equal for both the treatment and control group on each side of the cutoff, while in columns (1), (3) and (5) trends are allowed to differ between the treatment and control group. The DiRD estimate is the difference in mean outcomes between those born just after December 31, 1990, relative to the same difference in the previous year. Outcomes are measured between 2008 and 2018. Columns (1) and (2) show the likelihood of being charged during this time period, and columns (3) and (4) the total number of charges in this time period. Standard errors are clustered on the municipality level and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table B.6: Effects of Compulsory Child Care for Six-Year-Olds on Criminal Charges, Ages 18-25

	Charged		Number of charges	
	(1)	(2)	(3)	(4)
Panel A: All				
DiRD	-0.0208***	-0.0204***	-0.2008***	-0.1928***
	(0.0053)	(0.0051)	(0.0541)	(0.0529)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.18	0.18	0.73	0.73
Observations	117152	117152	117152	117152
Panel B: Males				
DiRD	-0.0293***	-0.0266***	-0.4000***	-0.3713***
	(0.0085)	(0.0082)	(0.0990)	(0.0977)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.27	0.27	1.22	1.22
Observations	60152	60152	60152	60152
Panel C: Females				
DiRD	-0.0129***	-0.0130***	0.0043	0.0106
	(0.0048)	(0.0048)	(0.0404)	(0.0448)
Controls	No	Yes	No	Yes
Pre-reform mean, control	0.07	0.07	0.21	0.21
Observations	57000	57000	57000	57000

Notes: This table displays the β_5 coefficient from equation (2), using linear trends, 180 days bandwidth and triangular weights (see section 6 for more information). The DiRD estimate is the difference in mean outcomes between those born just after December 31, 1990, relative to the same difference in the previous year. Outcomes are measured between ages 18-25. Columns (1) and (2) show the likelihood of being charged during this time period, and columns (3) and (4) the total number of charges in this time period. Standard errors are clustered on the municipality level and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

C Cash-for-Care

C.1 Tables

Table C.1: Effects of Cash-for-Care on Number of Charges, ITT

	≥ 2	≥ 5	≥ 10
	(1)	(2)	(3)
Panel A: All			
DiD	0.0024***	0.0015***	0.0006***
	(0.0004)	(0.0002)	(0.0001)
Controls	Yes	Yes	Yes
Pre-reform mean, 18-year-olds	0.01	0.00	0.00
Observations	2943581	2943581	2943581
Panel B: Males			
DiD	0.0049***	0.0025***	0.0011***
	(0.0006)	(0.0003)	(0.0001)
Controls	Yes	Yes	Yes
Pre-reform mean, 18-year-olds	0.02	0.00	0.00
Observations	1514156	1514156	1514156
Panel C: Females			
DiD	-0.0001	0.0005***	0.0002***
	(0.0003)	(0.0001)	(0.0001)
Controls	Yes	Yes	Yes
Pre-reform mean, 18-year-olds	0.00	0.00	0.00
Observations	1429425	1429425	1429425

Notes: This table displays the β_4 coefficient from equation (3), estimated separately for each quartile of fathers' income. The DiD estimate is the difference in mean outcomes in between 18-year-olds and 23-27-year-olds in 2016-2018, relative to the same difference in 2010-2013. Estimates are intention-to-treat. The outcomes are indicators for being charged at least 2, 5 and 10 times in a given year. Standard errors are robust and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table C.2: Effects of Cash-for-Care on Punishments, ITT

	Prison	Probation	Community service	Fine	Other
	(1)	(2)	(3)	(4)	(5)
Panel A: All					
DiD	0.0023***	0.0005***	-0.0004***	0.0133***	-0.0000
	(0.0001)	(0.0002)	(0.0001)	(0.0007)	(0.0000)
Controls	Yes	Yes	Yes	Yes	Yes
Pre-reform mean, 18-year-olds	0.00	0.00	0.00	0.04	0.00
Observations	2943581	2943581	2943581	2943581	2943581
Panel B: Males					
DiD	0.0041***	0.0008***	-0.0006**	0.0219***	-0.0000
	(0.0002)	(0.0003)	(0.0002)	(0.0012)	(0.0000)
Controls	Yes	Yes	Yes	Yes	Yes
Pre-reform mean, 18-year-olds	0.00	0.00	0.00	0.06	0.00
Observations	1514156	1514156	1514156	1514156	1514156
Panel C: Females					
DiD	0.0003***	0.0002	-0.0002**	0.0043***	0.0000
	(0.0001)	(0.0001)	(0.0001)	(0.0007)	(0.0000)
Controls	Yes	Yes	Yes	Yes	Yes
Pre-reform mean, 18-year-olds	0.00	0.00	0.00	0.02	0.00
Observations	1429425	1429425	1429425	1429425	1429425

Notes: This table displays the β_4 coefficient from equation (3). The DiD estimate is the difference in mean outcomes in between 18-year-olds and 23-27-year-olds in 2016-2018, relative to the same difference in 2010-2013. Estimates are intention-to-treat. The outcomes are indicators for being sentenced to a specific type of punishment in a given year. Standard errors are robust and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table C.3: Effects of Cash-for-Care on Criminal Charges, ITT, Phase-in

	Charged		Number of charges	
	(1)	(2)	(3)	(4)
Panel A: All				
DiD Phase-in	-0.0005	0.0010	0.0010	0.0070***
	(0.0007)	(0.0007)	(0.0024)	(0.0025)
Controls	No	Yes	No	Yes
Pre-reform mean, 18-year-olds	0.04	0.04	0.07	0.07
Observations	2943581	2943581	2943581	2943581
Panel B: Males				
DiD Phase-in	0.0002	0.0029**	0.0038	0.0146***
	(0.0012)	(0.0012)	(0.0044)	(0.0045)
Controls	No	Yes	No	Yes
Pre-reform mean, 18-year-olds	0.06	0.06	0.11	0.11
Observations	1514156	1514156	1514156	1514156
Panel C: Females				
DiD Phase-in	-0.0012*	-0.0007	-0.0018	-0.0007
	(0.0007)	(0.0007)	(0.0016)	(0.0016)
Controls	No	Yes	No	Yes
Pre-reform mean, 18-year-olds	0.02	0.02	0.02	0.02
Observations	1429425	1429425	1429425	1429425

Notes: This table displays the β_3 coefficient from equation (3). The DiD estimate is the difference in mean outcomes in between 18-year-olds and 23-27-year-olds in 2014-2015, relative to the same difference in 2010-2013. Results for post years 2016-2018 from the same estimations are shown in table 9. Estimates are intention-to-treat. Outcomes are measured yearly. Columns (1) and (2) show the likelihood of being charged in given year, and columns (3) and (4) the yearly total number of charges. Standard errors are robust and reported in parentheses.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Chapter 2

Surviving a Mass Shooting

Surviving a Mass Shooting

Prashant Bharadwaj[*] Manudeep Bhuller[†] Katrine V. Løken[‡] Mirjam Wentzel[§]

Abstract

We use data on all middle and high school aged children who survived a mass shooting incident on July 22, 2011 in Utøya, Norway, to understand how such events affect survivors, their families, and their peers. Using a difference-in-differences design to compare survivors to a matched control group, we find that in the short run children who survive have substantially lower GPA (nearly 0.5 SD) and increased utilization of health care services and more mental health diagnoses. In the medium run, survivors have fewer years of schooling completed and lower labor force participation. Parents and siblings of survivors are also impacted, experiencing substantial increases in doctor visits and mental health diagnoses. However, there appear to be limited impacts on school aged peers of survivors. While this event affected the entire country, we show that survivors and their families bear significant costs despite robust social safety nets and universal access to healthcare.

Keywords: mass shooting, test scores, labor markets
JEL codes: J13, O15, K14

Acknowledgments: This research was conducted as part of the Starting Grant project 'Criminality, Victimization and Social Interactions' (CIVICS 757279) funded by the European Research Council. We thank seminar participants at UC San Diego, NBER SI 2020 (Children's Program), Stanford, and UT Austin for helpful comments. The paper substantially improved as a result of conversations with Eli Berman, Marika Cabral, Petra Persson, Maya Rossin- Slater, and Agne Suziedelyte.

[*]University of California San Diego and NBER. E-mail: prbharadwaj@ucsd.edu
[†]University of Oslo, Statistics Norway, IZA, and CESifo. E-mail: manudeep.bhuller@econ.uio.no
[‡]Norwegian School of Economics, Statistics Norway, IZA and CESifo. E-mail: katrine.loken@nhh.no
[§]Norwegian School of Economics & Statistics Norway. E-mail: mirjam.wentzel@nhh.no

1 Introduction

Exposure to childhood trauma is widely recognized as an important factor determining long run labor market outcomes. While children are subjected to various forms of trauma, ranging from health shocks to parental death and financial distress, the impacts of childhood exposure to violence and gun violence in particular has been on the forefront of media and policy attention.[1] This is in part due to well publicized mass shootings and school shootings in the US, but also increased attention towards gun violence in other countries Sturup, Rostami, Mondani, Gerell, Sarnecki, and Edling (2019). The broad public interest in children exposed to such events could indicate that the costs of victimization for this group are particularly large Cook and Ludwig (2002). However, empirical evidence on such costs is limited, primarily as a consequence of data availability and challenges in linking pre- and post event outcomes to individual victims.

In this paper we quantify the effects of surviving an episode of mass shooting on school aged children in Norway in the short and medium term. In addition, we document spillover effects of surviving such events on children's families and their school aged peers. The empirical challenge in such an exercise stems from the usual concern that children who are exposed to such events are different from those who are not, and could have fared differently even absent this exposure. In our setting, the context of mass shooting is one where it is unlikely that individual children were targeted for unobserved reasons, and we also leverage detailed administrative data and a matching strategy to identify “similar" children who were not directly exposed to the event in a difference-in-differences (DiD) design.

On July 22, 2011 around 600 individuals, mostly school aged children from all over Norway, were attending a youth camp on the island of Utøya when a politically motivated mass shooting took place. Using detailed administrative data on individual demographics and parental characteristics (including earnings), we match survivors to observably similar children who were not on the island and who also did not attend the same school as the affected children. These matched children form the control group for the affected survivors, which allows us to compare outcomes before and after the event across survivors and controls in a DiD design. With this design, we can account for time-invariant unobserved heterogeneity across survivors and matched controls as well as common changes over time. Under a standard parallel trends assumption, this allows us to estimate the effects of exposure to mass shooting on the survivors.

We consider a series of different outcomes for survivors, their families, and their peers to draw a comprehensive picture of the impacts of mass shooting exposure. For school outcomes (e.g., GPA), we compare survivors and controls who took the same grade exams before and after the event, effectively relying on variation across different birth cohorts in the timing of event. For

[1] For an excellent recent review on the costs of victimization see Bindler, Ketel, and Hjalmarsson (2020).

survivors' health outcomes (e.g., doctor visits) and their parental outcomes, we can follow all survivors and controls both before and after the event, and rely on a DiD design with individual fixed effects. While for survivors' medium-term outcomes (e.g., employment and earnings), we do not observe pre-event outcomes for most individuals, and rely therefore on a simpler treatment-control comparison based on matching techniques.

We find that survivors aged 14-15 at the time of event obtained lower test scores (by 0.5 SD) in middle school, and survivors aged 14-18 were less likely to finish high school on time (by 21 percentage points) after the event. Regardless of the age at exposure, survivors had increased utilization of health care services, with more medical general practitioner visits (60% increase relative to control mean) and more psychological diagnoses (400% increase relative to control mean) in the aftermath of the shooting. By 2018, approximately seven years after the incident, survivors were 14% less likely to have attained a high school diploma and 15% less likely to have completed any college education. Survivors were also 5 percentage points less likely to be employed in 2018, and had 13% lower earnings (not statistically significant) conditional on being employed. Parents of survivors were not meaningfully impacted in terms of employment or earnings, although they do see significant increases in doctor visits and mental health diagnoses. The same is true for survivors' siblings. We also find that siblings suffer in terms of test scores in middle school by scoring nearly 0.2 SD lower. We find no economically meaningful or statistically significant effects on schooling or health outcomes among school peers.

In writing this paper we contribute to a rich literature on the short and long run consequences of exposure to various types of violence during childhood across multiple disciplines.[2] Some of this work also focuses on the role of acts of terror (a term also used to describe the Utøya incident) on student test scores Shany (2016); Auger, Seymour, and Roberts Jr (2004) and mental health Otto, Henin, Hirshfeld-Becker, Pollack, Biederman, and Rosenbaum (2007). More specifically, our paper adds to some of the recent work in economics focusing on gun violence and its effects on children. To provide a few key examples: Ang (2021) focuses on the effects on African American students who are exposed to police shootings in their neighborhood, Rossin-Slater, Schnell, Schwandt, Trejo, and Uniat (2020) focus on anti-depressant use by students after fatal school shootings, Koppensteiner and Menezes (2021) explore the effects of homicides on educational performance and human capital investments among students in Brazil, and Gershenson and Tekin (2018) examine the effects on school performance of students exposed to the "Beltway Sniper" in 2002 near Washington DC.

Three contemporaneous papers are closely related to this paper. First is the important work by Levine and McKnight (2020) who examine the consequences of being exposed to school shootings

[2] Some of the papers focusing specifically on trauma due to violence include Finkelhor, Turner, Shattuck, Hamby, and Kracke (2015), Moffitt (2013), Sharkey (2010) and Hurt, Malmud, Brodsky, and Giannetta (2001).

on test scores and well being of students in the US. Using several quasi-experimental techniques and first focusing on the Sandy Hook incident in Newtown, CT in 2012, they find that average test scores in schools are lower and chronic absenteeism higher after this event. They also find that these effects spillover to nearby schools. They next examine the Columbine shooting incident from 1999 in Columbine, CO and offer evidence that is consistent with the idea that affected students were more likely to commit suicide in subsequent years. Second is the excellent paper by Cabral, Kim, Rossin-Slater, Schnell, and Schwandt (2021) where they examine using a matched DiD strategy (where *schools* are matched as opposed to individuals as in our case) the impact of being exposed to school shootings in Texas between 1995-2016. They document adverse short term impacts on absence, high school graduation, and early labor market outcomes. Third is the work by Deb and Gangaram (2021) where they identify long run impacts on risky behavior, declines in health and well being, and worse labor market outcomes for children exposed to school shootings relative to their unexposed peers. We build on such findings by leveraging individual level data where we can follow students over time in terms of test scores and health visits, and also examine spillovers more precisely by examining family members and peers individually.

Finally, this paper provides additional evidence specifically on the impacts of the Utøya mass shooting. To highlight just a few: Alne and Serdarevic (2020) examine the impacts of losing a child during this event on parental labor market outcomes, Stene and Dyb (2015) and Dyb, Jensen, Glad, Nygaard, and Thoresen (2014) provide descriptive accounts of healthcare utilization by the survivors of the attacks, Gjerland, Pedersen, Ekeberg, and Skogstad (2015) examine health impacts on rescue workers, and Hernæs (2021) examines sickness absence responses in the short run at the municipality level. By examining the effects on test scores, high school completion, labor force partition, and earnings, and also by considering family and peer effects, we add meaningfully to this important area of research.

2 The July 22, 2011 incident in Utøya

On July 22, 2011, Anders Behring Breivik committed two acts of domestic terrorism in Norway. The first was a car bomb explosion in the capital Oslo, and the second, a few hours later at a summer camp on the island of Utøya. The camp was organized by AUF, a youth wing of the ruling (at the time) Norwegian Labor Party (AP). During the hour or so long period of indiscriminate killing, Breivik killed 69 people. A majority of those killed were between the ages of 16-18; the youngest victims were 14 years old, and the oldest was 51. Among the survivors, 66 were shot and wounded, while 585 survived without physical injuries (Helsedirektoratet, 2012).[3] Since this was a youth

[3]Given the small sample size of children who faced direct physical injury as a result of this incident, we are unable to fully distinguish victims with physical and non-physical injuries for ethical and statistical reasons.

camp, a large fraction of the survivors were middle and high school aged children, who came from all over Norway.

There was a massive response from the government following this attack. An independent commission examined the police responses and preparedness for these types of threats (see NOU (2012:14)). In terms of treatment of the victims and their families, since 2003, Norway has had a national scheme for compensating victims of criminal injuries. Survivors as well as the bereaved family members of those murdered at Utøya received compensation through this scheme, and the amounts paid out were typically higher than the compensation given to other victims of violent crime Nilsen, Langballe, and Dyb (2016).[4] There was also an increased allocation of resources to municipalities with survivors to take care of those that needed more follow-up services and additional expenses for health specialists Helsedirektoratet (2012). These payments and resources are in addition to what is generally provided under a generous welfare state with universal health care and insurance schemes (such as sickness benefits and disability benefits) to help those in need.

3 Data

We link several administrative data sources from Norway for this paper. These include the victim database from the Norwegian Police Directorate and population level panel data drawn from school, health, and tax registers.

3.1 School data

School registers contain data on all middle and high school students. For the analysis concerning younger survivors we use the 10th grade data from middle school and this data is available from 2005-2016. We have two measures of performance in middle school: grade point average (GPA) and national exam scores. Given compulsory schooling laws, all individuals are required to be enrolled in 10th grade in Norway, and as a consequence, we observe GPA for nearly every resident child aged 16 between 2005-2016. This GPA is the average of the grades in all subjects. Grades are normalized for each academic year-cohort separately. To identify peers, we consider pupils of the same age who attended the same middle school in grade 10. Nearly all schools in Norway are public and school switching is uncommon. Exam scores are from national exams, where everyone takes the same exam and is externally graded. The scores are standardized for each academic

[4] The Norwegian victim support scheme is targeted at the victims of violence (including victims of sexual assault and abuse) and aims at covering the loss of income, besides covering extraordinary medical expenses that result from the victimization, and providing redress and injury compensation. To receive compensation, a victim must (a) file a police report concerning the incident and (b) submit a written application to the National Office for Victim Compensation. Subject to evaluation and approval, the Office will compensate the victims for their losses.

year-cohort separately.

For high school students we use enrollment and completion data, which we have from 2002 to 2018. We do not use data on high school GPA as attending high school is a choice in Norway, and part of the effect on survivors could be that they drop out of high school. We therefore choose to focus on completion of grade 1 and grade 3 in high school on time and on not getting a high school diploma by 2018, i.e., approximately seven year after the event. Academic track programs in high school are 3 years, while vocational track programs can be longer. For peers, we can only study the completion of grade 3 in high school. The reason is that we need to define peers in high school before July 22, 2011, as the victims might drop out or change high school after the event, making the peers endogenous. Peers are therefore defined in grade 1 in high school, and only for those starting high school before July 22, 2011.

For higher education data, we use enrollment and completion data, available from 2002 to 2018, and information on the highest level of education completed, available up to 2018. We create an indicator equal to one if the individual is registered as enrolled in any college or university level education in any year between 2012 and 2018 and we also create an indicator equal to one if the individual has completed any college or university level education by 2018.

3.2 Health data

To analyze health outcomes, we use the Control and Payment of Reimbursement to Health Service Providers (KUHR) database, available from 2006-2018, with information on visits to general practitioner (GP), emergency units (ER), and referrals to specialists. All Norwegians belong to the patient list of a certified GP. The register includes information on each patient's unique personal identifier, date and time of contact, and diagnosis according to the ICPC-2-diagnosis codes. For GP visits, we sum all GP consultations in each calendar year for survivors, and in each academic year (August-June) for peers. For psychological diagnosis we create an indicator equal to one, for each calendar year, if the individual has any psychological diagnoses in the ICPC-2 (P01-P99) and ICD-10 diagnosis codes in that year. For peers, the indicator is equal to one if there is a psychological diagnosis during the academic year.

3.3 Labor market data

To analyze labor market outcomes for parents, we use data from tax registers and matched employer-employee records, available from 1987-2018. Earnings include labor earnings, sickness benefits and parental leave benefits, but not e.g. unemployment benefits, disability insurance, pension income or entrepreneurial income. We create an employment indicator, being equal to one if earnings are positive. For earnings, we use the log of earnings and drop observations with zero earnings.

The earnings measures are CPI adjusted with 2015 as the base year. We also include a dummy equal to one if parents are married as of January 1st in a given year. Another important outcome that we analyze for parents is their sickness leave. We create an indicator equal to 1 if an individual has a physician-certified sickness leave in a given year.[5]

3.4 Victim data

Our victim data come from an administrative police register, which comprises all crimes reported to the Norwegian police, including the Utøya incident. This database contains information on the month and place of the crime, which makes it possible for us to identify survivors of Utøya; see Figure 1, panel (a). Using this database, we can identify individuals who became victimized in July 2011 in the municipality of Hole (where Utøya is located). A person is defined as a survivor if the registered crime is attempted murder (*drapsforsøk*), and not murder (*drap*). A few of the registered victims were not registered residents in Norway, and hence they are excluded from our analytical sample. Figure 1, panels (b) and (c) show that most of the survivors were school aged children and came from municipalities all over Norway. Tables B1-B2 in the Online Appendix show additional descriptives on survivors.

3.5 Matching

We match a survivor to controls who were not directly affected using Coarsened Exact Matching (CEM) based on the protocols described as follows. To be a match, the control individual must have the exact same value on all the following variables: birth year, gender, mother's and father's education measured in 3 levels in 2010 (completed middle school or less, completed at least some high school, and completed at least some university education), mother's and father's age in 2010 in tertiles (for each birth cohort), mother and father's earnings in 2010 in tertiles (for each birth cohort), and the centrality of municipality of residence, where we use Statistics Norway's centrality classes (6 levels) to distinguish areas with high or low population density. We also include categories for missing data for parental education, age, and earnings.

To reduce the risk of matching survivors to controls that knew a survivor or a deceased victim and thus may also have been partially impacted ("contaminated"), we further remove potential controls living in the same municipality or city district and born in the same year as a survivor or a deceased victim. To identify residence of the survivor and the matched control, we use city districts for the four largest cities in Norway, Oslo, Bergen, Trondheim, and Stavanger and otherwise use the

[5]In Norway, sick leave from work for 3 days or less is self-certified. Some workplaces have an "inclusive work life" agreement, which allows employees to take self-certified absences for up to 8 days. Additionally, parents can take up to 10 days of "sick child leave" per year to take care of sick children or children with special needs. Neither self-certified sick leave nor "sick child leave" incidences are included in the parental sickness leave measure.

municipality of residence on January 1, 2011. The CEM procedure thus ensures that a survivor is matched to controls born the same year who also share other background characteristics, yet reside in a different municipality or city district than the survivor.

While we exclude the same aged school peers of survivors or deceased victims from the pool of potential controls, a survivor could still be matched to controls residing in a municipality or city district where other survivors of a different age also resided. Thus, while our matching approach can tackle concerns about controls being impacted by directly affected same aged survivors, this does not accommodate unrestricted local spillovers across cohorts or broader population-wide influences.[6] In the analysis of spillover effects on school peers, we accordingly compare the outcomes of survivors' same aged school peers to the outcomes of their matched controls' same aged school peers.

For medium-term outcomes, we follow the matching procedure as described above, but further match on 10th grade exam score quintiles (for each cohort). We only include matched individuals taking the 10th grade exam before July 22, 2011, i.e., cohorts born in 1986-1995. Cohorts born before 1986 are excluded, as middle school data is only available from 2002.

4 Estimation

We use a difference-in-differences (DiD) approach with the following main specification:

$$Y_{it} = \beta Surv_i + \gamma Post_t + \eta Surv_i \times Post_t + \lambda X_{it} + \varepsilon_{it}$$

where Y_{it} is the outcome for individual i in period t. $Surv_i$ is an indicator equal to one if the individual is a survivor, or a family member or a peer of the survivor, and zero if the individual is in the control group. For school outcomes, $Post_t$ is an indicator equal to one if individual i is from a birth cohort that completed the grade pertaining to Y_{it} in the 2011/12 academic year or later, and zero otherwise. For health outcomes of survivors, as well as parental outcomes, $Post_t$ is simply an indicator for whether the outcome is measured in year 2012 or later. And X_{it} is a set of controls, including gender, parental education and age, and the centrality of municipality.

For school outcomes, we only observe each individual once in each regression. The DiD estimate is then the difference between survivors and control individuals that "should" be finishing the grade after July 22, 2011, minus the equivalent for survivors and control individuals that "should"

[6] In practice, concerns about school-level spillovers across different cohorts are relevant for a small fraction of cases. For instance, in the analysis of middle school outcomes, only 7.8% of matched controls who finished middle school after July 2011 went to a school where there was another survivor in the three adjacent cohorts. Excluding these observations does not materially affect our results for middle school outcomes (available upon request). An alternative would be to match survivors to controls who went to a school with no survivors or deceased victims.

have finished the grade before July 22, 2011. The main identifying assumption is that given the matching, the counterfactual cohort profile for the survivors (or their families or peers) in the absence of Utøya would have been identical to the one for the matched control group. This approach allows us to account for secular changes across birth cohort that are unrelated to the event. We follow this approach also for younger siblings and peers of survivors. And for peers, we use the same approach for health outcomes.

For health outcomes of survivors, as well as parental outcomes, we can observe individuals over time. For these outcomes, we can thus include individual fixed effects, and exploit within-person changes in outcomes before and after the event across survivors and controls. The DiD estimate is then the difference in outcomes of all survivors (or parents of survivors) and all control individuals from 2011 and onwards, minus the equivalent difference in the years before.

For middle school outcomes, as well as high school outcomes and health outcomes for peers, we cluster standard errors at the school-cohort level. For health outcomes of survivors we use robust standard errors. For family outcomes we cluster standard errors on the mother's personal identifier.

For medium-term outcomes, we cannot use the difference-in-differences approach, as we do not have a pre-period. As mentioned in Section 3.5, we match on middle school 10th grade exam scores when we consider medium-term outcomes. This is because we observed level differences between survivors and control individuals in exam scores (and GPA), which may correlate with or drive differences in future outcomes. We hope to reduce this by also matching on exam scores. To this end, a simple treatment-control comparison is performed:

$$Y_i = \beta Surv_i + \lambda X_i + \varepsilon_i$$

The control observations are weighted in the estimations using the following formula:

$$W = \frac{m_C}{m_T} \times \frac{m_T^s}{m_C^s}$$

where W is the weight for each individual, m_C and m_T are the number of controls and survivors in the estimation sample, and m_C^s and m_T^s are the number of controls and survivors in the same stratum, i.e., with the same coarsened characteristics in the matching procedure. In the peer analysis, each survivor is matched to one focal control, so we perform unweighted estimations.

5 Results

Figure 2 graphically presents the DiD estimates of the effects for survivors. Each dot is the difference between the survivor mean outcome and the matched control mean outcome at each post and pre event year, minus the equivalent in the (academic) year before the event. The vertical bars next to each dot provide the 95% confidence intervals.

Figure 2, panels (a) and (b) plot the effects on middle school exam scores and GPA relative to the academic year 2010/11 (for brevity, referred to as year "2011" in the figure and in the following), which ended around a month before the Utøya event on July 22, 2011. We see no significant differences in the years prior to 2011, while there is a drop in 2012 and 2013. Table 1 provides the treatment effect estimates, with standard errors in parenthesis, and shows a drop of around 0.3 of a standard deviation on exam scores (column 1), and 0.5 SD decrease in GPA (column 2), where only the latter is significant at the 1% level.

Figure 2, panels (c) and (d) show similar patterns for the completion of 1st and 3rd grade in high school on time. Table 1 confirms that these are significant impacts, with 19 and 21 percentage points (columns 3 and 4) reductions in the probabilities of finishing the first and third year in high school on time, respectively. For individuals that were not old enough to have finished high school before the event, i.e., cohorts 1993-1995, we find a 13 percentage points (column 7) reduction in the probability of ever receiving a high school diploma by 2018, compared to a matched control group. For the same individuals, we also find a 7 percentage points (column 8) reduction in the probability of ever enrolling in any higher education by 2018. Considering all cohorts of survivors, we find a 6 percentage points (column 9) reduction in the probability of completing any higher education by 2018, compared to a matched control group.

Next, in Figure 2, panels (e) and (f), we see large increases in the utilization of health services for survivors (represented by round dots). There are large and immediate impacts on the number of GP visits and on the incidence of psychological diagnoses in the year of event. More strikingly, survivors also have a persistently higher utilization of health services over the following five year period after 2011. As seen in Table 1, columns (5)-(6), the number of GP visits increase by 0.64 (from a base of about 1) and a large portion of this is related to the incidence of psychological diagnoses, which increases by 36 percentage points from a base of around 10%. Finally, we consider the medium-run labor market impacts on survivors, estimating that survivors had 5 percentage points lower probability of being employed in 2018, compared to a matched control group. We also find that survivors had 13% lower labor market earnings conditional on being employed, although this effect is not significant.

Turning to the spillovers on family members, Table 2 summarizes the main estimates for mothers (panel A), fathers (panel B) and siblings (panel C). We first consider impacts on mothers' and fathers' labor force participation, log-earnings, the incidence of any sickness absence from work, and marital status. Only the impact on mothers' sickness absence is significant and economically meaningful; we find that mothers increase sickness leave take up by around 28%. This effect is mirrored when we examine doctor visits (column 5) and mental health diagnoses (column 6) for mothers. Figure 2, panel (f), further shows that the impacts on the incidence of psychological diagnoses are sizeable for both mothers (squares) and fathers (triangles), and these impacts persist during the five year period after the event.[7] Table 2, panel C, considers the impacts on the survivors' siblings. Siblings also see lower test scores (columns (7)-(8)) and increases in mental health diagnoses (column 10). While the impacts on siblings are substantial, with 0.22 SD lower middle school GPA and exam scores and around 25% increase in mental health diagnoses, these effect magnitudes are smaller than for what we find for survivors.

Finally, we study the impacts on same aged school mates of the survivors. Table 3 considers GPA and exam scores for middle school peers, completion of 3rd year for high school peers and GP visits and psychological diagnosis for both middle and high school peers. Panel A considers peers of either gender, while panels B-C show impacts by peer gender. While there is a significant negative impact on the test scores of female peers and a small increase in mental health diagnoses for female peers, since these are the only statistically significant effects across a range of peer outcomes, we do not wish to place undue focus on these effects. Our reading of Table 3 suggests that there are largely muted effects on the survivors' same aged school peers.

6 Conclusion

This paper shows that children who survive gun violence face tremendous costs in terms of school performance and mental health in the short run. In the medium run, survivors have lower years of education and worse labor market outcomes. Parents and siblings of survivors are also impacted in terms of mental health and school test scores.

While this setting allows us to examine the effects of a mass shooting event on survivors and those around them, we want to be upfront that this was a unique event (like most mass shootings) and there are challenges in drawing broad conclusions from this study that might apply to the case of school shootings or other forms of gun violence in the United States and elsewhere. The event we study was one of the deadliest gun attacks on Norwegian soil and the government response to victims, survivors, and their families was tremendous. The attack was also not on school premises as most mass shootings that involve children typically are, which means schools were not directly affected and there was no resorting of students and resources. Norway is also a setting where gun violence and mass shootings of this kind is extremely rare. While some of these factors (social support and government assistance) imply that the results we find are *net* of such compensatory actions, other factors such as the rarity of this event in this context might imply that in places where gun violence is more common, student, peer, and family reactions might be more muted. Finally,

[7] For brevity, we don't show the confidence intervals for the family members' effect estimates in Figure 2, panels (e)-(f). These can be found in Figures B1 (mothers), B2 (fathers) and B3 (siblings) in the Online Appendix.

our design relies on comparing those who are directly affected to a "control" group who might also be impacted by this event. Ultimately the empirical design only allows us to pick up *relative* effects and does not capture broad, overall impacts on the population. These considerations should be taken into account when thinking about how these results help make sense of similar events in different contexts.

References

ALNE, R., AND N. SERDAREVIC (2020): "The Cost of Terrorism: Network Effects and the Economic Impact of Child Loss," *Available at SSRN 3517441*.

ANG, D. (2021): "The effects of police violence on inner-city students," *Quarterly Journal of Economics*, 136(1), 115–168.

AUGER, R. W., J. W. SEYMOUR, AND W. B. ROBERTS JR (2004): "Responding to terror: The impact of September 11 on K-12 schools and schools' responses," *Professional School Counseling*, pp. 222–230.

BINDLER, A., N. KETEL, AND R. HJALMARSSON (2020): "Costs of Victimization," *Handbook of Labor, Human Resources and Population Economics*, pp. 1–31.

CABRAL, M., B. KIM, M. ROSSIN-SLATER, M. SCHNELL, AND H. SCHWANDT (2021): "Trauma at School: The Impacts of Shootings on Students' Human Capital and Economic Outcomes," Discussion paper, National Bureau of Economic Research.

COOK, P. J., AND J. LUDWIG (2002): "The costs of gun violence against children," *The Future of Children*, pp. 87–99.

DEB, P., AND A. GANGARAM (2021): "Effects of School Shootings on Risky Behavior, Health and Human Capital," Discussion paper, National Bureau of Economic Research.

DYB, G., T. JENSEN, K. A. GLAD, E. NYGAARD, AND S. THORESEN (2014): "Early outreach to survivors of the shootings in Norway on the 22nd of July 2011," *European Journal of Psychotraumatology*, 5(23523), 1–9.

FINKELHOR, D., H. TURNER, A. SHATTUCK, S. HAMBY, AND K. KRACKE (2015): *Children's exposure to violence, crime, and abuse: An update*. Citeseer.

GERSHENSON, S., AND E. TEKIN (2018): "The effect of community traumatic events on student achievement: Evidence from the beltway sniper attacks," *Education Finance and Policy*, 13(4), 513–544.

GJERLAND, A., M. J. B. PEDERSEN, Ø. EKEBERG, AND L. SKOGSTAD (2015): "Sick-leave and help seeking among rescue workers after the terror attacks in Norway, 2011," *International Journal of Emergency Medicine*, 8(1), 31.

HELSEDIREKTORATET (2012): *Læring for bedre beredskap – Helseinnsatsen etter terrorhendelsene 22. juli 2011 [Lessons for Improved Preparedness – Health Efforts After the Acts of Terror on July 22, 2011]*. Norwegian Directorate of Health (In Norwegian).

HERNÆS, Ø. (2021): "Going through Hell: Increased Work Effort in the Aftermath of Terrorism in Norway," *Scandinavian Journal of Economics*, 123(1), 216–237.

HURT, H., E. MALMUD, N. L. BRODSKY, AND J. GIANNETTA (2001): "Exposure to violence: Psychological and academic correlates in child witnesses," *Archives of Pediatrics & Adolescent Medicine*, 155(12), 1351–1356.

KOPPENSTEINER, M. F., AND L. MENEZES (2021): "Violence and Human Capital Investments," *Journal of Labor Economics*.

LEVINE, P. B., AND R. MCKNIGHT (2020): "Exposure to a School Shooting and Subsequent Well Being," *Unpublished Manuscript*.

MOFFITT, T. E. (2013): "Childhood exposure to violence and lifelong health: Clinical intervention science and stress-biology research join forces," *Development and Psychopathology*, 25(4pt2), 1619–1634.

NILSEN, L. G., Å. LANGBALLE, AND G. DYB (2016): *"Men hva er det egentlig ment for?" Voldsoffererstatning til de overlevende etter terrorangrepet på Utøya 22. juli 2011 ["But What Is It Meant For?" Victim Compensation for the Survivors of the Terror Attack on July 22, 2011]*. Norwegian Centre for Violence and Traumatic Stress Studies (In Norwegian).

NOU (2012:14): *Rapport fra 22. juli-kommisjonen [Report of the 22 of July Commision]*. Office of the Prime Minister, Norway (In Norwegian).

OTTO, M. W., A. HENIN, D. R. HIRSHFELD-BECKER, M. H. POLLACK, J. BIEDERMAN, AND J. F. ROSENBAUM (2007): "Posttraumatic stress disorder symptoms following media exposure to tragic events: Impact of 9/11 on children at risk for anxiety disorders," *Journal of anxiety disorders*, 21(7), 888–902.

ROSSIN-SLATER, M., M. SCHNELL, H. SCHWANDT, S. TREJO, AND L. UNIAT (2020): "Local exposure to school shootings and youth antidepressant use," *Proceedings of the National Academy of Sciences*, 117(38), 23484–23489.

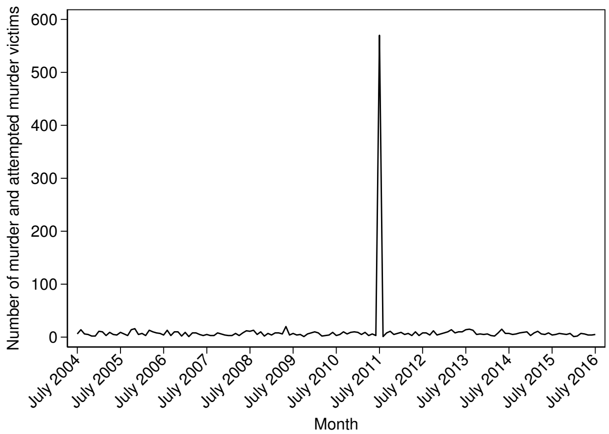
SHANY, A. (2016): “Too Scared for School? The effects of terrorism on Israeli student achievement,” Discussion paper, Working Paper. Hebrew University of Jerusalem.

SHARKEY, P. (2010): “The acute effect of local homicides on children’s cognitive performance,” *Proceedings of the National Academy of Sciences*, 107(26), 11733–11738.

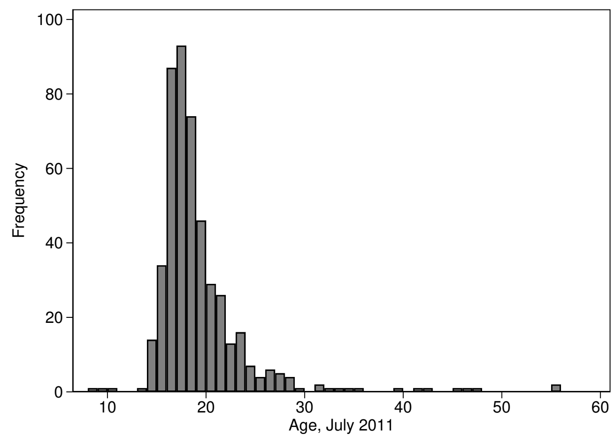
STENE, L. E., AND G. DYB (2015): “Health service utilization after terrorism: a longitudinal study of survivors of the 2011 Utøya attack in Norway,” *BMC Health Services Research*, 15(1), 158.

STURUP, J., A. ROSTAMI, H. MONDANI, M. GERELL, J. SARNECKI, AND C. EDLING (2019): “Increased gun violence among young males in Sweden: a Descriptive National Survey and International Comparison,” *European Journal on Criminal Policy and Research*, 25(4), 365–378.

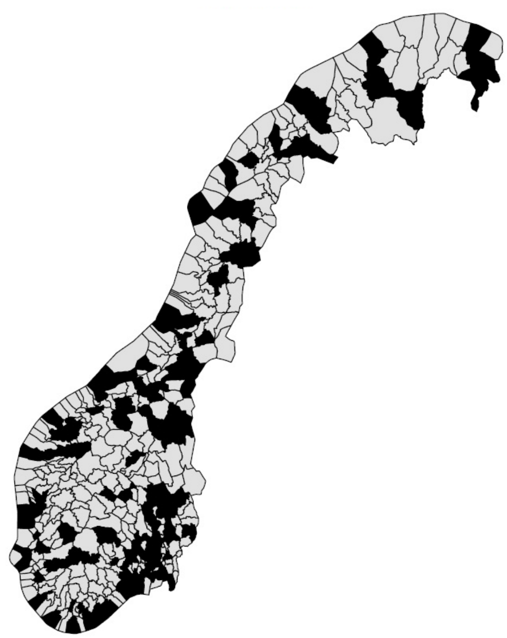
Figure 1: Descriptive Figures.



(a) The Number of Murder or Attempted Murder Victims from July 2004 to July 2016.



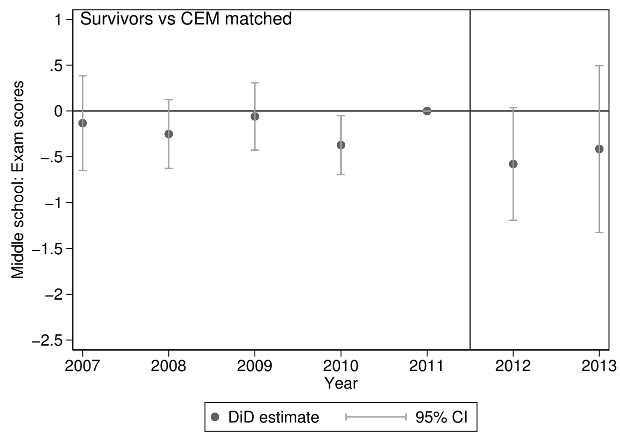
(b) The Age Distribution of the Utøya Survivors at the Time of Incident.



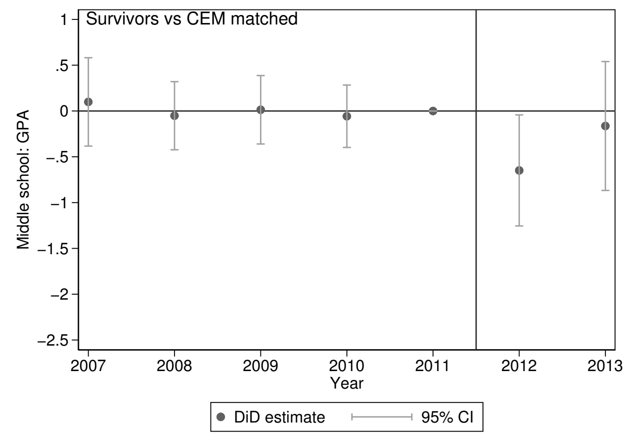
(c) The Map of Norway Illustrating the Residence Municipalities of the Utøya Survivors.

Notes: Panel (A) shows the number of individuals registered as victims of a murder or an attempted murder in Norway in each month from July 2004 to July 2016. Panel (B) shows the age distribution of the Utøya survivors, limited to residents of Norway. Panel (C) shows on a map of Norway where the municipalities from which the Utøya survivors hailed from are illustrating in black areas, defined by the mother's residential municipality in 2010 (with some exceptions).

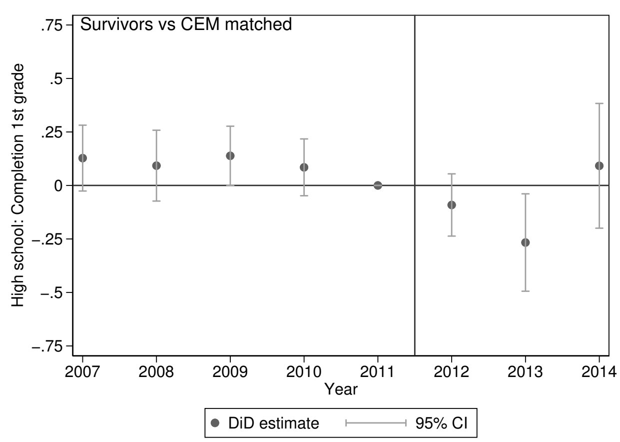
Figure 2: Graphical DiD Evidence: Comparisons of Survivors and Matched Controls



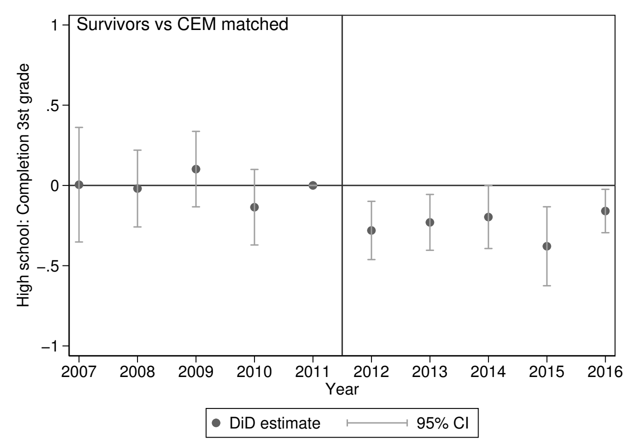
(a) Middle School: Exam Scores.



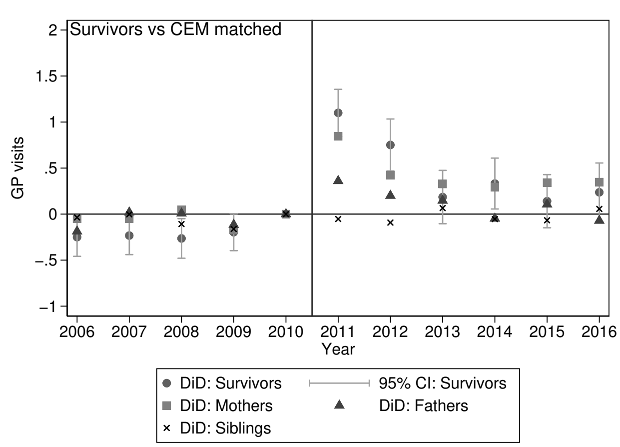
(b) Middle School: GPA.



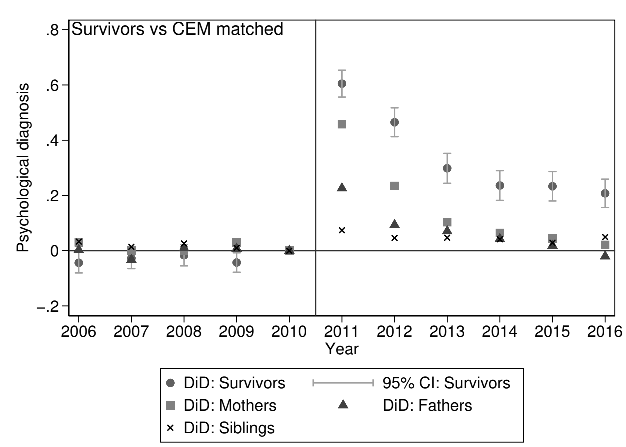
(c) High School: Completed 1st Grade.



(d) High School: Completed 3rd Grade.



(e) Number of GP Visits.



(f) Psychological Diagnosis.

Notes: Survivors have been matched to control individuals using coarsened exact matching (CEM) on birth year, gender, parental earnings, education and age, and the centrality of the residential municipality. The sample in panels (a) and (b) includes cohorts 1991-1997. Panel (c) includes cohorts 1990-1997 and panel (d) 1988-1997. Panel (e) and (f) includes cohorts 1983-1997, as well as the point estimates for the mothers, fathers and siblings of survivors. The DiD estimate is the difference in weighted mean outcome between the survivors and control individuals, minus the equivalent in 2010. In panel (a)-(d) we compare outcomes for individuals finishing a certain level of education in a given year, to those finishing in academic year 2010/11. In panel (e) and (f) we observe outcomes for all individuals in the sample in each year. In panel (a) and (b), outcomes are standardized within each academic year, and standard errors are clustered on the school-cohort level. Each figure reports 95% confidence intervals.

Table 1: Impacts on the Survivors.

	Middle School		High School		Health	
	(1)	(2)	(3)	(4)	(5)	(6)
	Exam Scores	GPA	Completion 1st Grade	Completion 3rd Grade	Numer of GP visits	Psychological diagnosis
Panel A: Short-term						
DiD	-0.326	-0.527***	-0.189***	-0.213***	0.644***	0.367***
	(0.233)	(0.196)	(0.0592)	(0.0543)	(0.0849)	(0.0165)
Pre-event mean, survivors	0.385	0.393	0.806	0.681	1.057	0.0954
N	6401	6605	7567	8217	9032	9032

	Education			Labor Market	
	(7)	(8)	(9)	(10)	(11)
	High School Diploma	Ever College Enrollment	Ever College Completion	Employment	Log-Earnings
Panel B: Medium-term					
Treated-Control Difference	-0.129***	-0.0712*	-0.0638*	-0.0489**	-0.133
	(0.0317)	(0.0392)	(0.0326)	(0.0197)	(0.0929)
Dependent mean, control	0.906	0.697	0.604	0.940	12.22
N	1143	1143	1962	1962	1671

Notes: In columns (1)-(6), survivors born 1983-1997 were matched to control individuals using coarsened exact matching (CEM) on birth year, gender, parental earnings, education and age, and centrality of municipality. In column (7)-(11), the same matching is performed adding 10th grade exam scores for a restricted set of cohorts. Columns (1) and (2) display outcomes at the end of 10th grade for cohorts 1991-1997 during the academic years 2006/07-2012/13. Columns (3) and (4) display outcomes for cohorts 1990-1997 and 1988-1995, respectively, during academic years 2006/07-2013/14 and 2006/07-2013/14. Columns (5) and (6) include all matched individuals for each year 2006-2016. In columns (7)-(8), we further restrict our analysis to cohorts 1993-1995, i.e., individuals not old enough to have finished high school before July 22, 2011. In columns (7) and (9)-(11), outcomes are measured in 2018. In column (8) we measure higher education enrollment in any of the years 2012-2018. The DiD estimates (columns (1)-(6)) are the differences in weighted mean outcomes between the survivors and the control individuals after the academic year 2010/11 (calendar year 2010 for health outcomes), minus the equivalent before. Outcomes in columns (1) and (2) are standardized for each year. The estimates in columns (7)-(11) are the differences in weighted mean between survivors and the control individuals. Standard errors are shown in parenthesis, and clustered at the school-cohort level for columns (1) and (2), and otherwise robust standard errors are used.

$^{*} p < 0.10$ $^{**} p < 0.05$ $^{***} p < 0.01$

Table 2: Impacts on the Family Members of Survivors.

	Labor Market			Marital Status	Health	
	(1)	(2)	(3)	(4)	(5)	(6)
	Employment	Log-Earnings	Sickness Leave	Married	Number of GP visits	Psychological diagnosis
Panel A: Mothers						
DiD	0.00242	0.00310	0.0671***	-0.00367	0.474***	0.142***
	(0.0134)	(0.0242)	(0.0179)	(0.0146)	(0.161)	(0.0169)
Pre-event mean, treated	0.890	12.85	0.241	0.611	2.830	0.212
N	8505	6962	8505	8504	8505	8505
Panel B: Fathers						
DiD	0.00455	-0.00361	0.00912	-0.0138	0.171	0.0743***
	(0.0108)	(0.0314)	(0.0141)	(0.0163)	(0.133)	(0.0151)
Pre-event mean, treated	0.897	13.13	0.181	0.644	2.161	0.141
N	8247	6634	8247	8247	8247	8247

	Education		Health	
	(7)	(8)	(9)	(10)
	Middle School Exam Scores	Middle School GPA	Number of GP visits	Psychological diagnosis
Panel C: Siblings				
DiD	-0.227*	-0.220*	0.0377	0.0311***
	(0.132)	(0.128)	(0.0818)	(0.0119)
Pre-event mean, treated	0.396	0.318	1.524	0.120
N	2992	3196	12054	12054

Notes: Survivors born 1983-1997 were matched to control individuals using coarsened exact matching (CEM) on birth year, gender, parental earnings, education and age, and centrality of municipality. Their family members were identified using mother and father personal identifiers. In columns (1)-(6) and (9)-(10), outcomes are measured in years 2006-2016. In columns (7)-(8), outcomes are measured in academic years 2006/07-2015/16, and limited to survivors' younger siblings. The DiD estimates are the differences in weighted mean outcome between the family members of survivors and the control individuals after 2010 (academic year 2010/11), minus the equivalent before. Employment is defined as having any positive labor income in a given year. Sickness leave is defined as having any physician-certified sickness leave in a given year. Earnings are CPI adjusted. Marital status is measured January 1st the following year. Exam scores and GPA are standardized. Standard errors are shown in parenthesis, and always clustered on mother ID.

$^{*}p < 0.10$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table 3: Impacts on the School Peers of Survivors.

	Middle School		High School	Health	
	(1) Exam Scores	(2) GPA	(3) Completion 3rd Grade	(4) Number of GP Visits	(5) Psychological Diagnosis
Panel A: All Peers					
DiD	-0.0694	-0.0251	0.00679	-0.0600	0.00802
	(0.0519)	(0.0574)	(0.0278)	(0.0468)	(0.00580)
Pre-event mean, peers of survivors	0.00316	-0.0181	0.540	1.379	0.101
N	51700	52290	84717	115492	115492
Panel B: Female Peers					
DiD	-0.161***	-0.0224	-0.0184	-0.0683	0.0140*
	(0.0565)	(0.0656)	(0.0269)	(0.0731)	(0.00798)
Pre-event mean, peers of survivors	0.164	0.212	0.649	1.721	0.114
N	24940	25562	41306	56269	56269
Panel C: Male Peers					
DiD	0.0256	-0.0225	0.0301	-0.0540	0.00218
	(0.0620)	(0.0718)	(0.0320)	(0.0378)	(0.00654)
Pre-event mean, peers of survivors	0.164	-0.234	0.432	1.051	0.0885
N	25948	26728	43411	59223	59223

Notes: Survivors born 1983-1997 were matched to one control individual using coarsened exact matching (CEM) on birth year, gender, parental earnings, education and age, and centrality of municipality. Peers are individuals in the same school and cohort as the survivor or matched control individual. In columns (1) and (2), peers are defined in the 10th grade, and cohorts 1991-1997 are included. In column (3), peers are defined in the 1st grade of high school, and cohorts 1988-1994 are included. In columns (4) and (5), we combine the samples of column (1) and (2), and only include the academic year individuals were supposed to finish the 10th grade in middle school or the 3rd grade in high school. We include academic years 2006/07-2012/13 in all regressions. The DiD estimates are the differences in mean outcomes between the peers of survivors and the peers of control individuals after the academic year 2010/11 (calendar year 2010 for health outcomes), minus the equivalent before. Outcomes in columns (1) and (2) are standardized for each year. Standard errors are shown in parenthesis, and clustered at the school-cohort level.

$^{*} p < 0.10$ $^{**} p < 0.05$ $^{***} p < 0.01$

Online Appendix A: Details on Data and Sample Selections

Survivors

- *Middle school*
  - Academic years: 2006/07-2012/13. Note that 2012/13 is the last year with a sufficient number of survivors to be included in the sample.
  - Cohorts: 1991-1997.
  - Only survivors with a matched control individual are included.
  - Pairs of survivors-controls are removed if either the survivor or the control individual are not found in the middle school grade registry the year they turned 16.
  - Only observe each individual once in each regression.
- *High school*
  - Academic years: Grade 1: 2006/07-2013/14. Grade 3: 2006/07-2015/16. The last year with with a sufficient number of survivors to be included is 2013/14 (2015/16) for grade 1 (3).
  - Cohorts: Grade 1: 1990-1997; Grade 3: 1988-1997.
  - Only survivors with a matched control individual are included.
  - High school completion of grade 1 and grade 3 is coded as zero if individual is not found in the high school registry data set in the relevant year.
  - Only observe individual once in each regression.
- *Health*
  - Calendar years: 2006-2016.
  - Cohorts: 1983-1997.
  - Only survivors with a matched control individual are included.
  - Individuals are observed every year. If they are not found in the KUHR-dataset in the relevant year, variables are set to zero.
- *Medium-term outcomes*
  - Years: Enrollment in higher education in years 2012-2018. High school diploma and completed higher education by 2018. Employment and earnings in 2018.

- Cohorts: 1993-1995 for high school diploma and enrollment in higher education. 1986-1995 for all other outcomes.
- If either the survivor or the none of the matched control individuals are alive and living in Norway in 2018, then the survivor-control pair is removed from the sample.
- Note that different cohorts are measured at different ages and we use the last available year of data for each outcome (except enrollment).

PARENTS

- Parents are found with mother's and father's ID in the population registry.
- Parents of more than one survivor (or control individual) are included in the estimation as many times as the number of their children for each year. Since each survivor has a control individual, there will be an equal number of parent observation in total.
- All parents of children born in 1988-1997.
- Only parents of matched survivors are included in the sample.
- Restrict to individuals living by December 31, 2016.
- Residents in Norway for the whole period (has used immigration and emigration dates).
- *All variables*
  - Years: 2006-2016
  - Estimations for mothers and fathers are run separately.

SIBLINGS

- Individuals are defined as siblings if they have the same mother.
- The oldest sibling is chosen as the focal victim if both siblings were victimized.
- We remove control individuals for victims with older siblings that were also victims.
- We remove control individuals for victims without any (non-victim) siblings.
- We remove victims with control individuals without siblings.
- Siblings are defined as younger if they are strictly younger than the oldest victimized sibling. Siblings born in the same year as the oldest victimized sibling are excluded.
- *Health*
  - All siblings.

- Years: 2006-2016.

- *Middle school exam score and GPA*
  - Only younger siblings.
  - Academic years: 2006/07-2015/16
  - Cohorts: 1991-2000

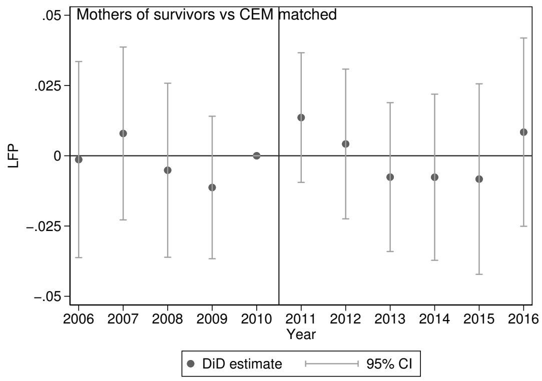
PEERS

- In the following, we call the survivors and their matched controls as the focal individuals.
- *Middle school*
  - Peers are pupils in the same school and cohort as the focal individual.
  - We only include schools of survivors that have a match.
  - Only observe each individual once.
  - Academic years: 2006/07-2012/13. Note that 2012/13 is the last year with a sufficient number of survivors to be included in the sample.
  - Cohorts: 1991-1997.
- *High school*
  - Peers are pupils in the same high school grade 1 and cohort as the focal individual.
  - We only include those starting high school before July 22, 2011 (time of event), otherwise high school choice and enrollment are potentially endogenous.
  - We only include schools of survivors that has a match.
  - Academic years: 2006/07-2012/13. We cannot observe longer, as they have to have started high school before July 22, 2011.
  - Cohorts: 1988-1994.
- *Health*
  - We pool peers from middle school and high school sample.
  - We observe them in the same academic year as they are in the middle school or high school sample.
  - Only observe each individual once.
  - Academic years: 2006/07-2012/13 (August-June).

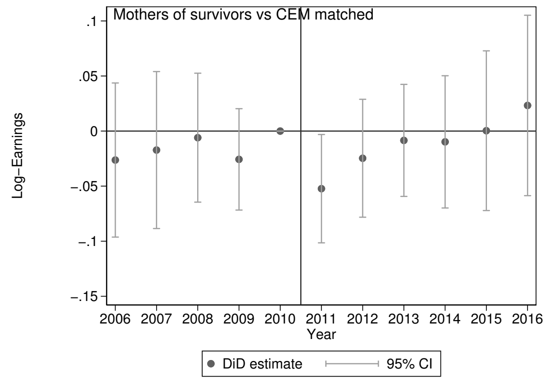
Online Appendix B: Additional Results

Additional Figures

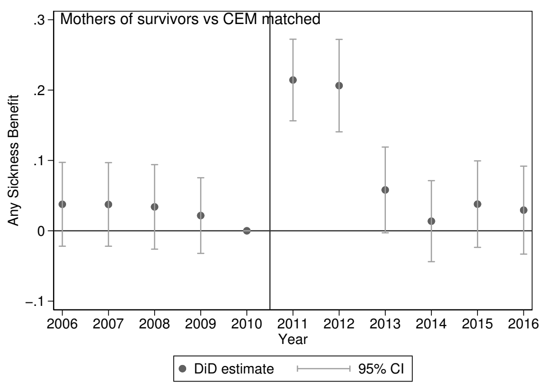
FIGURE B1: Yearly DiD Estimates: Comparisons of Survivors' Mothers and Matched Controls



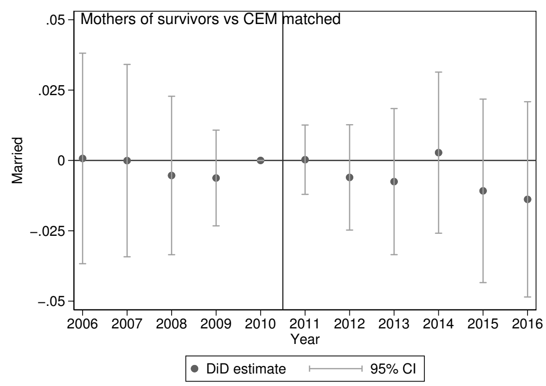
(A) Employment



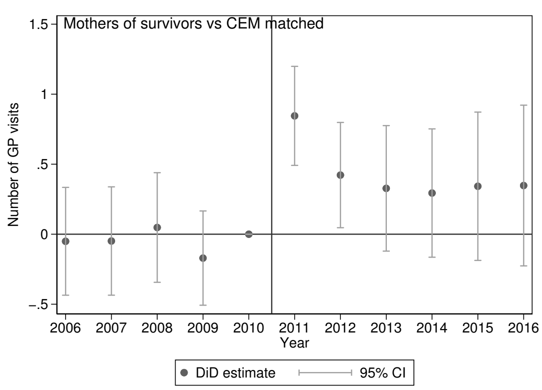
(B) Log-Earnings



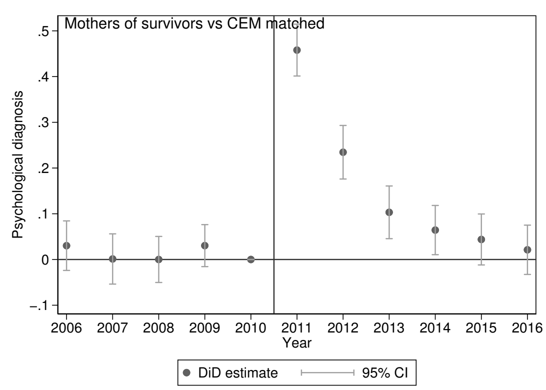
(C) Any sickness benefit



(D) Married



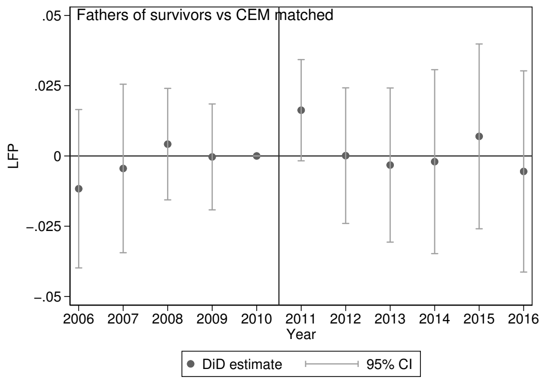
(E) Number of GP visits



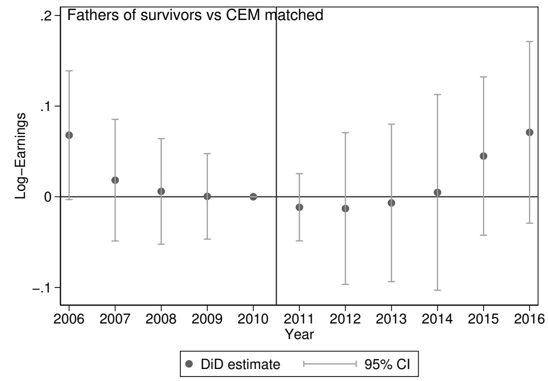
(F) Psychological diagnosis

Notes: Survivors have been matched to control individuals using coarsened exact matching on birth year, gender, parental earnings, education and age, and the centrality of the residential municipality. The DiD estimate is the difference in weighted mean outcome between the mothers of survivors and the control individuals, minus the equivalent in 2010. LFP and any sickness benefit is defined as having any positive labor income or sickness benefit in a given year. Earnings are CPI adjusted. Being married is measured January 1st the following year. Standard errors are clustered on mother ID. 95% confidence intervals.

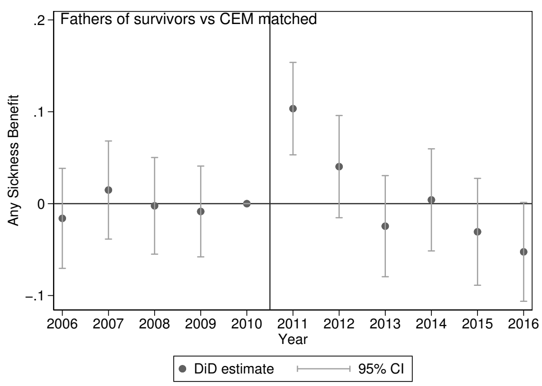
FIGURE B2: Yearly DiD Estimates: Comparisons of Survivors' Fathers and Matched Controls



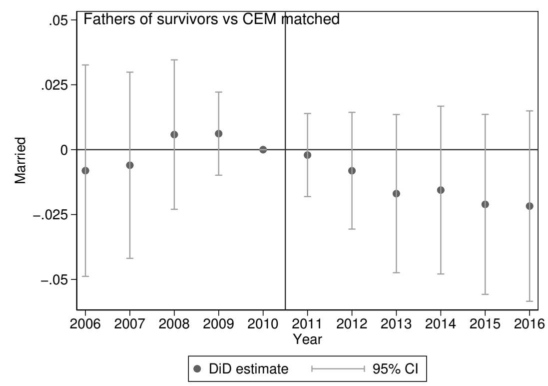
(A) Employment



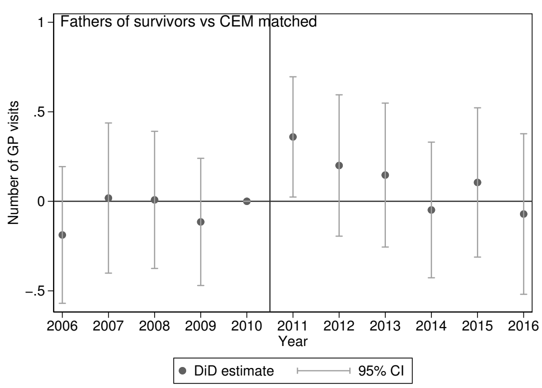
(B) Log-Earnings



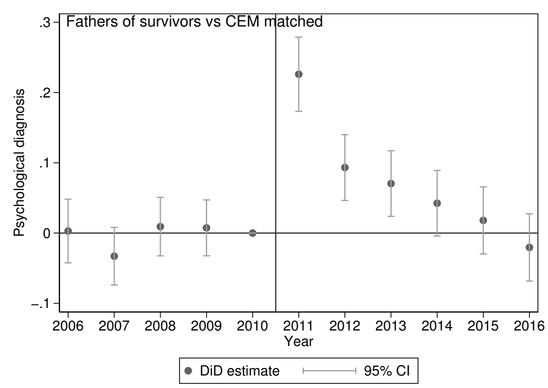
(C) Any sickness benefit



(D) Married



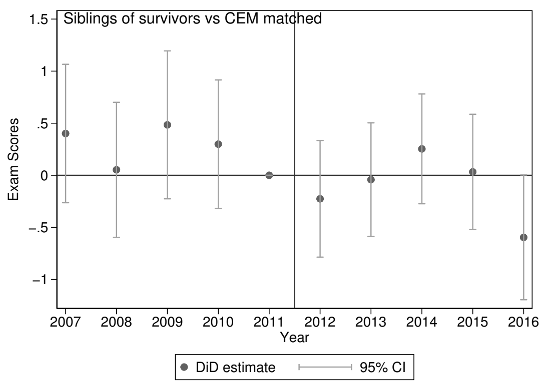
(E) Number of GP visits



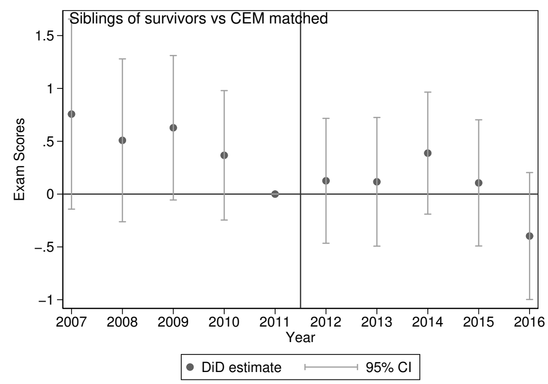
(F) Psychological diagnosis

Notes: Survivors have been matched to control individuals using coarsened exact matching on birth year, gender, parental earnings, education and age, and the centrality of the residential municipality. The DiD estimate is the difference in weighted mean outcome between the mothers of survivors and the control individuals, minus the equivalent in 2010. LFP and any sickness benefit is defined as having any positive labor income or sickness benefit in a given year. Earnings are CPI adjusted. Being married is measured January 1st the following year. Standard errors are clustered on mother ID. 95% confidence intervals.

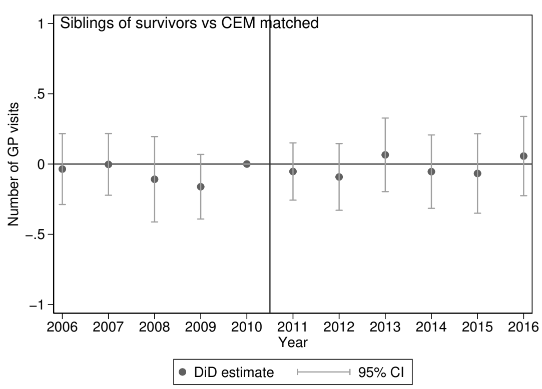
FIGURE B3: Yearly DiD Estimates: Comparisons of Survivors' Siblings and Matched Controls



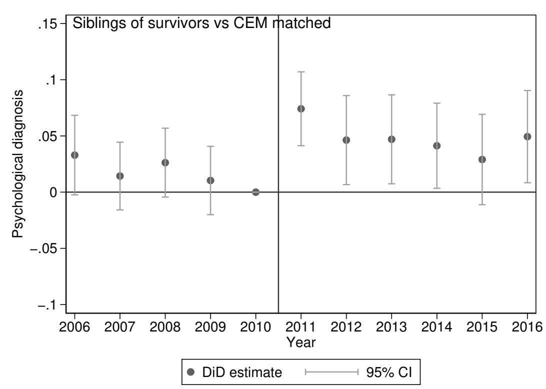
(A) Middle School: Exam Scores



(B) Middle School: GPA



(C) Number of GP visits



(D) Psychological diagnosis

Notes: Survivors have been matched to control individuals using coarsened exact matching on birth year, gender, parental earnings, education and age, and the centrality of the residential municipality. The DiD estimate is the difference in weighted mean outcome between the siblings of survivors and the control individuals, minus the equivalent in 2010. The sample panel (a) and (b) includes only younger siblings and columns (c) and (d) include all siblings. Standard errors are clustered on mother ID. 95% confidence intervals.

ADDITIONAL TABLES

TABLE B1: Descriptive Statistics – Characteristics of the Survivors

	Mean
Female	0.46
Mother's age at birth	28.56
Mother's earnings in 1000 NOK, 2010	351.74
Mother with higher education	0.48
Father's age at birth	31.53
Father's earnings in 1000 NOK, 2010	473.87
Father with higher education	0.37
Parents live in urban area	0.36
N	438

Notes: The sample consists of survivors born in 1983-1997. Survivors without matched control individuals from the Coarsened Exact Matching (CEM) procedure are left out of the sample. Earnings are only reported if positive.

TABLE B2: Descriptive Statistics – Outcomes of the Survivors

	Pre-event Mean	Post-event Mean
Middle School		
Exam scores	0.38	-0.16
GPA	0.39	-0.28
High School		
Completion 1st grade	0.81	0.66
Completion 3rd grade	0.68	0.44
Health		
Number of GP visits per year	1.06	2.31
Psychological diagnosis in a year	0.10	0.52
Medium term		
High school diploma		0.78
Enrolled in higher education		0.63
Completed higher education		0.54
Labor force participation		0.89
Earnings in 1000 NOK		299.22

Notes: The full sample consists of survivors born in 1983-1997. Survivors without matched control individuals from the Coarsened Exact Matching (CEM) procedure are left out of the sample. For middle school outcomes, we exlude cohorts born before 1991. For 1st (3rd) grade completion in high school, we exlude cohorts born before 1990 (1988). For high school diploma and enrolled in higher education, the sample is limited to survivors born 1993-1995. For the remaining medium-term outcomes, the sample is limited to those born in 1986-1995. Middle school, high school and health are measured in both the pre-period, while medium term outcomes are measured in (or up to) 2018. Earnings are only reported if positive.

Chapter 3

Income, Consumption and Savings around Retirement

Income, Consumption, and Savings around Retirement*

Ola Vestad† and Mirjam Wentzel‡

Abstract

This paper studies the evolution of household income, consumption, and savings around retirement in the context of a major reform of the public pension system in Norway. The 2011 reform implied a substantial increase in work incentives for workers covered by the early retirement scheme AFP and a reduction in the early retirement age for workers not covered by AFP, and more flexibility in terms of pension claiming and retirement decisions for both groups of workers. These changes in incentives and choice sets allow us to provide new evidence on the links between the structure of public pension systems and household consumption and savings behavior around retirement. Difference-in-differences analyses around the early retirement age reveal substantial positive reform impacts on household net income. The positive impacts on household net income are passed through to both savings and consumption, and contribute to a flattening of consumption profiles for workers faced with the flexible post-reform system. Event study analyses around retirement reveal decreasing consumption profiles for pre-reform cohorts and flat profiles for post-reform cohorts. These findings suggest that the flexible post-reform pension system leaves workers in a better position to maintain their pre-retirement levels of consumption than the more rigid pre-reform system.

Keywords: pension reform, consumption, savings
JEL codes: D12, D14, E21, H55, J14, J26

---

*This research was conducted as part of the project 'Globalization, Labor Market Restructuring, and Economic Policy' (project # 295901) funded by the Norwegian Research Council. We thank seminar participants at the NHH-PSE NORFACE Workshop, NHH-UiB PhD Workshop, Stockholm University, and Norwegian School of Economics for helpful comments and suggestions. The paper substantially improved as a result of conversations with Katrine V. Løken.

†Statistics Norway, Research Department. E-mail: olavestad@gmail.com

‡Norwegian School of Economics, FAIR and Statistics Norway. E-mail: mirjam.wentzel@gmail.com

1 Introduction

In response to dramatic demographic changes and the associated financial pressures on PAYGO old-age pension systems, many countries have implemented reforms to increase the work incentives for older individuals. While an extensive literature studies the relationship between financial incentives of pension rules and individual retirement behavior,[1] less is known about the effects of (changes in) pension incentives on households' net income, savings, and consumption expenditure. Understanding how labor supply and pension claiming decisions are passed through to household consumption and savings behavior is key to understanding the broader welfare implications of pension system designs and pension reforms. In this paper, we provide new evidence on the behavioral responses to pension incentives. We study the evolution of household income, consumption, and savings around retirement, and draw on the 2011 reform of the Norwegian pension system to investigate how the consumption and savings behavior of households is affected by the structure of the public pension system.

The pre-reform Norwegian pension system can be characterized by three main features: (i) a regular retirement age of 67 and an early retirement age of 62, with eligibility for pension claiming prior to the regular retirement age restricted to workers covered by the early retirement scheme AFP (AFP workers); (ii) no increases in future pensions for AFP workers postponing pension claiming past the early retirement age; and (iii) a restrictive earnings test[2] for AFP workers claiming pensions prior to the regular retirement age. The 2011 pension reform brought major changes along all three dimensions: (i) both AFP and non-AFP workers can now start claiming pensions from age 62; (ii) the pension system has actuarially neutral pension adjustments for all claiming ages between 62 and 75; and (iii) there is no earnings testing. These three changes together imply that the decision to claim pension benefits is now largely disentangled from the decision to withdraw from the labor market, and the regular retirement age at 67 is effectively abolished.

The 2011 reform affected AFP and non-AFP workers in distinctively different ways. For AFP workers, it implied a substantial increase in the net returns to work after the early retirement age, but no change in the early retirement age. For non-AFP workers, it implied a reduction in the early retirement age, but no change in the work incentives

[1] See e.g., Gruber and Orszag (2003), Song and Manchester (2007), Haider and Loughran (2008), Brown (2013), Staubli and Zweimüller (2013), Vestad (2013), Hernæs et al. (2016), Manoli and Weber (2016), Brinch et al. (2018b), Gelber et al. (2020).

[2] The "earnings test" is a common institutional feature of public pension systems, and it affects those who might consider combining pension receipt with continued work: If labor earnings pass a certain threshold, pension benefits are reduced by some fraction. The earnings test comes in addition to regular income taxes (on both pensions and earnings), implying high implicit tax rates on labor earnings for individuals claiming pensions.

around the (early) retirement age. Previous studies have documented substantial labor supply responses to the improved work incentives for private sector AFP workers (Hernæs et al., 2016; Brinch et al., 2018b; Kruse, 2021), and (at most) moderate reductions in labor supply in response to the reduction in the early retirement age for non-AFP workers (Hernæs et al., 2016; Brinch et al., 2018b).[3] Moreover, claiming pensions at the early retirement age is much more common under the post-reform system, while exiting the labor market at the same age is much less common. This implies that many workers chose to take the new opportunity to combine pension receipt with continued work (Brinch et al., 2018b).

The goal of this paper is to provide new evidence on the links between the structure of public pension systems and household consumption and savings behavior around retirement. Our work draws on two strengths of the Norwegian environment. First, the detailed Norwegian register data cover the full population of registered citizens and allow us to link employment, earnings, and benefits receipts and characterize the paths to retirement for different types of workers. Unique person and household level identifiers allow us to link individuals to their households, and by linking administrative tax records of the households' income and wealth to data on their sales and purchases of assets, we obtain comprehensive measures of household income, savings, and wealth and a derived measure of consumption expenditure. Second, the 2011 reform brought major changes in the financial incentives and choice sets for workers approaching the early retirement age, and thereby provides a unique opportunity to compare the consumption and savings behavior of households under two different pension systems: a rigid pre-reform system and a flexible post-reform system.

Our first set of analyses investigates the evolution of household income, consumption, and savings around the early retirement age, comparing workers reaching the early retirement age right before and after the reform. These analyses yield three main findings. First, the reform had substantial positive impacts on household net income, reflecting the large fraction of workers combining early pension claiming with continued work in the post-reform pension system. Second, the positive impacts on household net income are reflected in both savings and consumption profiles. The shifts in savings profiles mirror the shifts in income profiles, and we find that the positive reform impacts on household savings are almost entirely due to increased savings in liquid assets (bank deposits). Consumption profiles are weakly decreasing for pre-reform cohorts and almost entirely flat for post-reform cohorts. This might suggest that the flexible post-reform pension

[3]Kruse (2021) studies spousal spillover effects of the 2011 reform, and finds that the labor supply of female public sector workers is positively affected by their private sector husbands' postponed retirement, while male public sector workers are unaffected by their private sector wives' postponed retirement.

system leaves workers in a better position to maintain their pre-retirement levels of consumption than the more rigid pre-reform system. Third, while household net savings are either negative or moderately positive around the early retirement age, both pre- and post-reform cohorts are accumulating wealth as they pass the early retirement age. This highlights the importance of capital gains (mainly on real estate) for the evolution of elderly Norwegian households' wealth levels.

With our second set of analyses, we revisit two puzzles in the household consumption and retirement literature: the "retirement consumption puzzle", that consumption drops upon retirement; and the "retirement savings puzzle", that households continue to accumulate wealth as they transition into retirement. While the existing literature is almost exclusively based on various sources of survey data,[4] which tend to suffer from problems such as small sample sizes, under-reporting of spending, and non-random response rates, we draw on the exceptionally rich Norwegian register data. We also provide novel evidence on how the structure of pension systems interacts with these puzzles.

We conduct event study analyses around retirement and the first year with pension income, and contrast the behavior of workers faced with the rigid pre-reform system with the behavior of workers faced with the flexible post-reform system. Event study plots around the first year with pension income reveal patterns resembling the ones we documented with our analyses of income, consumption, and savings around the early retirement age: For pre-reform cohorts, income, savings, and consumption decline around the first year with pension income. For post-reform cohorts, both income and savings increase while the consumption profile is flat (or weakly increasing). In contrast, event study plots around retirement show that both income and savings decline at the onset of retirement, and average net savings are negative after retirement for both pre- and post-reform cohorts. Consumption profiles are weakly decreasing, and somewhat flatter for post-reform cohorts than for pre-reform cohorts, while wealth profiles are flat or weakly increasing around retirement. Together, these findings suggest that Norwegian households are able to maintain a consumption level that is similar to their pre-retirement consumption level as they transition into retirement. And this stability of consumption is achieved without decumulating assets: negative net savings are offset by capital gains on financial and real assets (mainly real estate), leaving the levels of wealth largely unchanged around the onset of retirement.

To summarize our results in the light of the retirement consumption and savings puzzles,

[4]See e.g., Banks et al. (1998); Bernheim et al. (2001); Hurd and Rohwedder (2003); Aguiar and Hurst (2005); Ameriks et al. (2007); Haider and Stephens (2007); Hurst (2008); Battistin et al. (2009); Aguiar and Hurst (2013). Agarwal et al. (2015) and Olafsson and Pagel (2018) are two recent contributions using transactions data (rather than survey data) to study the changes in consumption around retirement.

we first note that our analyses reveal decreasing consumption profiles for pre-reform cohorts and flat (or moderately increasing) profiles for post-reform cohorts. This might suggest that more flexible pension systems leave workers in a better position to maintain their levels of consumption than less flexible ones. Second, the declining consumption profiles for pre-reform cohorts are somewhat puzzling, considering their non-decreasing levels of wealth: One might have expected to see a decumulation of assets along with a flat consumption profile rather than a decline in consumption and a flat (or increasing) wealth profile. This puzzle could be related to the fact that Norwegian households hold a large fraction of their wealth in real estate, which might imply that moderate adjustments in wealth levels are hindered by large transaction costs.

The remainder of the paper proceeds as follows. Section 2 reviews the key features regarding the Norwegian pension system and the 2011 reform, describes our data and sample selection, and presents descriptive evidence on the impacts of the reform on employment, earnings, and pension receipts for the individual workers in our sample. Section 3 estimates the impacts of the reform on household income, consumption, and savings using a difference-in-differences framework, comparing the behavior of pre- and post-reform workers before and after the early retirement age. Section 4 presents the results of event study analyses around retirement and the first year with pension income, contrasting the behavior of workers faced with the rigid pre-reform system with the behavior of workers faced with the flexible post-reform system. The final section concludes.

2 Institutional setting, data, and descriptive evidence

In this section, we first review the key features regarding the Norwegian pension system and the 2011 reform. We then provide information on our data sources and sample selection. Finally, we present descriptive evidence on employment, earnings, and pension receipts around age 62 for the individual workers in our sample, to illustrate the impacts of the reform on labor market outcomes.

2.1 The Norwegian pension system and the 2011 reform

The Norwegian pension system consists of several different parts, with the most important one for a majority of the pensioners being the National Insurance Scheme (NIS). The NIS is a universal coverage pay-as-you-go pension system, with a minimum pension provided to all Norwegian citizens and an additional earnings related pension for those who have had sufficient earnings throughout their working lives. Prior to the reform, pension benefits from the NIS were only available from age 67, but the contractual early retirement scheme AFP, which covered public sector workers and about 50% of all private sector workers, provided covered workers with pension benefits between ages 62 and 67. The AFP scheme was rather generous, in that benefits from age 62 were the same as they would have been had the worker continued working until age 67. But postponed claiming was not compensated with higher future benefits, and there was a strict earnings testing system in place.

The 2011 reform brought major changes to both the public pension system (the NIS) and the private sector AFP scheme, with AFP now being completely integrated with the NIS in terms of a supplemental benefit for eligible workers. In the post-reform system, both the NIS and the AFP pension schemes can be characterized by three main features: (i) claiming of pensions can take place at any age between 62 and 75; (ii) annual benefits are subject to actuarial adjustments based on the mean expected longevity of each cohort; and (iii) there is no earnings testing.

Our analyses focus on private sector workers,[5] distinguishing between two groups of workers that were affected by the reform in very different ways. Workers employed in private sector firms affiliated with the early retirement scheme AFP (AFP workers) experienced pronounced changes in work incentives as a result of the reform, with the earnings testing and the implicit tax on continued work past the early retirement age being completely removed. For AFP workers, the early retirement age did not change with the reform. Workers not covered by AFP (non-AFP workers), on the other hand, got access to public pension benefits from the NIS up to five years earlier than before, at actuarially neutral terms, while their work incentives around the (early) retirement age were left unchanged.

The changes in pension claiming incentives for the first cohorts of AFP workers that were faced with the post-reform system can be summarized as follows.[6] In the pre-reform system, pension wealth was decreasing with age at claiming, and annuitized annual pensions, defined as total pension wealth divided by the expected number of years with pension receipt, were close to constant. In the post-reform system, pension wealth is independent of claiming age for individuals with expected longevity equal to the average for their cohort, and annuitized annual pensions are increasing with age at claiming. Individuals with expected longevities below (above) the average for their cohort will

[5]Public sector workers are also covered by a contractual AFP pension scheme, but a breakdown in negotiations about the pension reform in the public sector left the public sector AFP scheme largely unchanged. For this reason we confine the analysis to private sector workers.

[6]See Brinch et al. (2018b) for further details.

increase (decrease) their total pension payouts by claiming early.[7]

The impacts of the reform on expected pension wealth are somewhat involved and depend on the income histories and the chosen claiming ages of affected workers. Brinch et al. (2018b) show that for a hypothetical individual with 40 years of earnings at age 62, earning 6 Basic Amounts[8] every year (which is close to the average yearly income of an AFP worker), the pre-reform system is more generous than the post-reform system, if the worker chooses to start claiming pensions prior to age 64.5. This is largely due to a change in the indexation of pension benefits: While pension benefits in the pre-reform system were indexed according to the average wage growth in the economy, pension benefits in the post-reform system are indexed according to average wage growth minus 0.75%. Hence, although initial pension levels are largely unchanged for the first cohorts faced with the post-reform system, conditional on claiming pensions close to the early retirement age, the difference in indexation results in more pronounced differences in pension levels at older ages. For workers choosing to postpone claiming to older ages, e.g., age 67, initial pension levels will be higher than they would have been in the pre-reform system, but the less generous indexation implies a convergence towards the corresponding pre-reform pension level.

2.2 Data sources and sample selection

Our empirical analyses are based on several administrative data sources that we link together using unique identifiers for firms, individual workers, and their households. We first construct a sample consisting of individuals who were in their late fifties or older at the time of the reform and employed in a private sector firm at age 58. Using unique household identifiers, we link the workers in our individual sample to their households and construct a household panel data set with comprehensive measures of household income, savings, and wealth and a derived measure of consumption expenditure, covering the period 2007–2014.

Using longitudinal administrative registers provided by Statistics Norway, covering every Norwegian resident from 1967 onwards, we obtain individual demographic information

[7]Brinch et al. (2018a) study the relationship between pension claiming incentives and claiming decisions in the Norwegian pension system, after the 2011 reform, and find that individuals who choose to claim pensions early are to a large extent those who gain in terms of expected lifetime pensions from claiming early.

[8]The Basic Amount (BA) is a central feature of the public pension system in Norway. It is adjusted every year, with a nominal rate of growth varying between 2 and 13% since its introduction in 1967, and from the late nineties onward in accordance with the average wage growth in the economy. In 2014, the most recent year covered by our data, one BA was equal to NOK 87,328, which at the time of writing corresponds to about USD 10,200.

such as birth and death dates, gender, and level of education. These registers also allow us to identify recipients of AFP, disability, and old age pensions, and the first month of benefit receipt for each type of pension.

Our main source of data on employment and earnings is a matched employer-employee panel data set, consisting of annual tax records of the universe of workers. This register is used in the administration of sickness benefits and therefore subject to extensive quality controls. The dataset includes information on total earnings, contracted hours, and start and stop dates for each job spell. We use the information on earnings and the start and stop dates to construct measures of employment and retirement. We obtain information on the AFP membership status of private sector firms from the AFP membership register, and use this information to classify workers into two groups according to the AFP affiliation of their main employer at age 58: AFP and non-AFP workers.

Our baseline individual sample consists of private sector workers who were employed and did not receive disability or unemployment insurance benefits at the end of the calendar year in which they turned 58. The sample includes a set of pre-reform cohorts, consisting of individuals who turned 62 within a window of four years prior to the reform (born 1945–1948), and a set of post-reform cohorts consisting of individuals who were younger than 62 in 2010 (born 1949–1955).[9]

Claiming AFP pensions and exiting the labor force at or shortly after age 62 was a very common response to the incentives in the pre-reform regime, which implies that the AFP covered workers that postponed claiming and continued working past that age is a strongly selected group. Following Brinch et al. (2018b), we therefore exclude post-2010 observations of the pre-reform cohorts. We also exclude observations of pre-reform cohorts prior to 2007, to make sure that the sample contains the same number of cohorts at each age on both sides of the reform. Hence, our empirical results are based on observations of pre-reform cohorts over the period 2007–2010, and on observations of post-reform cohorts over the period 2007–2014.

Using unique household identifiers, we link the workers in our individual sample to their households and construct a household panel data set with comprehensive measures of household income, savings, and wealth and a derived measure of consumption expenditure. These measures are based on administrative tax records containing detailed information on all sources of annual income, including earnings, self-employment income, capital income, and cash transfers, as well as most types of assets holdings and liabilities.

[9]The cohorts born later than 1952 did not reach the early retirement age within our observation window, but are included in the estimation sample to allow for flexible controls for time trends in our difference-in-differences framework.

Income data are reported in annual amounts, while the values of assets holdings and liabilities are measured as of the last day of each year.

A key challenge in studying behavioral responses to pension reforms and the evolution of consumption and savings around retirement is the lack of reliable longitudinal data on consumption and savings. The existing literature on household consumption and retirement is almost exclusively based on various sources of survey data, which tend to suffer from problems such as small sample sizes, under-reporting of spending, and non-random response rates.[10] As demonstrated by Browning and Leth-Petersen (2003), an alternative to using survey data is to create measures of consumption using longitudinal data on income and assets and the accounting identity that total consumption expenditure is equal to income plus capital gains minus the change in wealth over the period. Eika et al. (2020) perform a similar exercise, and show that the combination of tax records on income and wealth with individual level information on financial and real estate transactions offers a unique opportunity to construct reliable measures of household consumption expenditure. Moreover, the ability to link tax records on income and wealth with individual level data on asset transactions has two other key advantages. First, it allows the researcher to directly observe net savings from data on sales and purchases of assets without having to measure or make assumptions about capital gains or changes in net wealth. Second, data on real estate transactions allow the researcher to observe market values of real estate rather than relying on tax assessment values from tax records, which differ significantly (and unsystematically) from actual market values. We use the measures of household income, consumption, savings, and wealth from Eika et al. (2020) here, and refer the reader to their paper for more details.

Table 1 describes our main analysis samples, dividing the AFP and non-AFP workers in our individual sample into four different groups in accordance with our difference-in-differences framework: pre- and post-reform cohorts of workers observed at ages 59–61 and 62–65. For both AFP and non-AFP workers, the pre- and post-reform cohorts are very similar in terms of observable characteristics. We note that more than 70% of the samples are men, and that around one fifth have some university level education. The average CPI adjusted earnings are higher for younger cohorts, reflecting a real earnings increase in Norway over the period. The average years of work experience is around 22-23 years, and having received sickness leave benefits at age 58 is fairly common. And while rather large fractions of the AFP workers are employed in large firms within the manufacturing sector, the non-AFP workers are more often observed in smaller firms and less often in the manufacturing sector.

[10] See Pistaferri (2015) for a discussion.

The table also reports household disposable income, savings, consumption, and wealth, measured at age 58 and reported per household member.[11] Net wealth corresponds to about five annual incomes for AFP workers and about six annual incomes for non-AFP workers, while net savings at age 58 are generally low and negative for three of the eight groups.

2.3 Descriptive evidence on employment, earnings, and pension receipt

Previous studies have documented substantial labor supply responses to the improved work incentives for private sector AFP workers (Hernæs et al., 2016; Brinch et al., 2018b; Kruse, 2021), and (at most) moderate reductions in labor supply in response to the reduction in the early retirement age for non-AFP workers (Hernæs et al., 2016; Brinch et al., 2018b). Moreover, claiming pensions at the early retirement age is much more common under the post-reform system, while exiting the labor market at the same age is much less common, which implies that many workers chose to take the new opportunity to combine pension receipt with continued work (Brinch et al., 2018b). In this section, we reproduce some of these results for the individual workers in our sample, in terms of descriptive evidence on the impacts of the reform on employment, earnings, and pension receipts.

Figure 1 shows employment and pension claiming rates and mean labor earnings and pension income by age for AFP workers. Comparing pre- and post-reform cohorts of workers at the same ages, we see substantial differences in employment and pension claiming rates from age 62 onward: Employment rates are about 15 percentage points higher for the post-reform cohorts at ages 62 and 63, while pension claiming rates are about 25 percentage points higher. Differences in annual earnings are also substantial (about 18 and 35 percent at ages 62 and 63), while mean pension income for post-reform cohorts corresponds to about twice the mean pension income for pre-reform cohorts. From Figure 2, we see that while non-AFP workers appear to be largely unaffected by the reform in terms of employment rates and earnings, about 29 percent of post-reform workers chose to take the opportunity to claim pension benefits at the early retirement age.[12]

[11] For household level variables, we exclude a small number of observations with negative values of consumption in each year t, and the top and bottom percent in the distribution of consumption in years t and $t-1$. Appendix Table A.1 shows that these restrictions have little impact on the composition of the sample and on average income, consumption, savings, and wealth.

[12] Small fractions of non-AFP workers are claiming AFP pension benefits at ages 62–65. This is partly due to job changes after age 58, and partly due to mis-classification; that we select the wrong main job at age 58 or that we classify a firm as non-affiliated while in reality it was affiliated with the private sector AFP scheme.

The differences revealed in Figures 1 and 2 are broadly in line with the estimated reform effects in the existing literature. We also note that both employment rates and mean labor earnings evolve in similar ways prior to age 62 for pre- and post-reform cohorts of workers, which lends support to the parallel trends assumption in our difference-in-differences framework.

3 The impacts of the reform on income, consumption, and savings

In this section, we investigate the impacts of the 2011 reform on household income, consumption, and savings for AFP and non-AFP workers. We first present descriptive evidence on the evolution of income, consumption, and savings around the early retirement age, comparing workers reaching the early retirement age right before and after the reform. Against this background, we then turn to a difference-in-differences framework to estimate reform impacts on household income, consumption, and savings and their main components.

3.1 Descriptive evidence

Figure 3 shows mean income, savings, consumption, and net wealth by age for AFP workers. The income profiles in the upper left panel reveal substantial differences across pre- and post-reform cohorts of workers, reflecting the large fraction of workers combining early pension claiming with continued work in the post-reform system. A decline in mean income starting around age 60 for pre-reform cohorts is replaced by a sharp increase from age 61 to age 63 for post-reform cohorts, while mean income is weakly decreasing from age 63 onward for both pre- and post-reform cohorts. Similar differences are visible also in terms of the savings profiles in the upper right panel. For pre-reform cohorts, mean net savings are mostly negative and close to zero at all ages. For post-reform cohorts, mean net savings are positive at all ages, increasing from age 61 to 63 and declining from age 63 to age 65. The consumption profiles in the lower left panel show that household consumption expenditure is decreasing for pre-reform cohorts and close to constant for post-reform cohorts. Finally, the lower right panel shows that mean net wealth is increasing for both pre- and post-reform cohorts, with the differences in savings levels being reflected in a steeper wealth profile for post- than for pre-reform cohorts.

The declining consumption profile for pre-reform cohorts is somewhat puzzling, considering their increasing levels of wealth: One might have expected to see a decumulation of assets along with a flat consumption profile rather than a decline in consumption and an increasing wealth profile. This puzzle could be related to the fact that Norwegian households hold a large fraction of their wealth in real estate,[13] which might imply that moderate adjustments in wealth levels are hindered by large transaction costs.

Figure 4 provides a corresponding set of age profiles for non-AFP workers. The differences across pre- and post-reform cohorts of non-AFP workers are somewhat less pronounced than for AFP workers, but show similar patterns. The more moderate differences for non-AFP workers are in line with the different implications of the reform in terms of pension claiming and work incentives for the two groups of workers. For AFP workers, the reform implied a substantial increase in the net returns to work after the early retirement age, but no change in the early retirement age. For non-AFP workers, it implied a reduction in the early retirement age, but no change in the work incentives around the (early) retirement age. These changes in incentives led to substantial increases in early pension claiming for both groups of workers, and to increased employment and earnings for AFP workers only.

Taken together, this descriptive evidence might suggest that the flexible post-reform pension system leaves workers in a better position to maintain their pre-retirement levels of consumption than the more rigid pre-reform system. And while household net savings are either negative or moderately positive around the early retirement age, both pre- and post reform cohorts are accumulating wealth as they pass the early retirement age. This highlights the importance of capital gains (mainly on real estate) for the evolution of wealth levels for elderly Norwegian households.

3.2 Difference-in-differences estimation

To estimate the impacts of the 2011 reform on household income, consumption, and savings, we follow Brinch et al. (2018b) and use a difference-in-differences (DiD) framework comparing the outcomes of pre- and post-reform cohorts before and after the early retirement age. In particular, we estimate the following DiD specification;

$$Y_{a,t} = \alpha + \gamma_a + \lambda_t + \sum_{k=60}^{65} \beta_k \mathbf{1}[a = k, G(t) = 1] + \varepsilon_{a,t}, \tag{1}$$

where a and t denote age and calendar year, respectively, γ_a and λ_t are age and year fixed effects, and $G(t)$ is an indicator for post-reform observations (years 2011–2014). The coefficients of interest β_k capture the impacts of the 2011 reform, estimated as the

[13]Eika et al. (2020) show that real estate is the key component of gross wealth for most Norwegian households: Real estate accounts for more than 70% of total gross wealth for all households, and real estate's fraction of gross wealth is below .5 only for households in the top 1% of the distribution of net wealth.

difference in mean outcomes before and after the reform at each age $k \in \{60, \ldots, 65\}$, relative to the before-after difference in mean outcomes at age 59. The coefficients β_{60} and β_{61} are included to assess the validity of the identifying assumption, i.e., that the changes in outcomes for pre-reform workers are valid counterfactuals for the changes in outcomes for post-reform workers. We estimate equation (1) for each of the outcomes net income, savings, consumption, and net wealth. To summarize the estimates and estimate reform impacts on the main components of household net income and savings, we also construct a single estimate of the reform impact at ages 62–65 by replacing the age-specific indicators $\mathbf{1}[a = k, G(t) = 1]$ and the corresponding coefficients β_k with a single post-treatment indicator $\mathbf{1}[a \geq 62, G(t) = 1]$ and a single coefficient β.

Figure 5 plots the DiD estimates of β_k from the specification in equation (1), for AFP workers at each age $k \in \{60, \ldots, 65\}$. Prior to age 62, household income, consumption, and savings evolve in similar ways for pre- and post-reform workers. The estimates at age 61 for income and savings are marginally significant at the 95% level, but the point estimates are close to zero and very close to the age 60 estimates. The similarity in pre-trends supports the identifying assumption, and indicates that the changes in outcomes in the pre-reform period are valid counterfactuals for the changes in outcomes in the post-reform period. In line with the descriptive evidence presented above, the DiD estimates at ages 62–65 show significant differences in income, consumption, and savings from age 62, with the largest differences occurring at age 63 or 64. For net wealth, the DiD estimates are much less precise and indicate a divergence in trends prior to age 62. These results should therefore be interpreted with caution.

Estimates of the reform impacts at ages 62–65 on income, consumption, savings, and wealth are provided in Table 2. On average, the reform led to an increase in net income of about NOK 40,000, which corresponds to approximately 10% of mean income for pre-reform cohorts of workers in the same age range. 65% of the reform impact on net income is reflected in increased savings and the remaining 35% in increased consumption. The DiD estimates on income, consumption, and savings are precisely estimated and not sensitive to including a set of controls for worker characteristics. The estimates for net wealth, in contrast, are imprecise and not robust to the inclusion of controls.

Tables 3 and 4 report DiD estimates on the main components of income and savings, respectively. The estimates reported in Table 3 show that the positive impact on net income is driven by a substantial positive impact on transfers, which includes both AFP and NIS pensions, and a moderate positive impact on market income. These positive impacts are counteracted by a positive impact on taxes paid, which corresponds to about one third of the estimated impact on market income and transfers. For Table 4, we have decomposed net savings into three broad components: one-year changes in liquid capital (cash and bank deposits) and debt, respectively, and other savings. The DiD estimates for each of the three components show that the positive impact on net savings is mainly due to an increase in bank deposits (75% of the total effect), but also to a moderate reduction in debt (9%) and an increase in other savings (16%).

Corresponding DiD estimates for non-AFP workers are provided in Figure 6 and Tables 5–7. The age specific DiD estimates in Figure 6 are broadly similar to the corresponding estimates for AFP workers in Figure 5. Prior to age 62, household income, consumption, and savings evolve in similar ways for pre- and post-reform workers, while the DiD estimates for net wealth are much less precise and indicate a divergence in pre-trends. For income and savings, the estimated reform impacts are positive and significantly different from zero from age 62 onward. The estimated impacts on consumption are positive and significantly different from zero from age 63.

Table 5 shows that on average, the reform had a positive impact on net income at ages 62–65 of about NOK 28,000, which corresponds to about 6% of mean income for pre-reform cohorts of workers in the same age range. 75% of the reform impact on net income is reflected in increased savings and the remaining 25% in increased consumption. Compared to the estimates for AFP workers in Table 2, the reform impact on net income is smaller, and an even larger share of the impact on net income is reflected in increased savings.

The estimates reported in Table 6 show a substantial positive impact on transfers, which is offset by a moderate negative impact on market income and a positive impact on taxes paid. There is also some evidence of a moderate positive impact on other income, which includes dividends, interest on deposits and liabilities, other capital income (e.g., returns on life insurance, taxable rental income, and capital income from abroad), and the value of owner-occupied housing services. The estimated impacts on the components of net savings in Table 7 show that the positive impact on net savings is mainly due to an increase in bank deposits also for non-AFP workers (64% of the total effect), but the estimated impact on other savings is larger than for AFP workers (30% of the total effect), although not significantly different from zero.

To summarize, our DiD estimates show that the 2011 reform had substantial positive impacts on household income, savings, and consumption around the early retirement age. In line with the different implications of the reform in terms of pension claiming and work incentives for AFP and non-AFP workers, the estimated reform impacts are generally larger for AFP workers than for non-AFP workers. For non-AFP workers, the positive impact on net income is entirely due to a positive impact on transfers, reflecting the substantial increase in early pension claiming and the moderate labor supply responses for this group. The different labor supply responses for the two groups of workers might also explain why a larger share of the estimated impact on net income is reflected in increased savings for non-AFP workers than for AFP workers, if the reform led to an increase in work related expenses for AFP workers but not for non-AFP workers.

4 Income, consumption, and savings around retirement

In this section, we revisit two puzzles in the household consumption and retirement literature: the "retirement consumption puzzle", that consumption drops upon retirement; and the "retirement savings puzzle", that households continue to accumulate wealth as they transition into retirement. The existing literature on these topics is either based on various sources of survey data (see e.g., Banks et al., 1998; Bernheim et al., 2001; Hurd and Rohwedder, 2003; Aguiar and Hurst, 2005; Ameriks et al., 2007; Haider and Stephens, 2007; Hurst, 2008; Battistin et al., 2009; Aguiar and Hurst, 2013) or on transactions data (Agarwal et al., 2015; Olafsson and Pagel, 2018)), with relatively small sample sizes. In contrast, our analysis is based on detailed Norwegian register data with comprehensive measures of household income, consumption, savings, and wealth for the full population of registered citizens. Drawing on these data and on policy variation from a substantial reform of the Norwegian pension system, we conduct event study analyses around retirement and the first year with pension income, and contrast the behavior of workers faced with the rigid pre-reform system with the behavior of workers faced with the flexible post-reform system.

4.1 Event study setup

We consider two separate events: retirement, defined as the first year without an active employment spell at the end of the year; and pension claiming, defined as the first year with pension income. For each of the two events, we recenter the data such that the event occurs at year zero, and follow workers from up to three years before to up to three years after the event.

We base our analysis on the household sample described in Section 2.2. Similar to the DiD analysis, we disregard pre-reform cohort events occurring after 2010, to maintain a sharp contrast between pre- and post-reform cohorts of workers. For the post-reform cohorts, we restrict attention to those who turned 62 prior to 2014 (born in 1952 or before), since age 62 is the earliest possible claiming age for both NIS and AFP pensions and since only a small fraction of the workers in our sample retire prior to age 62. With this setup, we observe retirement events occurring between ages 60 and 65 and claiming events occurring between ages 62 and 65. For non-AFP workers we restrict attention to the retirement events, since claiming pensions prior to age 67 was not an option under the pre-reform pension system.

4.2 Graphical evidence

Figure 7 shows event study plots around the first year with pension income for AFP workers. In line with the large fractions of AFP workers claiming pensions at the early retirement age, the figure reveals patterns resembling the ones we documented with our analyses of income, consumption, and savings around the early retirement age in Section 3. The differences between pre- and post-reform cohorts of workers are rather pronounced. For post-reform cohorts, both income and savings increase sharply at the onset of pension claiming. Wealth levels are increasing, while the consumption profile shows moderate increases in consumption around the onset of pension claiming. For pre-reform cohorts, in contrast, income, savings, and consumption decline around the first year with pension income, while wealth levels remain constant. And while the declines in income and savings are sharp and occur at the first full year with pension income, the decline in consumption is more gradual and starts one year prior to pension claiming.

Event study plots around retirement are provided in Figures 8 and 9, for AFP and non-AFP workers, respectively. The broad patterns are similar both across pre- and post-reform cohorts of workers and across AFP and non-AFP workers, except for somewhat larger changes in income and savings at the onset of retirement for AFP than for non-AFP workers. Both income and savings decline at the onset of retirement, average net savings are negative after retirement, while wealth profiles are flat or weakly increasing. Consumption profiles are weakly decreasing around retirement, and for AFP workers, somewhat flatter for post-reform cohorts than for pre-reform cohorts.

Together, these findings suggest that Norwegian households are able to maintain a consumption level that is similar to their pre-retirement consumption level as they transition into retirement. And this stability of consumption is achieved without decumulating assets: Negative net savings are offset by capital gains on financial and real assets (mainly real estate), leaving the levels of wealth largely unchanged around the onset of retirement. For AFP workers, the contrasts between pre- and post-reform cohorts seem to suggest that the flexible post-reform pension system leaves workers in a better position to maintain their pre-retirement levels of consumption than the more rigid pre-reform system.

5 Conclusion

In this paper we have studied the evolution of household income, consumption, and savings around retirement in the context of a major reform of the public pension system in Norway. The 2011 reform implied a substantial increase in work incentives for workers covered by the early retirement scheme AFP and a reduction in the early retirement age for workers not covered by AFP, and more flexibility in terms of pension claiming and retirement decisions for both groups of workers. We drew on these changes in incentives and choice sets to provide new evidence on the links between the structure of public pension systems and household consumption and savings behavior around retirement. Difference-in-differences analyses around the early retirement age revealed substantial positive reform impacts on household net income. We found that the positive impacts on household net income are passed through to both savings and consumption, and contribute to a flattening of consumption profiles for workers faced with the flexible post-reform system. Event study analyses around retirement revealed decreasing consumption profiles for pre-reform cohorts and flat profiles for post-reform cohorts. These findings suggest that the flexible post-reform pension system leave workers in a better position to maintain their pre-retirement levels of consumption than the more rigid pre-reform system.

Our analyses provide new evidence on how the structure of pension systems interacts with two puzzles in the household consumption and retirement literature: the "retirement consumption puzzle", that consumption drops upon retirement; and the "retirement savings puzzle", that households continue to accumulate wealth as they transition into retirement. To summarize our results in the light of these puzzles, we first note that our analyses reveal decreasing consumption profiles for pre-reform cohorts and flat (or moderately increasing) profiles for post-reform cohorts. This might suggest that more flexible pension systems leave workers in a better position to maintain their levels of consumption than less flexible ones. Second, the declining consumption profiles for pre-reform cohorts are somewhat puzzling, considering their non-decreasing levels of wealth: One might have expected to see a decumulation of assets along with a flat consumption profile rather than a decline in consumption and a flat (or increasing) wealth profile. This puzzle could be related to the fact that Norwegian households hold a large fraction of their wealth in real estate, which might imply that moderate adjustments in wealth levels are hindered by large transaction costs.

References

Agarwal, S., Pan, J., and Qian, W. (2015). The Composition Effect of Consumption around Retirement: Evidence from Singapore. *American Economic Review: Papers and Proceedings*, 105(5):426–31.

Aguiar, M. and Hurst, E. (2005). Consumption versus Expenditure. *Journal of Political Economy*, 113(5):919–948.

Aguiar, M. and Hurst, E. (2013). Deconstructing Life Cycle Expenditure. *Journal of Political Economy*, 121(3):437–492.

Ameriks, J., Caplin, A., and Leahy, J. (2007). Retirement Consumption: Insights from a Survey. *The Review of Economics and Statistics*, 89(2):265–274.

Banks, J., Blundell, R., and Tanner, S. (1998). Is There a Retirement-Savings Puzzle? *American Economic Review*, 88(4):769–788.

Battistin, E., Brugiavini, A., Rettore, E., and Weber, G. (2009). The Retirement Consumption Puzzle: Evidence from a Regression Discontinuity Approach. *American Economic Review*, 99(5):2209–26.

Bernheim, B. D., Skinner, J., and Weinberg, S. (2001). What Accounts for the Variation in Retirement Wealth among U.S. Households? *American Economic Review*, 91(4):832–857.

Brinch, C. N., Fredriksen, D., and Vestad, O. L. (2018a). Life Expectancy and Claiming Behavior in a Flexible Pension System. *The Scandinavian Journal of Economics*, 120:979–1010.

Brinch, C. N., Vestad, O. L., and Zweimüller, J. (2018b). Excess Early Retirement? Evidence from the Norwegian 2011 Pension Reform. Working Paper.

Brown, K. M. (2013). The Link Between Pensions and Retirement Timing: Lessons from California Teachers. *Journal of Public Economics*, 98:1–14.

Browning, M. and Leth-Petersen, S. (2003). Imputing consumption from income and wealth information. *The Economic Journal*, 113(488):F282–F301.

Eika, L., Mogstad, M., and Vestad, O. L. (2020). What Can We Learn about Household Consumption Expenditure from Data on Income and Assets? *Journal of Public Economics*, 189.

Gelber, A. M., Jones, D., and Sacks, D. W. (2020). Estimating Adjustment Frictions

Using Nonlinear Budget Sets: Method and Evidence from the Earnings Test. *American Economic Journal: Applied Economics*, 12(1):1–31.

Gruber, J. and Orszag, P. (2003). Does the Social Security Earnings Test Affect Labor Supply and Benefit Receipt? *National Tax Journal*, 45:1–22.

Haider, S. and Loughran, D. (2008). The Effect of the Social Security Earnings Test on Male Labor Supply: New Evidence from Survey and Administrative Data. *Journal of Human Resources*, 43(1):57.

Haider, S. J. and Stephens, M. (2007). Is There a Retirement-Consumption Puzzle? Evidence Using Subjective Retirement Expectations. *The Review of Economics and Statistics*, 89(2):247–264.

Hernæs, E., Markussen, S., Piggott, J., and Røed, K. (2016). Pension Reform and Labor Supply. *Journal of Public Economics*, 142:39 – 55.

Hurd, M. and Rohwedder, S. (2003). The Retirement-Consumption Puzzle: Anticipated and Actual Declines in Spending at Retirement. *NBER Working Paper 9586*.

Hurst, E. (2008). The Retirement of a Consumption Puzzle. *NBER Working Paper 13789*.

Kruse, H. (2021). Joint Retirement in Couples: Evidence of Complementarity in Leisure. *The Scandinavian Journal of Economics*, forthcoming.

Manoli, D. and Weber, A. (2016). Nonparametric Evidence on the Effects of Financial Incentives on Retirement Decisions. *American Economic Journal: Economic Policy*, 8(4):160–82.

Olafsson, A. and Pagel, M. (2018). The Retirement-Consumption Puzzle: New Evidence from Personal Finances. Working Paper 24405, National Bureau of Economic Research.

Pistaferri, L. (2015). Household Consumption: Research Questions, Measurement Issues, and Data Collection Strategies. *Journal of Economic and Social Measurement*, 40(1–4):123–149.

Song, J. G. and Manchester, J. (2007). New Evidence on Earnings and Benefit Claims Following Changes in the Retirement Earnings Test in 2000. *Journal of Public Economics*, 91(3):669–700.

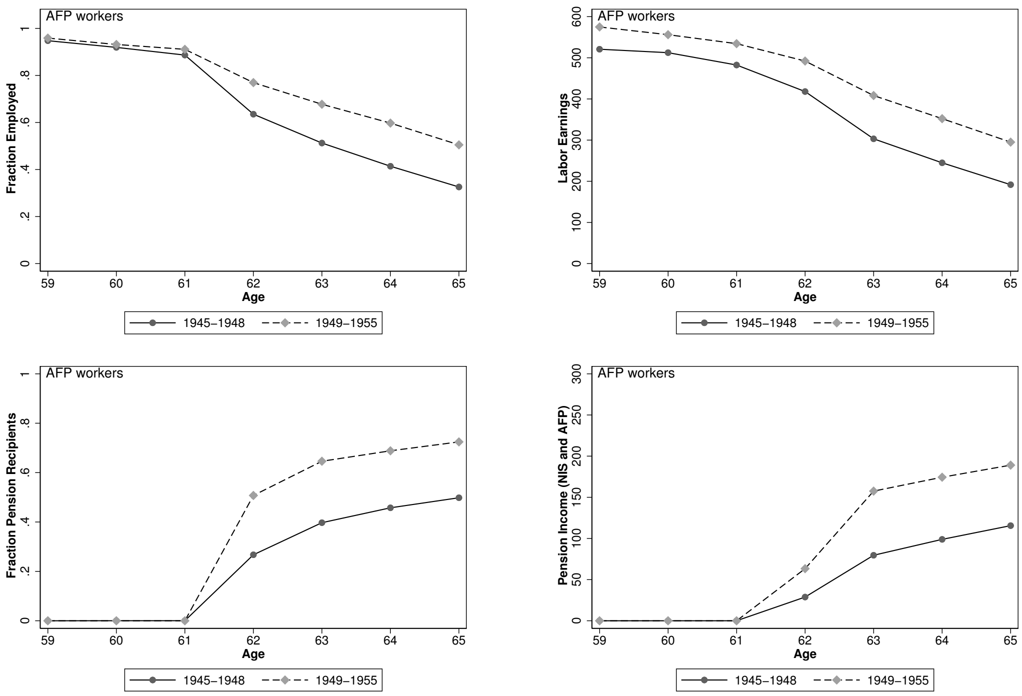
Staubli, S. and Zweimüller, J. (2013). Does Raising the Early Retirement Age Increase Employment of Older Workers? *Journal of Public Economics*, 108:17–32.

Vestad, O. L. (2013). Labour Supply Effects of Early Retirement Provision. *Labour Economics*, 25:98 – 109.

Figures

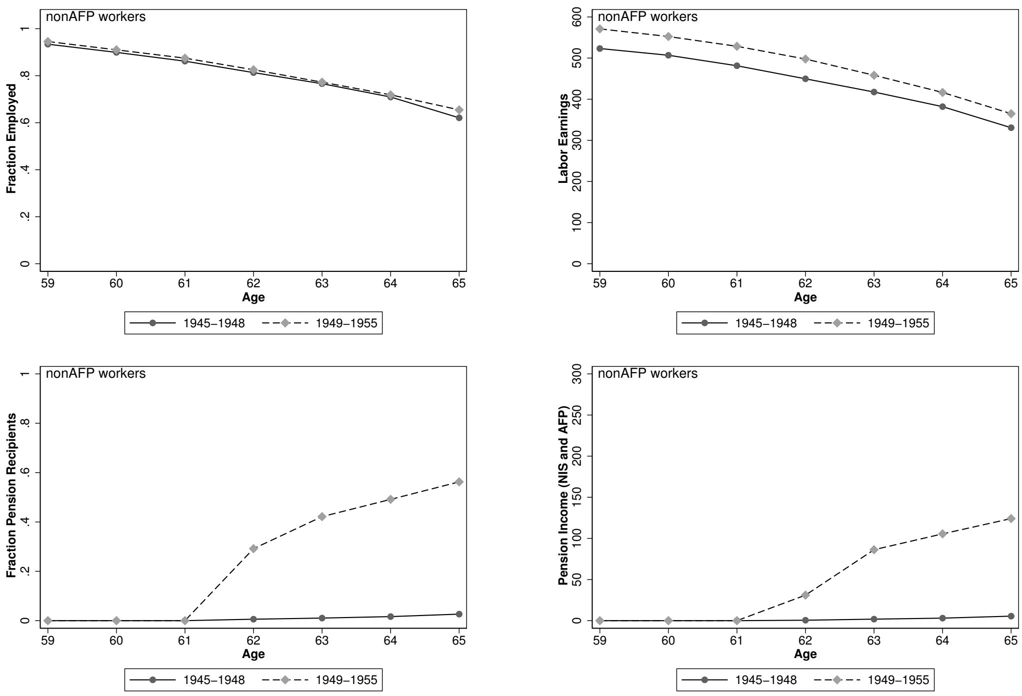
Descriptive evidence on employment, earnings and pension receipt

Figure 1: Employment, labor earnings, and pensions - AFP workers



Notes: This figure shows the evolution of employment, labor earnings, pension receipt, and pension income around age 62, for individuals who were employed by an AFP affiliated firm at age 58 ("AFP workers"). The sample used is the individual sample as defined in Section 2.2. Labor earnings and pension income are reported in NOK 1,000s and adjusted to 2014 prices.

Figure 2: Employment, labor earnings, and pensions - non-AFP workers

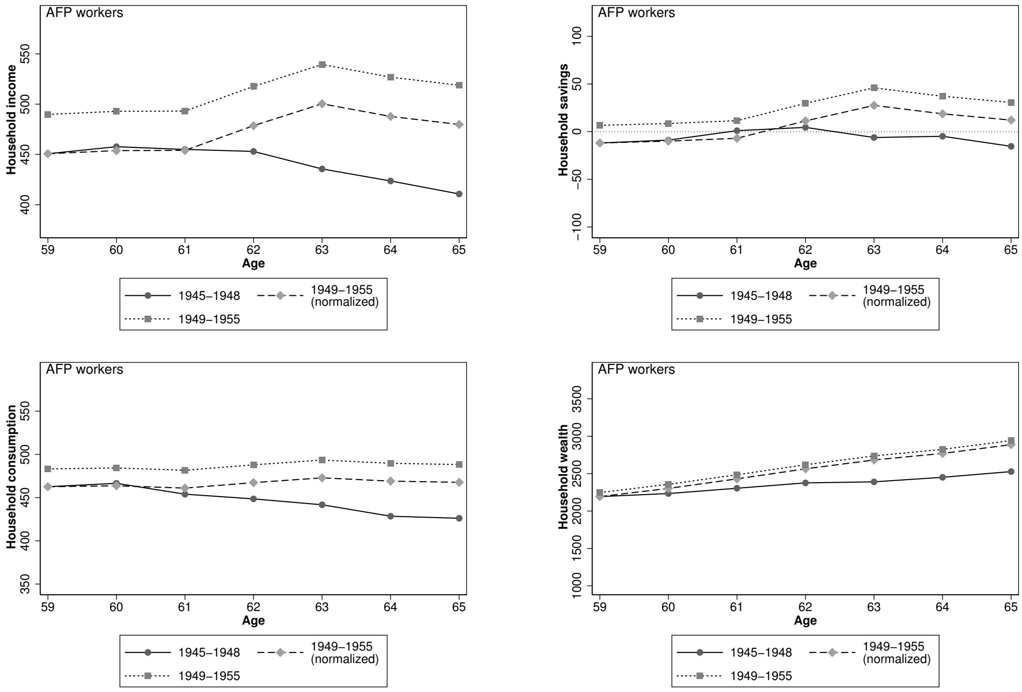


Notes: This figure shows the evolution of employment, labor earnings, pension receipt, and pension income around age 62, for individuals who were employed by a firm not affiliated with AFP firm at age 58 ("non-AFP workers"). The sample used is the individual sample as defined in Section 2.2. Labor earnings and pension income are reported in NOK 1,000s and adjusted to 2014 prices.

The impacts of the reform on income, consumption, and savings

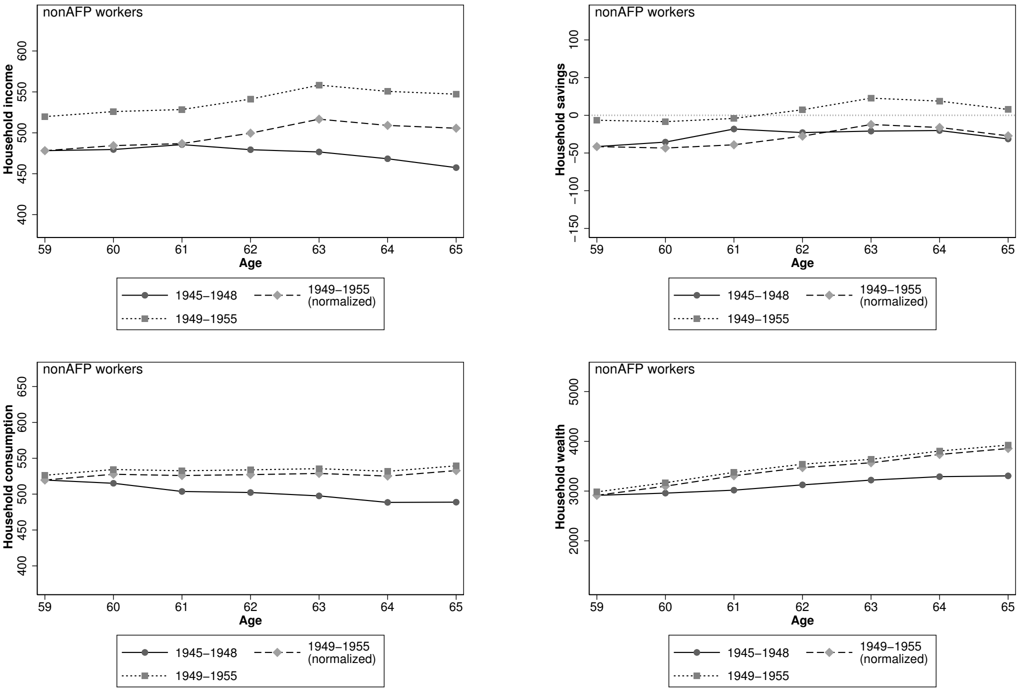
Descriptive evidence

Figure 3: Income, savings, consumption, and wealth around age 62 - AFP workers



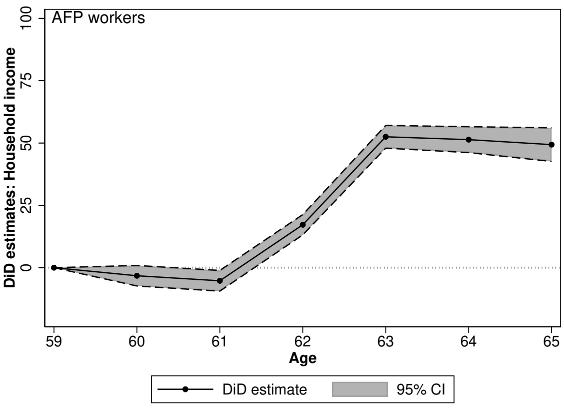
Notes: This figure shows the evolution of household income, savings, consumption, and net wealth around age 62, for the households of workers employed by an AFP affiliated firm at age 58 ("AFP workers"). The sample used is the household sample as described in Section 2.2. All variables are measured in NOK 1,000s, adjusted to 2014 prices, and reported per household member. The "1949-1955 (normalized)" series show the means of the respective variables for the post-reform cohorts after normalizing to the pre-reform cohorts level at age 59.

FIGURE 4: Income, savings, consumption, and wealth around age 62 - non-AFP workers

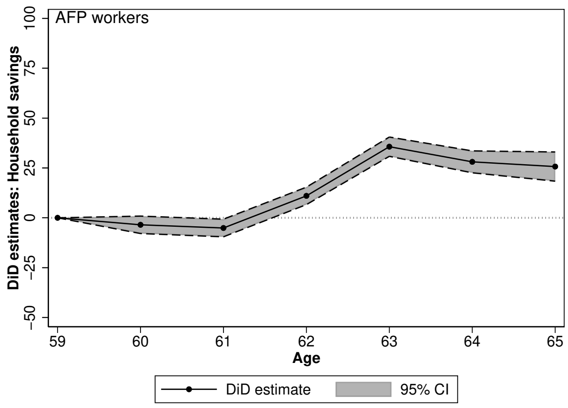


Notes: This figure shows the evolution of household income, savings, consumption, and net wealth around age 62, for the households of workers employed by a firm not affiliated with AFP at age 58 ("non-AFP workers"). The sample used is the household sample as described in Section 2.2. All variables are measured in NOK 1,000s, adjusted to 2014 prices, and reported per household member. The "1949-1955 (normalized)" series show the means of the respective variables for the post-reform cohorts after normalizing to the pre-reform cohorts level at age 59.

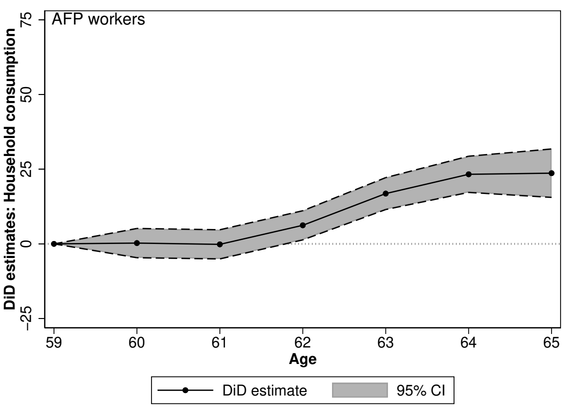
FIGURE 5: Age-specific DiD estimates for income, savings, consumption, and net wealth - AFP workers



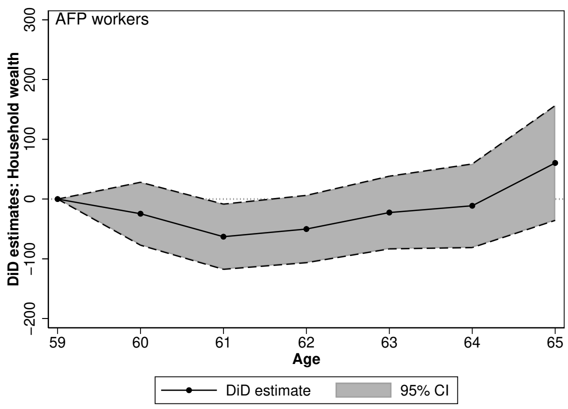
Income



Savings



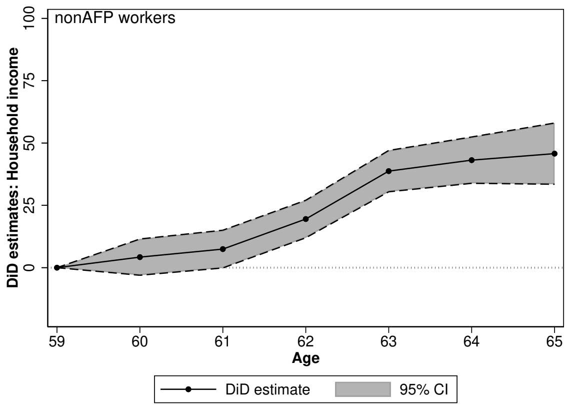
Consumption

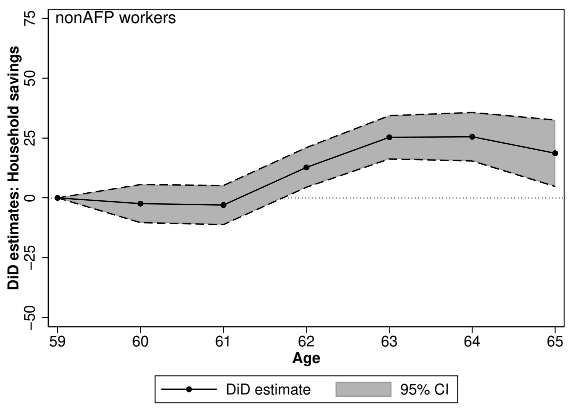


Wealth

Notes: This figure plots the estimated β_k *coefficients (along with 95% confidence intervals) from the specification in equation* (1)*, for AFP workers at each age* $k \in \{60, \ldots, 65\}$*. The sample used is the household sample as described in Section 2.2. All variables are measured in NOK 1,000s, adjusted to 2014 prices, and reported per household member. Robust standard errors are used in calculation of the confidence intervals.*

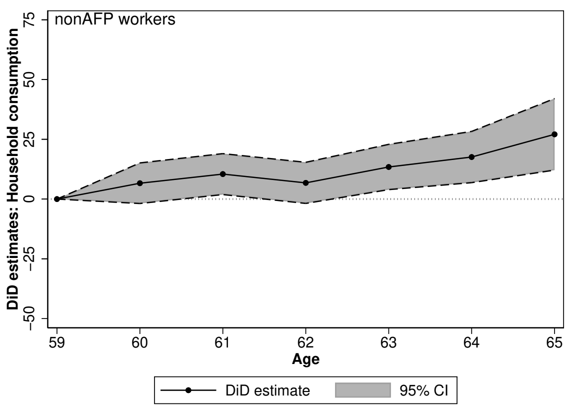
Figure 6: Age-specific DiD estimates for income, savings, consumption, and net wealth - non-AFP workers

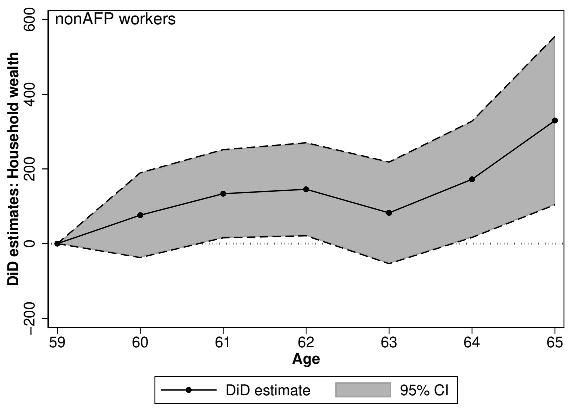




Income

Savings





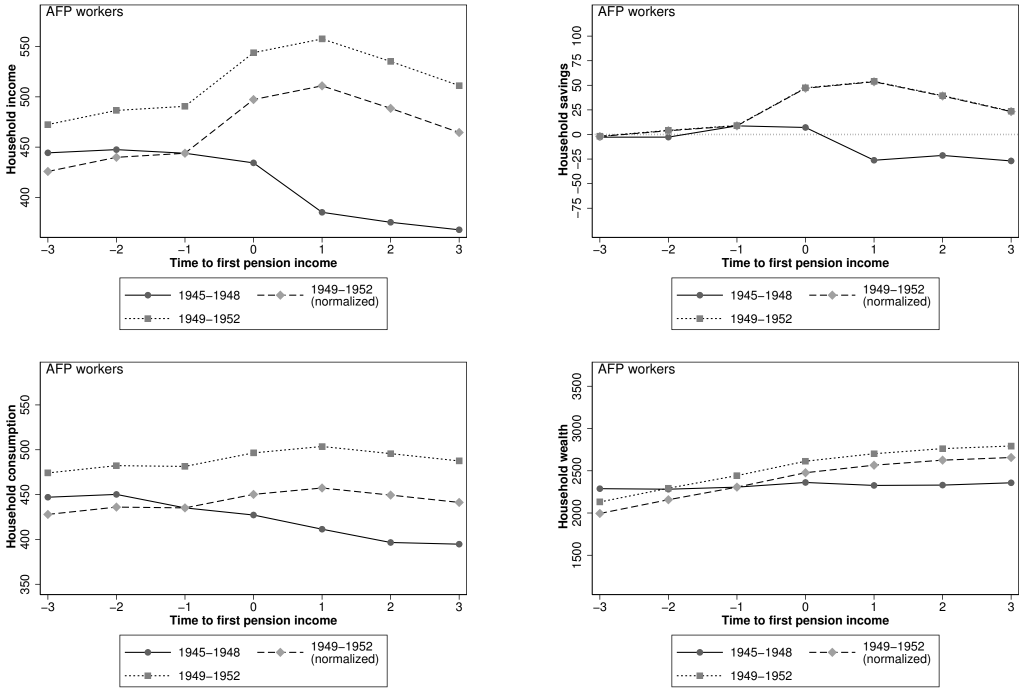
Consumption

Wealth

Notes: This figure plots the estimated β_k coefficients (along with 95% confidence intervals) from the specification in equation (1)*, for non-AFP workers at each age* $k \in \{60, \ldots, 65\}$*. The sample used is the household sample as described in Section 2.2. All variables are measured in NOK 1,000s, adjusted to 2014 prices, and reported per household member. Robust standard errors are used in calculation of the confidence intervals.*

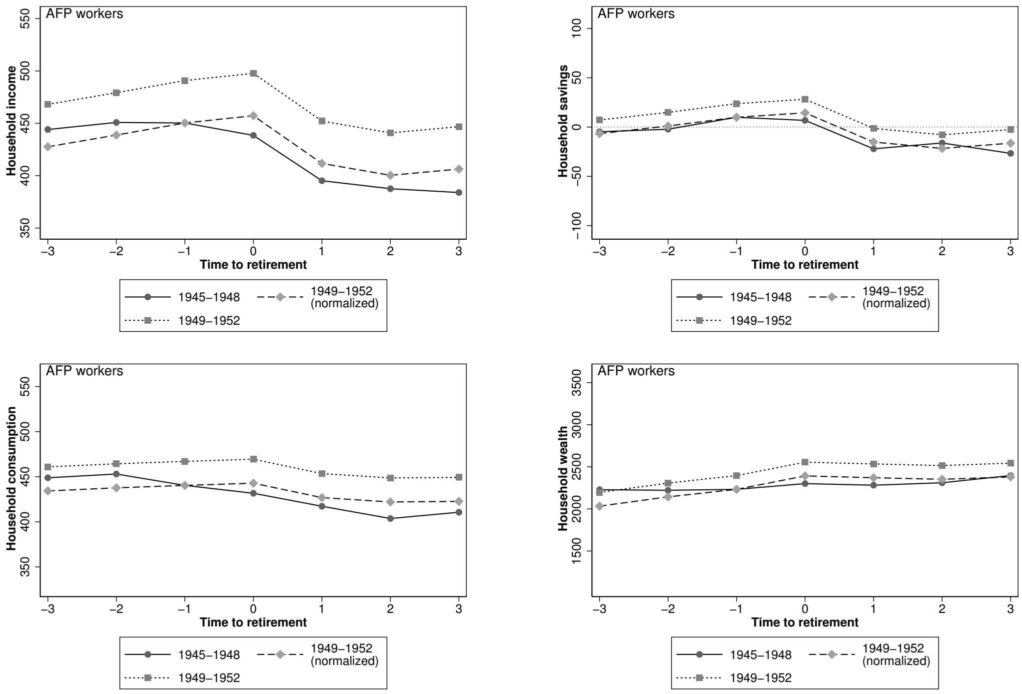
Income, consumption, and savings around retirement

Figure 7: Income, savings, consumption, and wealth by time to first pension income - AFP workers



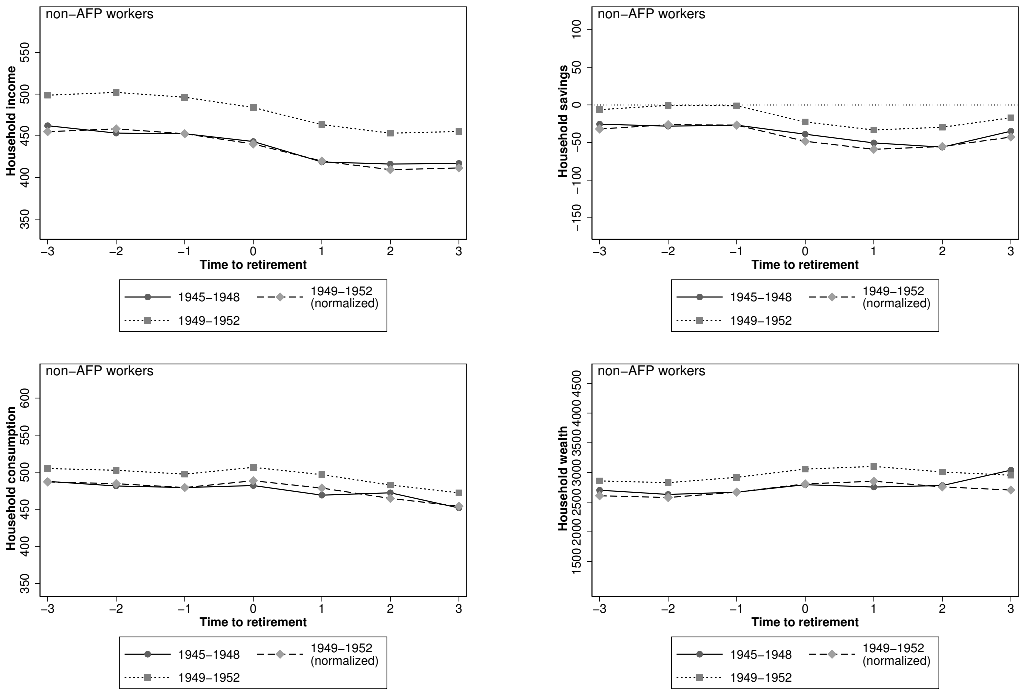
Notes: This figure shows the evolution of household income, savings, consumption, and net wealth by the number of years relative to the first year with pension income, for the households of workers employed by an AFP affiliated firm at age 58 ("AFP workers"). The sample used is the household sample described in Section 2.2, excluding cohorts born after 1952. All variables are reported in NOK 1,000s, adjusted to 2014 prices, and reported per household member. The "1949-1952 (normalized)" series show the means of the respective variables for the post-reform cohorts after normalizing to the pre-reform cohorts level at age 59.

Figure 8: Income, savings, consumption, and wealth by time to retirement - AFP workers



Notes: This figure shows the evolution of household income, savings, consumption, and net wealth by the number of years relative to the first year without an active employment spell at the end of the year, for the households of workers employed by an AFP affiliated firm at age 58 ("AFP workers"). The sample used is the household sample described in Section 2.2, excluding cohorts born after 1952. All variables are reported in NOK 1,000s, adjusted to 2014 prices, and reported per household member. The "1949-1952 (normalized)" series show the means of the respective variables for the post-reform cohorts after normalizing to the pre-reform cohorts level at age 59.

FIGURE 9: Income, savings, consumption, and wealth by time to retirement - non-AFP workers



Notes: This figure shows the evolution of household income, savings, consumption, and net wealth by the number of years relative to the first year without an active employment spell at the end of the year, for the households of workers employed by a firm not affiliated with AFP at age 58 ("non-AFP workers"). The sample used is the household sample described in Section 2.2, excluding cohorts born after 1952. All variables are reported in NOK 1,000s, adjusted to 2014 prices, and reported per household member. The "1949-1952 (normalized)" series show the means of the respective variables for the post-reform cohorts after normalizing to the pre-reform cohorts level at age 59.

Tables

Table 1: Descriptive statistics, pre- and post-reform

	Ages 62-65				Ages 59-61			
	Post		Pre		Post		Pre	
Variable	Mean	(sd)	Mean	(sd)	Mean	(sd)	Mean	(sd)
				AFP workers				
Demographics								
Male	0.72	(0.4)	0.74	(0.4)	0.72	(0.4)	0.73	(0.4)
Married	0.71	(0.5)	0.74	(0.4)	0.68	(0.5)	0.72	(0.4)
Higher education	0.20	(0.4)	0.18	(0.4)	0.22	(0.4)	0.19	(0.4)
Eastern Norway	0.38	(0.5)	0.41	(0.5)	0.37	(0.5)	0.39	(0.5)
Labor market								
Earnings	536.9	(289.5)	471.7	(219.7)	568.9	(301.1)	509.1	(263.1)
Small firm	0.012	(0.1)	0.014	(0.1)	0.012	(0.1)	0.013	(0.1)
Employed in manufacturing	0.31	(0.5)	0.36	(0.5)	0.28	(0.4)	0.33	(0.5)
Experience	23.3	(2.4)	23.3	(2.2)	23.3	(2.7)	23.3	(2.3)
Any sickness benefit	0.21	(0.4)	0.22	(0.4)	0.20	(0.4)	0.22	(0.4)
Household variables								
Net income	488.2	(267.4)	441.9	(316.4)	501.1	(257.5)	452.0	(249.1)
Savings	2.26	(812.3)	13.3	(491.1)	7.55	(331.0)	4.68	(582.7)
Consumption	485.9	(832.6)	428.5	(484.3)	493.5	(374.1)	447.3	(591.7)
Wealth	2204.0	(2843.0)	1808.4	(2222.3)	2211.1	(2728.1)	2007.7	(2491.5)
No. of individuals	42436		41943		46090		63865	
No. of observations	99989		97130		125525		122767	
				nonAFP workers				
Demographics								
Male	0.72	(0.4)	0.72	(0.4)	0.72	(0.4)	0.73	(0.4)
Married	0.73	(0.4)	0.76	(0.4)	0.71	(0.5)	0.75	(0.4)
Higher education	0.24	(0.4)	0.22	(0.4)	0.26	(0.4)	0.23	(0.4)
Eastern Norway	0.41	(0.5)	0.45	(0.5)	0.40	(0.5)	0.43	(0.5)
Labor market								
Earnings	536.7	(305.7)	458.5	(230.9)	568.7	(340.1)	506.3	(292.4)
Small firm	0.32	(0.5)	0.34	(0.5)	0.30	(0.5)	0.33	(0.5)
Employed in manufacturing	0.099	(0.3)	0.11	(0.3)	0.093	(0.3)	0.10	(0.3)
Experience	22.9	(2.9)	22.9	(2.9)	22.8	(3.5)	22.9	(2.9)
Any sickness benefit	0.18	(0.4)	0.18	(0.4)	0.18	(0.4)	0.18	(0.4)
Household variables								
Net income	524.8	(356.8)	511.1	(529.8)	533.2	(306.4)	491.9	(354.9)
Savings	-29.1	(787.2)	34.4	(801.8)	-8.96	(480.2)	-6.30	(563.4)
Consumption	553.9	(810.8)	476.7	(790.8)	542.2	(522.4)	498.2	(608.5)
Wealth	3009.6	(4108.5)	2441.0	(3605.3)	2957.8	(4101.7)	2664.0	(3593.6)
No. of individuals	20369		18504		24017		28964	
No. of observations	46546		41761		61602		54356	

Notes: This table displays descriptive statistics at age 58 for AFP and non-AFP workers. The sample used is the household sample described in Section 2.2. Higher education is defined as any university or college education. A small firm is a firm with no more than 10 employees. Experience is defined as the number of years with labor income larger than 1 Basic Amount between ages 34 and 57. Any sickness benefit is defined as having at least one period with more than 2 weeks of sickness absence in a year. Earnings, net income, savings, consumption, and wealth are measured in NOK 1,000s and adjusted to 2014 prices. Household variables are reported per household member.

Table 2: DiD estimates for net income, savings, consumption, and wealth - AFP workers

	Income		Savings		Consumption		Wealth	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DiD Estimate	39.126***	39.727***	25.364***	25.570***	13.762***	14.157***	2.630	16.282
	(1.314)	(1.156)	(1.403)	(1.399)	(1.546)	(1.453)	(18.078)	(17.373)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Pre-reform mean, ages 62-64	440.83	440.83	-1.08	-1.08	441.91	441.91	2397.55	2397.55
Individuals	122164	122164	122164	122164	122164	122164	122164	122164
Observations	445411	445411	445411	445411	445411	445411	445411	445411

Notes: This table shows the estimated β coefficient from the specification in equation (1)*, for AFP workers. The DiD estimator measures the difference in mean outcomes before and after the reform at ages 62–65 relative to the before-after difference in mean outcomes at ages 59–61. All variables are measured in NOK 1,000s, adjusted to 2014 prices, and reported per household member. Robust standard errors are reported in parentheses.* $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 3: DiD estimates for income components - AFP workers

	Total		Market		Transfers		Other		Tax and neg. transfers	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
DiD Estimate	39.126***	39.727***	9.749***	11.333***	48.227***	47.475***	0.181	0.554	19.032***	19.634***
	(1.314)	(1.156)	(1.765)	(1.494)	(0.760)	(0.748)	(0.966)	(0.958)	(0.783)	(0.642)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Pre-reform mean, ages 62-64	440.83	440.83	384.86	384.86	150.23	150.23	56.78	56.78	151.04	151.04
Individuals	122164	122164	122164	122164	122164	122164	122164	122164	122164	122164
Observations	445411	445411	445411	445411	445411	445411	445411	445411	445411	445411

Notes: This table shows the estimated β coefficient from the specification in equation (1), for AFP workers. The DiD estimator measures the difference in mean outcomes before and after the reform at ages 62–65 relative to the before-after difference in mean outcomes at ages 59–61. All variables are measured in NOK 1,000s, adjusted to 2014 prices, and reported per household member. Robust standard errors are reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4: DiD estimates for savings components - AFP workers

	Total		Liquid Capital		Debt		Other	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DiD Estimate	25.364***	25.570***	19.088***	19.147***	-2.237	-2.250	4.038	4.173*
	(1.403)	(1.399)	(1.545)	(1.543)	(2.137)	(2.137)	(2.521)	(2.521)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Pre-reform mean, ages 62-64	-1.08	-1.08	17.47	17.47	-14.06	-14.06	-32.61	-32.61
Individuals	122164	122164	122164	122164	122164	122164	122164	122164
Observations	445411	445411	445411	445411	445411	445411	445411	445411

Notes: This table shows the estimated β coefficient from the specification in equation (1), for AFP workers. The DiD estimator measures the difference in mean outcomes before and after the reform at ages 62–65 relative to the before-after difference in mean outcomes at ages 59–61. All variables are measured in NOK 1,000s, adjusted to 2014 prices, and reported per household member. Robust standard errors are reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 5: DiD estimates for income, savings, consumption, and wealth - non-AFP workers

	Income		Savings		Consumption		Wealth	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DiD Estimate	27.802***	29.579***	21.018***	21.273***	6.783**	8.306***	78.254**	99.904***
	(2.404)	(2.173)	(2.623)	(2.617)	(2.751)	(2.602)	(39.585)	(38.470)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Pre-reform mean, ages 62-64	476.04	476.04	-21.64	-21.64	497.68	497.68	3193.22	3193.22
Individuals	59268	59268	59268	59268	59268	59268	59268	59268
Observations	204265	204265	204265	204265	204265	204265	204265	204265

Notes: This table shows the estimated β coefficient from the specification in equation (1), for non-AFP workers. The DiD estimator measures the difference in mean outcomes before and after the reform at ages 62–65 relative to the before-after difference in mean outcomes at ages 59–61. All variables are measured in NOK 1,000s, adjusted to 2014 prices, and reported per household member. Robust standard errors are reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 6: DiD estimates for income components - non-AFP workers

	Total		Market		Transfers		Other		Tax and neg. transfers	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
DiD Estimate	27.802***	29.579***	-11.053***	-9.033***	46.148***	46.217***	4.874**	5.605**	12.168***	13.210***
	(2.404)	(2.173)	(2.794)	(2.394)	(1.037)	(1.014)	(2.244)	(2.213)	(1.418)	(1.228)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Pre-reform mean, ages 62-64	476.04	476.04	466.15	466.15	100.64	100.64	88.73	88.73	179.48	179.48
Individuals	59268	59268	59268	59268	59268	59268	59268	59268	59268	59268
Observations	204265	204265	204265	204265	204265	204265	204265	204265	204265	204265

Notes: This table shows the estimated β coefficient from the specification in equation (1)*, for non-AFP workers. The DiD estimator measures the difference in mean outcomes before and after the reform at ages 62–65 relative to the before-after difference in mean outcomes at ages 59–61. All variables are measured in NOK 1,000s, adjusted to 2014 prices, and reported per household member. Robust standard errors are reported in parentheses.* $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 7: DiD estimates for savings components - non-AFP workers

	Total		Liquid Capital		Debt		Other	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DiD Estimate	21.018***	21.273***	13.419***	13.593***	-1.222	-1.449	6.378	6.232
	(2.623)	(2.617)	(3.527)	(3.537)	(4.240)	(4.247)	(4.884)	(4.886)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Pre-reform mean, ages 62-64	-21.64	-21.64	16.71	16.71	-11.42	-11.42	-49.77	-49.77
Individuals	59268	59268	59268	59268	59268	59268	59268	59268
Observations	204265	204265	204265	204265	204265	204265	204265	204265

Notes: This table shows the estimated β coefficient from the specification in equation (1), for non-AFP workers. The DiD estimator measures the difference in mean outcomes before and after the reform at ages 62–65 relative to the before-after difference in mean outcomes at ages 59–61. All variables are measured in NOK 1,000s, adjusted to 2014 prices, and reported per household member. Robust standard errors are reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendices

A Tables

Table A.1: Descriptive Statistics

	Individual Sample		Household Sample	
	Private AFP	nonAFP	Private AFP	nonAFP
Demographics				
Male	0.72	0.72	0.73	0.72
Higher education	0.21	0.25	0.21	0.25
Eastern Norway	0.39	0.43	0.39	0.42
Labor market, age 58				
Earnings	544.14	543.60	538.51	536.42
Small firm	0.01	0.32	0.01	0.32
Manufacturing	0.31	0.10	0.31	0.10
Experience	23.12	22.46	23.27	22.81
Sickness benefit	0.21	0.18	0.21	0.18
Household variables				
Disposable income	490.89	556.78	477.34	527.18
Savings	7.19	0.07	9.85	-0.04
Consumption	483.70	556.71	467.48	527.23
Wealth	2304.14	3546.91	2156.24	3014.36
Individuals	127955	64237	122164	59268
Observations	507182	250439	445411	204265

Notes: This table displays descriptive statistics for AFP and non-AFP workers at age 58. The reported values are averages. In Columns (1) and (2) we use the individual sample and in Columns (3) and (4) the household sample, both described in Section 2.2. Monetary variables are measured in NOK 1'000s and adjusted to 2014 prices. Disposable income, savings, consumption, and wealth are measured at the household level and reported per household member.