Long-Term Effects of Maternity Ward Closures on Children

Milla Hägg* Mika Kortelainen†

December 1, 2025

Abstract

It is widely known that early-life health interventions can affect long-term outcomes later in life. This paper studies how the closures of maternity wards affect short-term health outcomes as well as educational and labour market outcomes of children in their adulthood using a quasi-experimental research design and nationwide administrative data sets from Finland. Using difference-in-differences approach in a staggered design, we find significant improvements in perinatal child health outcomes in the short run. In the long term, we find significant positive effects of closures on upper secondary educational attainment as well as employment and earnings.

JEL classification codes: H75; I11; I18; J13

Keywords: maternity care, closure, merger, healthcare consolidation, long-term outcomes

*University of Helsinki, Helsinki Graduate School of Economics, VATT Institute for Economic Research. Email: `milla.hagg@helsinki.fi`.

†University of Turku, Finnish Institute for Health and Welfare and INVEST Research Flagship Centre. Email: `mika.kortelainen@utu.fi`.

We acknowledge financial support from the Academy of Finland (Decision No. 325110). We are grateful for helpful comments from Maria Koch Gregersen, Daniel Knutsson, Petter Lundborg, Louise Schubert Paaske and Lauri Sääksvuori and as well as participants of the European Association of Labour Economists conference, Nordic Health Economics Study Group meeting and the PhD workshop of the Finnish Association of Health Economists. The statements, findings, conclusions, views and opinions contained and expressed in this paper are based on data from the Finnish Institute for Health and Welfare and Statistics Finland, processed by Statistics Finland. All errors are our own.

1 Introduction

Closing down small and rural hospitals to cut healthcare costs and increase efficiency in the provision of healthcare services has become a common policy globally. Typically, it is believed that healthcare production is associated with significant economies of scale and that increasing the average size of providers through closures of small hospitals and mergers of units can be beneficial for the organization of healthcare services. However, closures may also result in congestion and increased travel time to hospitals, influencing access to care and possibly resulting in deteriorating health outcomes for affected populations. Additionally, hospital and ward closures can be classified as health shocks, which especially in the early life have been documented to shape later outcomes. Although causal evidence on the effect of health shocks on short-term health outcomes has been established in the literature, much less is known about their long-term consequences, particularly the effects of hospital closures on educational and labor market outcomes in a modern healthcare context. We argue that a health shock experienced in the early childhood may be one of the channels affecting long-term outcomes of children. This question is relevant for policies related to the organization and provision of healthcare, as policy-making is often focused on efforts to reduce costs while overlooking broader social and economic impacts.

This paper sheds light onto the possible long-term effects of hospital closures, which are often overlooked in assessing cost containing policies. Our empirical strategy takes advantage of Finnish maternity ward closures. With the rigorous administrative data available, we are able to track patient flows and determine accurately the geographical areas that were affected by ward closures to form treatment and control groups in a difference-in-differences (DiD) design. We use the DiD estimator of Callaway and Sant'Anna (2021) to account for heterogeneity in treatment effects and different timing of treatments. We observe the universe of births in Finland and have administrative data on socioeconomic and health-related variables for mothers and children. We are able to follow the children born in the 1980's and 1990's up to year 2018 and observe whether they have obtained a upper secondary school diploma as well as their labor market status and earnings.

With a sample of data on births in Finland, we compare the affected treatment areas to unaffected control areas. In the short term, we look at child health outcomes and find significant positive health effects. We then proceed to study whether the closures affect the long-term outcomes of children, including education attainment and earnings and find significant effects in education attainment. The net effect of closures may be driven by heterogeneous effects in different areas. We follow the empirical strategy of Avdic, Lundborg, and Vikström (2024), where a ward closure will affect namely two areas. Firstly around the closing wards, there are mothers who need to travel to a different ward to give birth. These areas around the closing wards constitute the closure areas. Secondly, the remaining wards will receive an inflow of new patients that are coming from the closure areas. Mothers not living in closure areas, but giving birth in the remaining hospitals may thus also be affected. The areas around the remaining hospitals constitute the inflow areas. The net effect is given by the combination of effects in closure and inflow areas. The mothers are assigned to the areas depending on where they live, not depending on the ward they give birth in.

Our results suggest that closures result in statistically significant improvements in child health outcomes. There are improvements in both the closure and inflow areas, but only closure area effects are statistically significant. This suggests children may have benefited from the closures, for example, through a higher skill level of physicians and other staff in the remaining wards. After closures, there are more overall births taking place yearly in the remaining wards, but the results suggest this has not lead to congestion. The long-term results are consistent with the so-called early-life health shock hypothesis: there are statistically significant improvements in upper secondary education diploma attainment in closure areas. Employment at age 20 exhibits a statistically significant net increase and employment at age 25 a statistically insignificant net decrease. Disposable income at age 25 also shows a significant positive effect. These results suggest that there may be long-term effects policymakers should take into account when making closure decisions concerning provision of specialized healthcare services and questions on access to care.

This paper relates to three main strands of literature. First, it is to our knowledge one of the first papers to explore the relationship between access to maternity care and long-term effects in the modern context. Historically, healthcare unit openings have been documented to have significant long-term effects on labor and educational outcomes. For example, long-term effects have been studied in the context of maternity ward openings by Lazuka (2021), who find that opening maternity wards increased the share of hospital births and reduced early neonatal mortality in the short run. It also improved labour income, employment, schooling as well as reduced health-related disabilities in the long run. The empirical setting utilizes a reform which improved access to care in Sweden between 1931 and 1946. Bütikofer, Løken, and Salvanes (2019) study the effect of openings of child healthcare centers in Norway from the 1930s on and find that improved access to care increases completed years of schooling by 0.15 years and earnings by 2 percent. We are not aware of such studies on child healthcare unit closures. Although the above two studies show substantial improvements in both short- and long-term, one may argue the setting to differ from today substantially. To our knowledge, our paper is the first to provide evidence on the long-term effects of early-life health shocks caused by specialized healthcare unit closures in modern times: our closures take place in the 1990's.

Second, our paper relates to the short-term effects of maternity ward closures on infant health. Although closures may often be justified by concerns about patient safety, earlier research has suggested they may also have other effects on health. In all, evidence from the earlier literature on the short-term effects of maternity ward closures is divided. Avdic et al. (2024) study the causal effects of maternity ward closures in Sweden on maternal and infant health after the birth. They find the net effect of the closures to be negative for mothers, and positive for children. Grytten, Monkerud, Skau, and Sørensen (2014) use a similar setting of maternity ward closures in Norway, but instead study the effect of regionalization and local hospital closures on neonatal and infant mortality. The study finds no significant effects of hospital type on neonatal or infant mortality. Maternity ward closures and C-sections are also studied by Battaglia (2024), who finds rural maternity ward closures to create net benefits in quality of care for mothers.

Third, this paper relates to the long-term effects of early life health shocks. Although there is a multitude of factors that affect educational and labor market outcomes later in a child's life, the relationship of health shocks and long-term schooling and labor market outcomes has been well documented in earlier literature (Almond & Doyle, 2011; Almond, Doyle Jr, Kowalski, & Williams, 2010; Bharadwaj, Løken, & Neilson, 2013; Sievertsen & Wüst, 2017). This literature is mostly concentrated on treatments other than hospital or ward closures or openings, but we argue that a perinatal health shock resulting from a ward closure can be one of the channels affecting a child's educational and labor market outcomes.

This paper proceeds with section 2 on the institutional background. Section 3 describes the data used for analysis. In section 4, the empirical strategy is introduced, and section 5 reports our main results. Section 6 concludes the paper.

2 Institutional background

The number of maternity wards in Finland has been decreasing over the past 20 years. In 1987, there were 53 operating wards, whereas in 2019 only 24 wards were left. The closures have been driven by the concern for patient safety due to a decreasing number of births as well as the need to reduce costs. Figure 1 illustrates the development of the locations and numbers of maternity wards in Finland. The circles indicate wards that were operating in June 2019. The triangles indicate wards that have been closed over time. The four maps show the locations of the wards between 1987–2018. Many of the wards closed in the so-called first wave of closures were in Northern parts of Finland or located relatively near other existing wards. These are the closures studied in this paper. In the second and third waves, the closures were parts of more structured centralization policies associated with requirements of a sufficient number of yearly births in wards.

In the 1970's, approximately 70 percent of all births in Finland took place in university hospitals or central hospitals. Up until then, the main tool in improving patient safety had been advocating hospital births instead of home births. After that, there were two main objectives: regionalization and centralization. The main idea of regional-

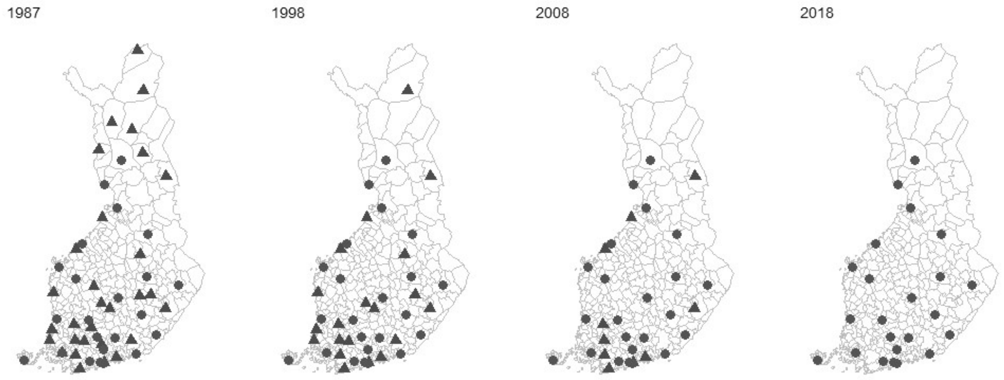


Figure 1: Locations of the maternity wards in 1987, 1997, 2007 and 2018

ization was to integrate smaller hospital units to the national healthcare service, making different parts of the care pathway accessible to everyone. Centralization policies were about improving patient safety. The aim was no longer only to prevent maternal deaths, but also reduce the risk of neonatal deaths. This often meant adopting new technologies related to childbirth, but also having a more diverse set of specialists in the same hospital. (Viisainen, 2000)

Until 1990, the funding of the hospitals was regulated by the government through a funding program. The funding covered 39-65 percent of all healthcare expenses and was granted directly to hospitals. The aim was to ensure the quality of and access to care were similar in all parts of Finland. After 1993, these subsidies were capped and hospitals no longer received direct funding from the government, creating cost-cutting incentives for the municipalities in charge of funding healthcare services. This may have strengthened the status of smaller wards. It may have been in the interest of local politicians to fund the smaller wards more generously to ensure sufficient access to healthcare services. (Viisainen, 2000)

Nowadays maternity ward closures are often justified by cost savings, as maternity wards are a particularly costly part of a hospital. In Sweden in 2017 they accounted for 4 percent and in the US for 3.5 percent of all healthcare spending (Avdic et al., 2024; Torio & Moore, 2016). There are no cost approximations available from Finland, but child births and procedures related to it are among the most common procedures in inpatient care in Finnish hospitals with maternity wards (HUS, 2022). Hospitals with maternity wards are legally required to have immediate access to specialists in obstetrics and anesthesiology, who can perform emergency procedures, which makes offering maternity ward services costly for the hospital. The savings accumulated from closing a maternity ward have been estimated to be approximately 4 million euros annually, but they may be substantially larger: closing a ward will often cause hospitals to re-evaluate the need for other services, such as elective surgeries. (Nieminen, 2015)

Along with cost reductions, another justification for unit closures has been patient safety. If there is not a sufficient amount of annual births in a ward, it may have adverse effects on the skill level of physicians. Only a small fraction of all births require immediate medical attention due to complications: for example in 2018, only 0.8 percent of births required emergency cesarean sections, which is why the number of annual births can matter for maintaining the skill level of the physicians. By centralizing the care involving these procedures, the aim could be either or both minimize costs and maximize patient safety. In Finland, births mostly take place in the hospital. In 2017, there were a total of 50 151 births recorded and out of these births, only 243 occurred outside the hospitals as either before arrival to the hospital, at home unplanned or at home planned. Mothers typically give birth in hospitals nearest to them, although there are exceptions: If a mother is evaluated to be especially at risk, they may be redirected to another maternity ward typically situated in a university hospital.

The maternity care system in Finland is extensive and reaches most of the expecting mothers. Approximately 99.8 percent of all pregnant mothers receive prenatal care from a prenatal clinic, and there is little variation in the extent of care provided geographically. Care is in general perceived as reliable and accessible. Indicators, such as maternal mortality, also support this: between 2011-2020 there was a total of 14 maternal deaths associated with pregnancy, delivery and the perinatal period (Finnish Institute for Health and Welfare, 2022). Prenatal care is based on national treatment recommendations and laws. The aim is to have nationally equal care available for all pregnant mothers. Prenatal care is offered in prenatal clinics, which are operated by municipalities. During pregnancy, mothers are encouraged to visit a prenatal clinic, where they are offered the services of a nurse or midwife and a doctor specialized in prenatal care. The first visits are scheduled around pregnancy weeks 8–10. These visits are especially important for prenatal screening, which helps to detect risks in the pregnancy. First-timers are offered at least nine visits and mothers that have given birth at least once at least eight visits. These visits include an extensive health check for the whole family as well as two doctor visits. First-timers are also offered a home visit from a nurse or midwife around the 30th week of pregnancy. A nurse visits all mothers within a week after being discharged from the hospital after giving birth and in addition there is a follow-up within 5–12 weeks of childbirth. If at any stage there are abnormalities in the course of the pregnancy or the health of the mother or the fetus, the nurse or doctor can refer the mother for further examination. These examinations take place at a maternity clinic, which are generally located in larger hospitals with a maternity ward. Factors that may contribute to the pregnancy being risky include a high or low BMI, old or young maternal age, substance abuse, various diseases and conditions, lack of social support, genetics, previous caesarean sections or a history of obstetric complications such as miscarriages, neonatal deaths or stillbirths (Attilakos & Overton, 2012; Dhanjal, 2012).

Whereas prenatal care is a part of primary care and is offered in hospital units of all size, perinatal care and childbirth are managed by maternity wards in larger hospital units. According to Finnish law, a hospital can have a maternity ward, if they have a sufficient amount of midwifes and staff to assist in emergency surgeries, along with the required facilities and equipment. They must also be able to monitor the health of the fetus, infant and the mother and evaluate their need for care. The patients should have immediate access to specialists in obstetrics and anesthesiology or physicians specialized in other fields, but thoroughly familiar with obstetrics or anesthesiology. The hospital should also be able to provide a pediatrician or a physician with good knowledge of pediatrics and a possibility to receive advice from a specialized pediatrician. If the maternity ward in question is a centralized unit for high-risk mothers, the hospital also needs to have a physician specialized in neonatal care.

This paper focuses on the effect of the closures on the short-term health an long-term educational and labor market outcomes of children. We use upper secondary education as our main outcome, because the FOLK data module provided by Statistics Finland does not distinguish between pupils that have completed the mandatory comprehensive school and those who have not. In Finland, the law from 1968 ensures children must attend school until age 16. In 2020, mandatory schooling was extended to age 18. Although attending school was mandatory in our sample only until the age of 16, it is very common to proceed to upper secondary education after comprehensive school. For example, in 2017, 97 percent of all ninth grade graduates went on to upper secondary school the next year (Statistics Finland, 2022a). Students graduate from upper secondary school at the age of 18 or 19. After a year from graduating from upper secondary education, 26 percent were employed full-time, 56 percent were studying either full-time or part-time, 6 percent were unemployed and 12 percent classified as others. After a year from vocational upper secondary education, 57 percent were working full-time, 17 percent were studying full-time or part-time, 17 percent were unemployed and 9 percent were classified as others. (Statistics Finland, 2022b)

3 Data

This study is based on administrative patient-level register data from the Medical Birth Register of Finnish Institute for Health and Welfare (THL). The complete Medical Birth Register covers all births in Finland and our sample spans years from 1987 to 2014. The register gathers all the live births of infants with birth weight of over 500 grams and stillbirth where the fetus is over 22 weeks of gestational age.

The data set used in this study is a subsample (n=494 090) of the register data set (N=1 830 070). The time frame of births is restricted to 1987-1994. The closure wards and their closest inflow wards are presented in Table 1 for years 1988-1991. Whereas nowadays patients are often directed to multiple wards around a closing ward, in the 90's they were more typically assigned to just one remaining ward. We are able to verify this by tracking the patient flows from postal codes in closure catchment areas to remaining

Table 1: Maternity wards in sample

Closed ward	Type	Mean births	Closed	Inflow #1	Type	Distance (km)	Inflow #2	Type	Distance (km)
Riihimäki	RH	253	1990	Hyvinkää	RH	18	Kanta-Häme	CH	54
Pieksämäki	HC	233	1991	Varkaus	RH	49	Mikkeli	CH	76
Valkeakoski	RH	382	1991	TAYS	UH	32			
Ähtäri	RH	256	1991	South Ostrobothnia	CH	68			

wards as demonstrated in Appendix A. Typically, the closest wards are also in the same hospital district. The closure of Pieksämäki health center is an exception, where the closest ward in Varkaus is not within the same hospital district and the closest ward in the same district is Mikkeli Central Hospital.

The locations of the mothers are from population grid data provided by Statistics Finland. The average travel distance to the ward for the mothers is 37 kilometers. All distance measures used are geodesic and may therefore differ from the actual distances travelled to the maternity wards. To form the catchment areas, the municipal reforms over the time period need to be taken into account. The disbanded municipalities are merged with the remaining municipalities in the data. The difference-in-difference method requires observations both before and after the treatment. In this case, as the municipalities have ceased to be, they would not have both types of observations. By forming larger municipality areas and merging the data from different postal codes, the sample could be kept as representative as possible. The final data does not include births from catchment areas of closed wards in Sodankylä (1988), Pello (1988) and Kittilä (1989) or their inflow wards. This is because the register data is only available from 1987 and thus the lack of a pre-period.

The Medical Birth Register is the main data source for health outcomes and control variables. The short-term outcome variables in this study reflect acute changes in health, determined as little as possible by pre-birth health or behaviour in pregnancy. Essentially, all of the included outcomes are such that can be improved with the appropriate measures during childbirth. They are therefore suitable proxies for quality of care. (OECD, 2011) Based on this, the chosen child health outcomes are the APGAR score at 1 minute, indicators of low and high APGAR scores and perinatal mortality.

APGAR scores are commonly used standardized scores to assess the health of a newborn immediately after birth. The APGAR scores are based on the heart rate, promptness and vigor of the first respiratory efforts, and reflex response to certain stimuli, muscle tone, and color of the infant. The highest possible score is 10 and it represents the optimum condition of the infant. (Apgar, Holaday, James, Weisbrot, & Berrien, 1958). As a general rule, a score of 7 or above is considered excellent, whereas a score under 7 hints of moderate abnormality in the health of the infant. A score under 3 is considered very low. (Casey, McIntire, & Leveno, 2001).

The child health outcomes are ready-made diagnostic variables recorded by the health authorities. Recording of ICD-10 diagnoses associated with child birth began officially in 2004, and the diagnostic data from our sample period is therefore not reliable enough. Due to the quality of the diagnose data during the sample period, we do not study some otherwise commonly used diagnostic outcomes, such as lacerations of mothers or complications of children (OECD, 2017; Russo & Andrews, 2009). Whereas the recording of the ready-made variables into the register seem to not have significant differences between different areas within Finland, it is likely they are not directly comparable to results from other countries or to results from after 2004. Another assumption we must make is that diagnostic practices do not change over time within our sample period. This is a plausible assumption, as the sample period is relatively short.

Table 2 shows the descriptive statistics for the whole sample as well as control, closure and inflow. For the short-term perinatal health measures, the average APGAR 1-minute score was 8.7. Only 2 percent of the newborns had a score under 6 and 77 percent scored above 7. Perinatal mortality indicates whether a child has been either born dead or died within the first week. In general, perinatal mortality is not a very good measure for quality of care or even health in developed countries. In the sample, the share of children born dead or dying within a few weeks of birth was only 0.2 percent.

The health-related controls for the mothers are their parity and the number of visits to a prenatal ward. In our sample, the average number of pregnancy-related visits to the doctor is approximately 15, where as visits to a maternity polyclinic vary in regions between 2 to 3 visits. The parity denotes whether a mother has given birth before. The results could also be estimated with a sample of mothers giving birth the first time, but as approximately 60 percent of the mothers in the whole data set have given birth before, the sample would get insufficiently small for estimation purposes. The average distance to the closest ward in the total sample is 37 kilometers. For the closure areas this average is roughly 17 kilometers. The average stay at the maternity ward is 5.4 days, of which the majority (on average 4.6 days) takes place after the mother has given birth. The length of stay is on average longer in closure than inflow areas.

The Birth Register data is merged with Statistics Finland's FOLK data sets on socioeconomic control variables and the long-term child outcome variables. The control variables for the mothers include earnings, marital status, education and first language. The education level indicates the highest degree the mother has attained. The home language is a dummy indicating whether the mother speaks some other language than Finnish or Swedish as their first language. The variables are from the year the mothers gave birth, except for earnings which are from the previous year. The mothers in the sample are on average 29 years old and 77 percent of them are married. 80 percent of mothers have a upper secondary school diploma. As many mothers give birth multiple times, they may be included in the sample more than once with the control variables from each year they give birth.

The long-term educational and labor market outcomes include obtaining an upper secondary education diploma by the age of 20 as well as a variety of labor market participation outcomes and earnings outcomes at the ages of 20 and 25. We regard a person to have obtained an upper secondary education diploma, if they are recorded to have it by the age of 20 in the Statistics Finland FOLK 2020 personal data module. An upper secondary education diploma can be from upper secondary general school or an upper secondary vocational school. We also look separately at the outcome of matriculating from the upper secondary general school. In our sample the share of upper secondary education degrees is on average approximately 87 percent, which is slightly different to the national statistics (Statistics Finland, 2022a) mentioned before. There may be some students that have enrolled in upper secondary education, but not graduated, which could explain the difference. For the labor market outcomes, we regard a person to

Table 2: Descriptive statistics for sample 1987-1994

	All	Control	Closure	Inflow
Maternal characteristics				
Age	28.7	28.7	28.3	28.6
Married (%)	76.5	76.2	78.9	77.7
Upper secondary school degree (%)	80.3	79.9	80.9	81.7
Taxable income in year prior to childbirth	11 318	11 432	9961	10 919
General care specific indicators				
Distance to ward	37.0	38.1	17.4	30.3
Length of stay	5.4	5.3	5.6	5.5
Length of stay after delivery	4.6	4.6	5.3	4.7
Pregnancy-related visits	15.2	15.1	15.7	15.2
Visits to maternity polyclinic	2.3	2.4	2.9	2.0
Pregnancy and delivery specific characteristics				
First-timers (%)	39.7	39.8	36.2	39.3
Earlier births	1.0	1.0	1.1	1.0
Miscarriages (%)	23.9	23.9	21.0	23.9
Earlier pregnancies	1.45	1.45	1.47	1.42
Abortions (%)	6.9	6.9	9.3	6.8
Ectopic pregnancies (%)	0.8	0.8	0.7	0.8
Earlier Caesarean sections (%)	5.6	5.8	4.3	4.8
Mother under 18 (%)	0.5	0.5	0.6	0.5
Mother over 35 (%)	10.3	10.5	8.7	9.9
Care for risk of prematurity (%)	1.7	1.7	1.0	1.9
Care for high blood pressure (%)	2.4	2.4	1.4	2.6
Placenta praevia (%)	0.2	0.2	0.1	0.1
Birth weight	3 557	3 553	3 638	3 569
Short-term health outcomes				
APGAR 1 min	8.7	8.7	8.9	8.6
APGAR 1 min > 7 (%)	92.2	92.4	94.3	91.1
APGAR 1 min ≤ 7 (%)	7.8	7.6	5.7	8.9
Perinatal mortality (%)	0.3	0.3	0.1	0.3
Long-term outcomes				
Upper secondary education degree (%)	86.6	86.3	88.2	87.7
Upper secondary general degree (%)	48.2	48.5	43.8	47.2
Employment at 20 (%)	82.8	83.6	82.0	81.8
Employment at 25 (%)	81.3	81.1	82.2	82.2
Employment income at 25	20 699	20 640	20 674	20 974
Entrepreneurial income at 25	12 271	12 031	12 617	13 277
Disposable income at 25	18 086	18 055	17 996	18 240
Total debt at 25	34 799	34 328	34 711	36 981
Number of births	494 090	400 260	8 147	85 683

have been employed in a given year, if they have received either employment income or entrepreneurial income in that year. We look at different types of earnings including disposable income, earned employment income, earned entrepreneurial income and earned total income in state taxation which also includes capital income. We separately look at indebtedness measured in terms of total debt. It should be noted the data on earnings and employment data is measured on a yearly basis, so there may be some measurement error in terms of distinguishing between workers and non-workers, especially regarding the students. A student working part-time or just during the summer months will be considered as employed in our data.

4 Empirical strategy

The effects of maternity ward closures are studied through a staggered difference-in-difference setting, which has been adopted from Avdic et al. (2024). With the DiD setting, we aim to analyze the policy intervention of closing certain maternity wards and its effect on mothers giving birth and children being born in the closure and inflow areas. We study the differential effect of an intervention on two groups: the unaffected control group and the affected treatment groups. The difference between these two groups yields the net effect. The key identifying assumption in any DiD model is that the groups being compared would have parallel trends without the treatment. When trends are similar in both groups, the treatment will result in a single-sided deviation from the trend and dependent variables of interest vary on an aggregate level. This is the canonical DiD model. For example, the closures of maternity wards may result in health varying between different areas after the closures, but not within the areas due to the closure.

We start with a baseline model, where we study the net effect of closure by comparing treatment (closure and inflow) areas to control areas. We first study the short-term health outcomes of the children, and then proceed to study their long-term educational and labor market outcomes. Equation 1 represents the canonical two-way fixed effects (TWFE) estimator for both short-term and long-term outcomes. The baseline estimation equation could be thus written as

$$y_{iadt} = \alpha + \beta_C C_{at} + \lambda_a + \lambda_t + (t \times \lambda_d) + X'_{it}\beta_X + \epsilon_{iadt}, \tag{1}$$

where in the y denotes outcomes. The outcomes y for short-term health of children are APGAR 1 minute score, high APGAR, low APGAR and mortality. The long-term outcomes of the children are upper secondary education diploma attainment, employment at ages 20 and 25, disposable income, employment income and entrepreneurial income.

C_{at} indicates whether a given catchment area was subject to closure at time $t \geq T_c$, where T_c was maternity ward closure year. λ_a denote catchment area fixed effects and λ_t yearly fixed effects. $t \times \lambda_d$ are hospital district-level time trends.

In the controls X'_{it} for the short-term, we include the mother's age, marital status, education attainment of mother and ethnicity as well as pre-birth health characteristics of parity and number of visits to the maternity clinic during pregnancy. Optimally, we would also add health characteristics such as information about smoking and obesity, but unfortunately these are not available for the sample period used in this study. Pre-birth health characteristics are not included as controls in the long-term regressions.

It should be emphasized that the children are not assigned to control and treatment groups depending on the ward they were born in, but according to the catchment area where they live when they are born. This enables us to account for changes in the patient mix, because it allows us to study the patients from the same areas before and after the closures (Avdic et al., 2024). Additionally, it supports the stable unit treatment value assumption (SUTVA), that implies all spillovers are modeled. One concern in our setting is related to the mothers possibly giving birth to their multiple children in different wards. If the mother does not move between giving birth to her children, the births will all be assigned to a single treatment or control group.

In the second-stage, we study the heterogeneity of treatment effects in different areas. As mentioned, closure and inflow areas may be affected by the closures differently, so we study the effects in the different areas separately. Again, closure areas are the catchment areas of the closing maternity ward, and inflow areas are the catchment areas of the maternity ward that will accommodate the mothers coming from the closure area. There are control areas unaffected by maternity ward closures. One area will be only subject to one treatment over the time period studied and will also remain classified in the same control or treatment group over the entire period.

Using TWFE with a staggered setting and heterogenous treatment effects can lead to biased estimates (Callaway & Sant'Anna, 2021; De Chaisemartin & d'Haultfoeuille, 2020; Goodman-Bacon, 2021; Sun & Abraham, 2021). The effects in our case may be dynamic - remaining wards may not be able to adjust to changes immediately, but they could do so over time. Therefore, we argue that using early-treated areas as controls for later-treated areas would lead to a bias. To account for these factors, we use the Callaway-Sant'Anna (CS) estimator proposed by Callaway and Sant'Anna (2021). This is also where our study differs from Avdic et al. (2024), as they use earlier and later treated observations as controls and utilize the decomposition of Goodman-Bacon (2021). The CS estimator takes into account variation in treatment times, multiple time periods and heterogeneity in treatment effects. The estimator is essentially an inverse propensity weighted difference in cohort average treatment effects that are observed between the treated and untreated units for a cohort. Since the causal parameters the estimator yields allow for heterogeneity and dynamic effects, problems in the interpretation of two-way fixed effects regression with a staggered setting can be avoided. In our model, the units treated will stay such until the end of the period. The controls are never treated in the sense that they do not become treated within our sample period. Some of the wards in the control group may be closed eventually, but not within an anticipation period of three years after the end of our sample.

The parameter in question is a group-time average treatment effect (hereby ATT), that we observe for group g at time t. The groups are defined by the time period of the treatment. The group-time ATT gives us the ATT for a cohort of the units treated at the same point in time. The estimator is given by

$$ATT(g,t) = \mathbb{E}[(\frac{G_g}{\mathbb{E}[G_g]} - \frac{\frac{p_g(X)C}{1-p_g(X)}}{\mathbb{E}[\frac{p_g(X)C}{1-p_g(X)}]})(Y_t - Y_{g-1})], \quad (2)$$

where Y denotes the outcome, X the covariates and p propensity score. G_g is a dummy and will equal one when the unit is in group g, where g is the year of treatment. The denominator in the first term in the equation therefore gives us the average of G_g. Similarly, C is a dummy for being a part of the control group. It should be noted that the estimator only uses the base year as the pre-period: the other years before treatment are not included in the ATT estimation. Indeed, one of the features of the CS estimator is, that causal effect parameters can be identified in the staggered setting even when there are differences in observed characteristics of the groups, which may create non-parallel outcome dynamics. This means that the estimator does not technically require common pre-trends, but instead relies on the assumption that the trends would develop similarly if there was no treatment at all. (Callaway & Sant'Anna, 2021) However due to the treatment existing, the pre-trends are the only thing we observe and thus we also check for them and find no significant differences. With the never-treated controls, an equivalent estimator would be the one suggested by Sun and Abraham (2021). Another alternative is the estimator of Borusyak, Jaravel, and Spiess (2024), which has slightly stronger assumptions on the parallel trends than the CS estimator. With the limited data period we are able to observe before closures, we opt out of using of it.

With the group-time ATTs, we can also formulate an aggregated ATT. Essentially, this sums the group-time ATTs given by estimator (2). In our results, we will show single overall treatment effect summary parameters and aggregated parameters for event-study analysis. Simply put, the aggregation parameters are given by the form

$$\theta = \sum_{g \in G} \sum_{t=2}^{T} w(g,t) \cdot ATT(g,t), \tag{3}$$

where $w(g,t)$ are weighting functions. There are several possibilities for aggregation, but we are interested in an overall aggregation of the treatment effect. This can be achieved with a parameter taking the average of the identified group-time ATTs together.

This parameter is of the form

$$\theta^O_{sel} = \sum_{g \in G} \theta_{sel}(g) P(G = g | G \leq T), \tag{4}$$

where $\theta^O_{sel}(g)$ is the average effect of participating in the treatment for units in group g, where g is again defined by the treatment year. It computes the average effects for each group across all time period and then averages these effects. The parameter does not therefore place more weight on the groups that are presented in the sample longer or are treated for a longer period. This aggregated parameter can be interpreted in the same way as the standard DiD estimators in a canonical DiD setting. (Callaway & Sant'Anna, 2021)

The estimations are run with Stata and the `csdid` package. We cluster standard errors at the hospital district level to account for autocorrelation within the hospital district. Clustered standard errors are used to account for certain variables being possibly similar with each other within same regions.

Endogeneity of the closures is a potential concern in the identification. It is quite evident the closures are not decided upon completely randomly: the closed wards are typically smaller in size compared to the remaining wards. However, in the sample period, it could be argued the timing of treatment is more random. As we are looking at the first wave of closures, there are still many small wards with similar characteristics operating, which may as well have been closed instead of the closures we observe.

Another common concern in DiD settings is the stable unit treatment value assumption. As mentioned, part of the threat to the SUTVA is solved by assigning the mothers and children to the treatment group based on their place of residence rather than the wards they give birth in, but it does not rule out cases where the mothers moves from one location to another in between giving birth to her children. This is a weakness we are unable to tackle perfectly. It could be addressed to some degree by restricting the analysis to only first-born children, but it would result in a very small sample making inference otherwise unreliable.

5 Results and discussion

This section presents the results. In each table, we show the net ATT from the baseline regression and the separate ATTs for the closure and inflow areas. Each cell of the results shows the ATT, standard error and percentage difference from the baseline mean. The aim of looking at the different areas separately is to further understand the mechanism behind the changes in health: how the net effect is driven by differences in changes in closure and inflow areas.

5.1 Health outcomes

One can think of the maternity ward having an optimal number of births n^*. This n^* assures that learning-by-doing knowledge is at a sufficient level, which will be enough to minimize risks to health. If the number of births is under n^*, more yearly births will result in increased learning-by-doing and thus improve health outcomes. Similarly, it is possible that once the number of births exceeds n^*, the patient caseload becomes too large and begins to have negative effects on health. This could be caused by a number of reasons. Possible suggestions include overcrowding of hospital spaces or a too large number of patients per midwife or obstetrician. Therefore, the final net health effects will be determined by which group of patients the closures affect more and through which channels the effects are realized. The strength of different mechanisms could also vary in time, which is studied in the event study Figures 2, 3 and 4.

For the short-term health outcomes of children presented in Table 3, there is a statistically significant improvement in APGAR 1 minute scores. The net effect of magnitude 0.016 is driven by improvements in the closure areas. The closure areas experience an average improvement of 0.131, which corresponds to a 1.5 percentage point increase in relation to the baseline mean. The improved scores of the closure areas may hint of better quality of care for the newborns. The coefficient for the inflow areas is insignificant and smaller in magnitude, but positive. Again, improvements in the inflow wards may imply increased learning-by-doing and that there is no congestion in the ward although more patients are coming in. Before closures, the wards mothers

Table 3: Average treatment effects for child short-term health outcomes

	ATT		
	Total	Closure	Inflow
APGAR 1 min	0.016* (0.010) +0.2%	0.131*** (0.029) +1.5%	0.015 (0.018) +0.2%
APGAR 1 min>7	0.005 (0.005) +0.5%	0.061*** (0.014) +6.6%	0.005 (0.005) +0.5%
APGAR 1 min≤7	-0.005 (0.005) -6.4%	-0.061*** (0.014) -78.2%	-0.005 (0.005) -6.4%
Perinatal mortality	0.0001 (0.0001) +3.3%	-0.0007 (0.0004) -30.0%	-0.0002 (0.0001) -6.7%

NOTE Finnish data for period 1987-1994. Estimated using the CS aggregated estimator. Cell shows the estimator value, standard error and control mean difference. Each cell is a separate regression. Standard errors are clustered at the hospital district level. *p<0.1 **p<0.05 ***p<0.01.

from the closure areas gave birth in had an average of 324 yearly births, which increased to an average of 482 after the closures (see Appendix B). This means mothers moved to significantly larger hospitals.

Panel (a) in Figure 2 shows the event ATTs for closure areas on the left and inflow areas on the right. The closure areas exhibit a large increase in the ATT after the first year, which slowly decreases with time. A possible explanation for the lag is that mothers from the closure areas may still have been treated in the closed wards for the most part of their pregnancy, although they go give birth in an inflow ward. The effects seem to persist after the first treatment period for the closure areas more so than for inflow areas, where the ATT increases after the first year, but fluctuates slightly above zero over time.

The APGAR score spans from 0 to 10, so it is also important to understand which scores have increased. A score of 7 and above is considered excellent, so an increase in very good scores instead of the lowest scores may hint of inefficiencies. For policy, it would most likely be more important to increase the worst scores instead of making good scores even better. The likeliness of an APGAR 1 minute score of above 7 has increased especially in the closure areas, where the coefficient implies an 6.6 percent increase compared to the control mean. Mechanically, this change is also seen in the scores of 7 and below. The decrease compared to the control mean is very large, but this is due to relatively few babies receiving a score of 7 or below: the control mean is 7.7 percent. From Panel (b) in Figure 2, we see the dynamics closely follow those of Panel (a), although the ATT for getting a low or high APGAR score is positive and significant already in the treatment year in the closure areas. The coefficients level off over time after the initial increase.

As a robustness check, we use the birth weight of children as the outcome variable. Where the mother gives birth should not affect the birth weight of the child. Thus we would expect to see no significant effects of ward closures on the birth weight. Table 4 confirms this, as the net, closure or inflow area results are insignificant and the changes very small.

Table 4: Average treatment effects for birth weight

	ATT		
	Total	Closure	Inflow
Birth weight (g)	-8.594 (6.891) -0.2%	-9.438 (5.705) -0.3%	-7.988 (4.457) -0.2%

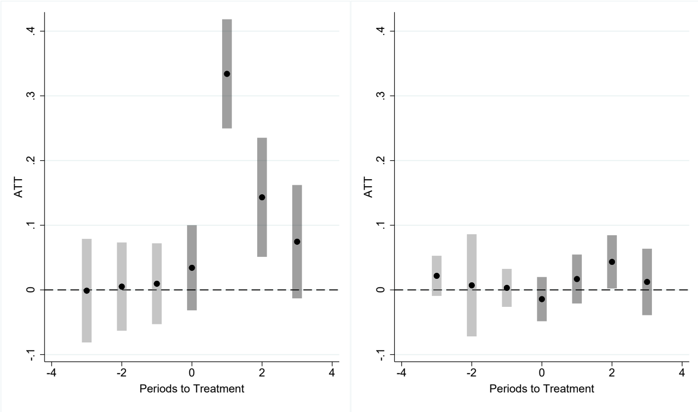
NOTE Finnish data for period 1987-1994. Estimated using the CS aggregated estimator. Cell shows the estimator value, standard error and control mean difference. Each cell is a separate regression. Standard errors are clustered at the hospital district level. *p<0.1 **p<0.05 ***p<0.01.

Our findings are mostly aligned with the findings of Avdic et al. (2024). They find significant reductions in fetal distress and infant birth trauma and insignificant positive net effects on APGAR 1-minute scores. The 1-minute scores are positively affected in the closure area and negatively affected in the inflow area. APGAR 5 minute and 10 minute score effects are negative, but small and insignificant. They also study the mechanisms behind the health effects and find closures leading to a larger number of births per midwife, which may be an indication of increased congestion in remaining wards. They do not find significant effects with distance. Since we do not have data on the number of midwifes in the wards, we are not able to comment on changes in the congestion. Figure 6 presented in the Appendix C shows the changes in distance were modest. We mainly observe does not seem to be a significant factor in the Finnish case either.

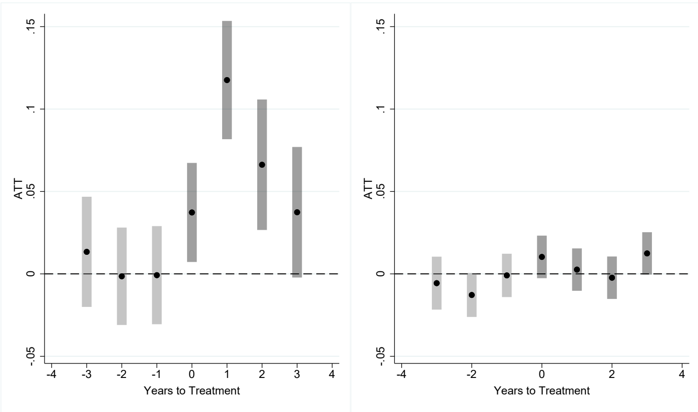
Perinatal mortality does not seem to be affected. As seen in Table 3 and in Appendix D, the mortality outcome is not significant for the closure or inflow areas. None of the ATTs are significant and seem to fluctuate around zero. Our finding of no significant changes to neonatal mortality is aligned with those of Grytten et al. (2014) from Norway, who also do not find hospital closures to be associated with changes in neonatal or infant mortality rates. In general, perinatal mortality can be considered an extreme measure of health in countries with developed healthcare systems.

When considering the short-term improvements in health outcomes, it is worthwhile to discuss what the margin being estimated is. As mentioned, the closure wards usually do not accommodate at-risk mothers, as they are designated to give birth at larger hospitals. It is possible some of these mothers are present in the sample, as they may have travelled to give birth at for example the inflow wards already pre-closure, and this is partially confirmed by the patient flows presented in Appendix A. However, since the treatment is assigned based on the area where the mothers live and not the ward, this does not pose a threat to the identification. If this was to induce a bias, assigning at-risk-mothers to wards would have to somehow change between the pre-period and post-period. We argue this is unlikely: the same procedures concerning at risk mothers were in place both before and after closures.

a



b



c

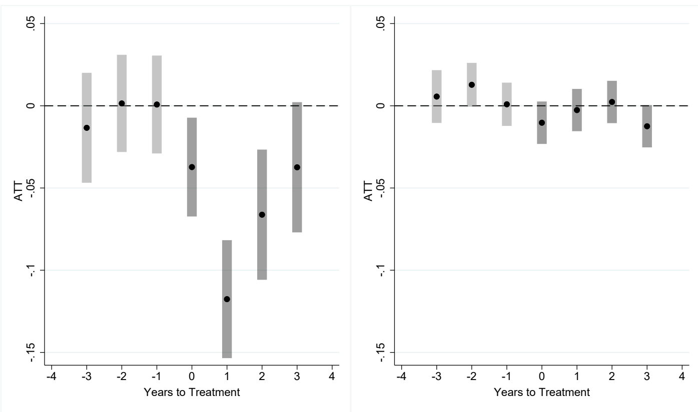


Figure 2: Aggregated event ATTs for closure and inflow areas **Panel a**: APGAR score at 1 minute, **panel b**: high APGAR score at 1 minute, **panel c**: low APGAR score at 1 minute

5.2 Educational and labor market outcomes

The results for child long-term outcomes are presented in Table 5. If we were to expect that the improvements observed in the short-term health outcomes would also channel into the long-term outcomes for children, we would expect to see a positive effect on schooling or labor market outcomes. It should be noted that the results do not imply a direct causal link between short-term health shocks and long-term socioeconomic outcomes, but the health shocks observed are one potential channel that long-term outcomes can be affected through. That is to say, we study the short-term and long-term effects of closures, not directly the effect of the short-term health shocks on the long-term outcomes.

For obtaining an upper secondary school diploma, we observe a sizeable and significant positive net effect and positive effects for closure areas. The results are shown in Table 5 and the event studies in Figure 3. A student that has completed upper secondary education has either a diploma from upper secondary general school or upper secondary vocational school. The net effect for upper secondary education is positive with a 2.2 percent increase compared to the baseline mean. It should be noted this is a considerable increase, as obtaining an upper secondary degree is very common in Finland. The magnitudes of the effects translate somewhat directly from the short-term to the long-term. Both in the short- and long-term the magnitude of the outcomes are larger in the closure areas. There do not seem to be similar effects in the probability to complete upper secondary general education. The effects are non-significant and close to zero as shown in Appendix 8. It would also be interesting to study the effects on the probability of completing tertiary (university) education. Unfortunately, we are only able to follow the children until the age of 25 and can only observe if they have graduated by then. With enrollment information, this kind of analysis would also be possible to conduct.

The effects for employment at ages 20 and 25 differ. The average starting age for tertiary education studies in Finland is 24 (OECD, 2019). The differences in employment outcomes at ages 20 and 25 could be explained by upper secondary graduates working

Table 5: Average treatment effects for long-term outcomes of children

	ATT		
	Total	Closure	Inflow
Education			
Upper secondary education diploma	0.019** (0.08) +2.2%	0.018* (0.010) +2.1%	0.002 (0.004) +0.2%
Labor market			
Employment at 20	0.005** (0.002) +0.6%	0.004 (0.004) +0.5%	0.005** (0.002) +0.6%
Employment at 25	-0.001 (0.001) -0.1%	-0.001 (0.001) -0%	-0.001** (0.001) -0%
Disposable income	170.5** (60.2) +0.9%	21.9 (168.2) +0.1%	216.2*** (63.3) +1.2%
Employment income	-190.7 (156.2) -0.9%	-270.2 (483.5) -1.3%	-112.6** (162.6) -0.5%
Entrepreneurial income	1 215.4 (945.0) +10.0%	2 950.5 (2 789.1) +23.3%	1 012.9 (987.5) +7.6%

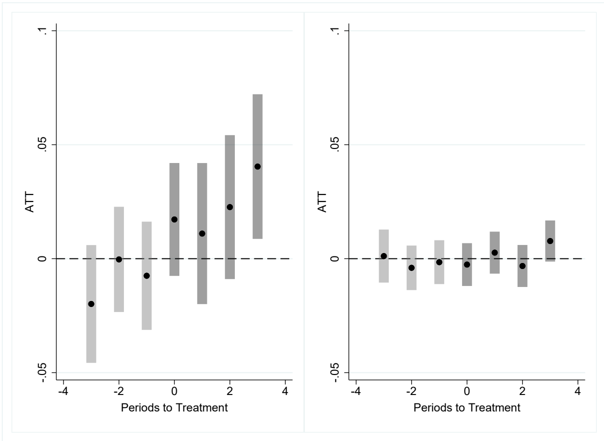
NOTE Finnish data for period 1987-1994. Estimated using the CS aggregated estimator. Cell shows the estimator value, standard error and control mean difference. Each cell is a separate regression. Standard errors are clustered at the hospital district level. *p<0.1 **p<0.05 ***p<0.01.

before entering tertiary education. Of course, being employed at age 25 may also hint of several different things. On one hand, the person may have already graduated from upper secondary or higher education and began working. On the other hand, they may have skipped higher education altogether. The economic implications are very small: the increase is under 1 percent. Interpreting the effects on employment is not quite as straightforward as education, and what may hinder the interpretation of is the definition of employment used. Employment is determined by whether you have received employment earnings or entrepreneurial earnings within the year. A large share of the individuals in the sample are employed at least in some form.

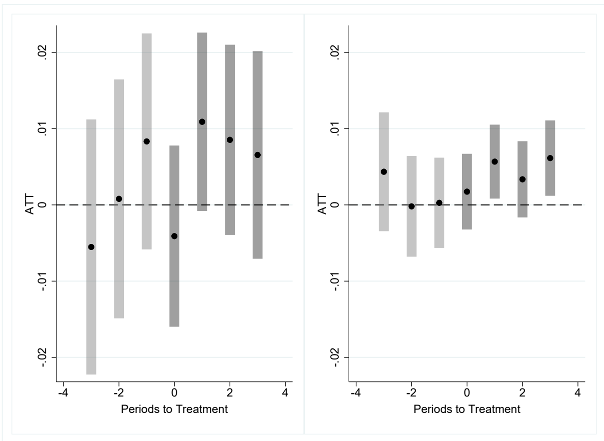
Whereas changes in employment may be difficult to interpret, different types of earnings may shed light onto the long-term outcomes. We compare outcomes of disposable income with employment and entrepreneurial earnings. Disposable income includes both earnings and social benefits, which are granted to for example students. We observe a positive effect on disposable income in both areas at age 25, and it is also significant in the inflow areas. Employment earnings at age 25 display a negative effect, but it is not significant. This is consistent with what we observe in our employment outcomes: if the children studied enter university studies around the age of 25, it would mean their employment income should decrease.

Of course, interpreting the results without information on hours worked is difficult. A smaller coefficient may imply less demanding work and thus lower wages. In the absence of the information on how much the individuals work, it could also imply smaller workloads in the closure areas and thus lower wages. It is difficult to state clearly what the economic intuition behind the result is. The results for disposable income at age 20, total taxable income and total debt are presented in Appendix 9. Disposable income at 20 shows some positive effects for inflow area and close to zero insignificant effects for closure areas. The total taxed income and total debt do not display any significant effects.

d



e



f

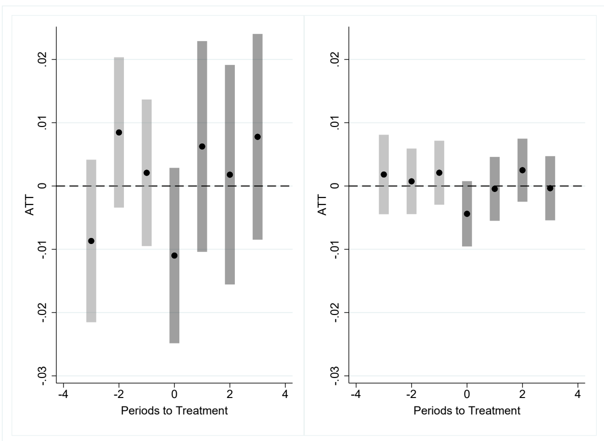
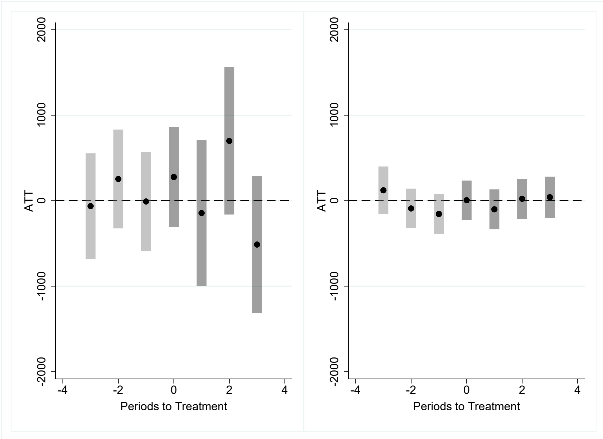
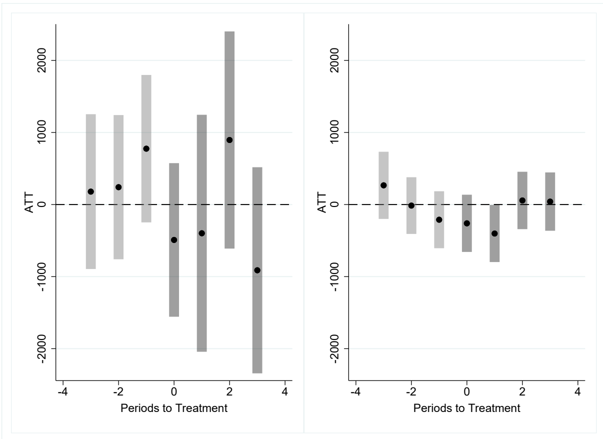


Figure 3: Aggregated event ATTs for closure and inflow areas **Panel d**: upper secondary education, **panel e**: employment at age 20, **panel f**: employment at age 25

g



h



i

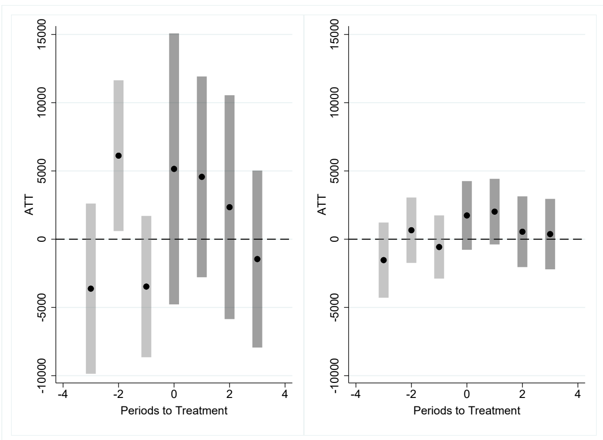


Figure 4: Aggregated event ATTs for closure and inflow areas **Panel g**: disposable income at age 25, **panel h**: employment income at age 25, **panel i**: entrepreneurial income at 25

Even though the results of the long-term outcomes can be explained within the early-life hypothesis framework, interpretations of causality should be careful. Although we are able to reliably control for socioeconomic variables concerning both parents and children, there may also be other factors affecting the long-term labor and educational outcomes of the children. For example, we are unable to fully observe the medical history of the children. In addition to early life health shocks, health shocks in other parts of life as well as other, completely unobservable shocks, may also matter for labor and educational outcomes.

Although our paper provides some evidence there may be long-term effects to ward closures, it should still be emphasized that the sample we use is from the 90's. Implications of ward closures today may result in different long-term outcomes. For example nowadays maternity wards in Finland are required by law to have 1 000 yearly births for the ward to keep operating. Therefore, the larger wards may be seriously congested as is and become even more congested as nearby wards close. Indeed to infer something about closures today, we still need to understand the mechanisms behind the effects of closures on both short-term and long-term outcomes better.

For policy-making, it would also be interesting to better understand the mechanisms behind the health shocks inflicted on mothers and children. With current data, we are however unable to reliably assess why it is we see the effects we see. The most important factor currently lacking is information on diagnoses and the staffing in the wards. As we are unable to control for how many doctors or midwifes were present, it is difficult to pin down whether the effects are due to congestion or changes in travel times to remaining wards. We leave this for future research.

6 Conclusions

Although maternity ward closures have become more common, we lack an understanding of the effects the closures may have on other factors than costs. This paper focuses on assessing the long-term implications closures may have on children, which have not been studied in the modern context. We use maternity ward closures that have occurred in the 1990's in Finland and study the short-term health outcomes and the long-term educational and labor market outcomes for children.

The results offer an insight into the previously unstudied long-run effects of specialized healthcare unit closures. To be able to evaluate the full effects of closure policies, which often aim at minimizing costs or improving patient safety, it is essential to also understand the effects they may have on health and socioeconomic outcomes in the long-term. Our evidence, in the light of the early-life health shock hypothesis, suggests there may be a link between long-run labor market and educational outcomes and short-term health shocks.

Our results show that closures had positive effects for the children in the short-term. We observe significant reductions in complications and improvements in the APGAR score, which is a standardized measure for the condition of a newborn infant 1 minute after birth. The improvements were larger in the closure areas than in inflow areas. This could be explained by access to better care, as the closure wards are smaller measured in yearly births than the inflow wards. Larger number of births can be seen to improve the skill level of physicians.

In the long-term, according to the early-life hypothesis, we would expect the positive health shocks to result in improvements in schooling and labor market measures. Our results show increases in attainment of an upper secondary education diploma in both closure and inflow areas. Whereas attaining upper secondary education can straightforwardly be interpreted as a positive improvement, understanding the effects of closures on employment and earnings is slightly more difficult. Our results point to increased employment at 25 and increased earnings, especially in the inflow areas. Understanding this mechanism would require more work.

The results in our paper are aligned with those in a paper closest to ours by Avdic et al. (2024). They find positive health effects in the short-term for children and the results are driven by improvements in closure areas. The closures resulted in the mothers and infants being moved to larger wards and albeit those wards being further away, on average the quality improved.

Our paper sheds light on a topic that has previously been unstudied. It is the first to show there may be long-term effects on children's educational and labor market outcomes from closing maternity wards. The results indicate that when forming policies about closures of specialized healthcare units, costs are not the only outcome that should be considered. Policy-makers need to also pay attention to short-term health outcomes as well as the effect of these short-term outcomes may have on the long-term educational and labor market outcomes.

References

Almond, D., & Doyle, J. J. (2011). After midnight: A regression discontinuity design in length of postpartum hospital stays. *American Economic Journal: Economic Policy*, *3*(3), 1-34.

Almond, D., Doyle Jr, J. J., Kowalski, A. E., & Williams, H. (2010). Estimating marginal returns to medical care: Evidence from at-risk newborns. *The Quarterly Journal of Economics*, *125*(2), 591–634.

Apgar, V., Holaday, D. A., James, L. S., Weisbrot, I. M., & Berrien, C. (1958). Evaluation of the newborn infant-second report. *Journal of the American Medical Association*, *168*(15), 1985–1988.

Attilakos, G., & Overton, T. G. (2012). Antenatal care. *Dewhurst's Textbook of Obstetrics & Gynaecology*, 42-49.

Avdic, D., Lundborg, P., & Vikström, J. (2024). Does health care consolidation harm patients? evidence from maternity ward closures. *American Economic Journal: Economic Policy*, *16*(1), 160–189.

Battaglia, E. (2024). The effect of hospital closures on maternal and infant health. *American Journal of Health Economics*, *forthcoming*.

Bharadwaj, P., Løken, K. V., & Neilson, C. (2013). Early life health interventions and academic achievement. *American Economic Review*, *103*(5), 1862–91.

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event study designs: Robust and efficient estimation. *Review of Economic Studies*, forthcoming.

Bütikofer, A., Løken, K. V., & Salvanes, K. G. (2019). Infant health care and long-term outcomes. *Review of Economics and Statistics*, *101*(2), 341–354.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230.

Casey, B. M., McIntire, D. D., & Leveno, K. J. (2001). The continuing value of the apgar score for the assessment of newborn infants. *New England Journal of Medicine*, *344*(7), 467–471.

De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–2996.

Dhanjal, M. K. (2012). Pre-conception counselling. *Dewhurst's Textbook of Obstetrics & Gynaecology*, 35-41.

Finnish Institute for Health and Welfare. (2022). *Perinataalitilasto - synnytt˙*

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of econometrics*, *225*(2), 254–277.

Grytten, J., Monkerud, L., Skau, I., & Sørensen, R. (2014). Regionalization and local hospital closure in Norwegian maternity care—the effect on neonatal and infant mortality. *Health Services Research*, *49*(4), 1184–1204.

HUS. (2022). *Tilinpäätös ja Toimintakertomus 2021.*

Lazuka, V. (2021). It's a long walk: Lasting effects of maternity ward openings on labour market performance. *The Review of Economics and Statistics*, 1–47.

Nieminen, K. (2015). Pienten synnytyssairaaloiden tulevaisuus. *Selvityshenkilön raportti, Sosiaali- ja terveysministeriön raportteja ja muistioita 2015:35*.

OECD. (2011). *Health at a glance.* Retrieved from `https://doiorg/101787/health_glance-2015-en`

OECD. (2017). *Health at a glance.* Retrieved from `https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-2017/obstetric-trauma_health_glance-2017-38-en`

OECD, D. (2019). *Society at a glance 2019.* ORGANIZATION FOR ECONOMIC.

Russo, C. A., & Andrews, R. M. (2009). *Potentially avoidable injuries to mothers and newborns during childbirth, 2006.* Agency for Healthcare Research and Quality.

Sievertsen, H. H., & Wüst, M. (2017). Discharge on the day of birth, parental response and health and schooling outcomes. *Journal of Health Economics*, *55*, 121–138.

Statistics Finland. (2022a). *Immediate transition to further studies by those having ended 9th grade of comprehensive school in, 2000-2020.* Retrieved from `https://pxnet2.stat.fi/PXWeb/pxweb/en/StatFin/StatFin__kou__khak/statfin_khak_pxt_11fy.px/table/tableViewLayout1/`

Statistics Finland. (2022b). *Main type of activity of completers of qualifications one year after graduation, 2007-2020.* Retrieved from `https://pxnet2.stat.fi/PXWeb/`

pxweb/en/StatFin/StatFin__kou__sijk/statfin_sijk_pxt_111l.px/table/tableViewLayout1/

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199.

Torio, C. M., & Moore, B. J. (2016). National inpatient hospital costs: the most expensive conditions by payer, 2013. statistical brief #204.

Viisainen, K. (2000). Choices in birth care-the place of birth.

Appendix

A Patient flows

Table 6: Patients of catchment area for Riihimäki Hospital (closed 1990)

	1987	1988	1989	1990	1991	1992	1993	1994
Riihimäki Hospital	**236**	**234**	**240**	**216**	**0**	**0**	**0**	**0**
Kanta-Häme Central Hospital	**197**	**185**	**188**	**282**	**501**	**513**	**437**	**494**
Hyvinkää Hospital	**<10**	**23**	**29**	**14**	**24**	**28**	**32**	**23**
Etelä-Karjala Central Hospital	0	0	0	0	0	<10	<10	0
Forssa Hospital	0	0	0	0	<10	<10	<10	0
Helsinki University Hospital	<10	13	13	<10	10	<10	<10	<10
Iisalmi Hospital	0	0	0	0	0	0	<10	0
Jokilaakso Hospital	0	<10	0	0	0	0	0	0
Jorvi Hospital	0	<10	<10	0	<10	0	<10	<10
Keski-Suomi Central Hospital	0	0	0	0	0	<10	<10	0
Kuopio University Hospital	0	<10	0	0	0	<10	0	0
Kuusankoski Hospital	0	<10	0	0	<10	0	0	0
Kymenlaakso Central Hospital	0	<10	0	0	<10	<10	<10	0
Kätilöopisto	<10	<10	12	0	<10	<10	<10	0
Lohja Hospital	0	<10	<10	0	<10	<10	<10	0
Loimaa Hospital	0	0	0	<10	0	0	<10	0
Mikkeli Central Hospital	0	0	<10	0	0	0	<10	0
Mänttä Hospital	0	0	<10	0	<10	0	0	0
Oulu University Hospital	0	0	<10	0	0	<10	0	0
Pohjois-Karjala Central Hospital	0	0	0	0	<10	<10	<10	0
Päijät-Häme Central Hospital	0	0	<10	0	<10	<10	<10	0
Salo Hospital	0	<10	<10	0	0	<10	0	0
Satakunta Central Hospital	0	0	0	0	<10	0	0	0
Savonlinna Central Hospital	<10	0	<10	0	0	0	0	0
South Ostrobothnia Central Hospital	0	0	<10	0	<10	0	<10	0
Tampere University Hospital	<10	<10	10	<10	<10	<10	<10	<10
Turku HC	0	0	<10	0	<10	<10	0	0
Turku University Hospital	0	<10	<10	0	<10	0	<10	0
Valkeakoski Hospital	0	0	<10	0	0	0	0	0
Vammala Hospital	0	0	0	0	0	<10	<10	0
Varkaus Hospital	0	0	<10	0	0	0	0	0

Table 7: Patients of catchment area for Pieksämäki Hospital (closed in 1991)

	1987	1988	1989	1990	1991	1992	1993	1994
Pieksämäki Hospital	**222**	**147**	**139**	**162**	**133**	**0**	**0**	**<10**
Mikkeli Central Hospital	**36**	**86**	**97**	**81**	**111**	**198**	**172**	**126**
Varkaus Hospital	**10**	**22**	**12**	**14**	**14**	**11**	**39**	**109**
Etelä-Karjala Central Hospital	0	0	0	0	<10	0	0	0
Helsinki University Hospital	0	<10	0	<10	<10	<10	<10	<10
Hyvinkää Hospital	0	0	<10	0	0	<10	0	0
Jorvi Hospital	0	0	<10	0	0	0	0	0
Kainuu Central Hospital	0	0	<10	0	0	0	0	0
Kanta-Häme Central Hospital	0	0	0	<10	0	0	0	0
Keski-Suomi Central Hospital	<10	<10	<10	0	0	0	<10	<10
Kuopio University Hospital	11	16	21	<10	<10	<10	<10	<10
Kuusankoski Hospital	<10	<10	0	0	0	0	0	0
Kätilöopisto	0	0	<10	0	<10	0	0	0
Kymenlaakso Central Hospital	0	0	0	0	<10	0	0	0
Lohja Hospital	0	0	<10	0	0	<10	0	0
Mänttä Hospital	0	<10	0	0	0	0	0	0
Pohjois-Karjala Central Hospital	0	<10	<10	0	<10	0	<10	0
Päijät-Häme Central Hospital	0	<10	<10	0	<10	0	0	0
Salo Hospital	0	<10	0	0	0	<10	0	0
Savonlinna Central Hospital	0	<10	<10	<10	0	<10	0	0
Turku HC	0	<10	0	0	0	0	0	0
Vaasa Central Hospital	0	0	0	0	<10	0	0	0

Table 8: Patients of catchment area for Valkeakoski Hospital (closed in 1991)

	1987	1988	1989	1990	1991	1992	1993	1994
Valkeakoski Hospital	**225**	**249**	**234**	**269**	**249**	**0**	**0**	**0**
Tampere University Hospital	**223**	**238**	**209**	**220**	**250**	**502**	**494**	**448**
Forssa Hospital	0	0	0	0	0	0	<10	0
Helsinki University Hospital	0	<10	<10	<10	0	0	<10	<10
Hyvinkää Hospital	0	0	0	0	<10	0	0	0
Jorvi Hospital	0	0	<10	0	<10	<10	<10	0
Kanta-Häme Central Hospital	0	<10	<10	<10	<10	<10	<10	<10
Keski-Suomi Central Hospital	0	<10	0	0	0	<10	0	0
Kuopio University Hospital	0	0	<10	0	<10	<10	0	0
Kymenlaakso Central Hospital	0	<10	0	0	<10	0	0	0
Kätilöopisto	0	<10	<10	0	0	0	0	0
Lohja Hospital	0	<10	<10	0	0	0	0	0
Loimaa Hospital	0	0	0	0	<10	0	0	0
Länsi-Pohja Central Hospital	0	0	0	0	<10	0	0	0
Oulaskangas Hospital	0	0	0	0	<10	0	0	0
Pohjois-Karjala Central Hospital	0	0	0	0	0	<10	0	0
Porvoo Hospital	<10	0	0	0	0	<10	0	0
Päijät-Häme Central Hospital	0	<10	<10	0	<10	<10	<10	0
Riihimäki Hospital	0	0	<10	0	0	0	0	0
Satakunta Central Hospital	0	<10	<10	0	0	0	0	0
South Ostrobothnia Central Hospital	0	0	0	0	<10	0	0	0
Turku HC	0	0	0	0	0	0	<10	<10
Turku University Hospital	0	<10	<10	<10	0	0	0	0
Vakka-Suomi Hospital	0	0	0	0	0	0	<10	0
Vammala Hospital	0	0	<10	0	0	<10	<10	0
Ähtäri Hospital	0	<10	0	0	0	0	0	0

Table 9: Patients of catchment area for Ähtäri Hospital (closed in 1991)

	1987	1988	1989	1990	1991	1992	1993	1994
Ähtäri Hospital	**143**	**176**	**142**	**124**	**70**	**0**	**0**	**0**
South Ostrobothnia Central Hospital	**28**	**30**	**31**	**37**	**76**	**133**	**121**	**109**
Forssa Hospital	0	0	<10	0	0	0	0	0
Helsinki University Hospital	0	0	<10	0	<10	0	0	0
Jokilaakso Hospital	0	0	0	0	0	<10	0	0
Jorvi Hospital	0	<10	<10	0	<10	<10	0	<10
Kainuu Central Hospital	0	0	0	0	0	<10	0	0
Keski-Pohjanmaa Central Hospital	0	<10	0	0	0	<10	0	0
Keski-Suomi Central Hospital	<10	<10	<10	0	<10	<10	<10	0
Kittilä Health Center	<10	0	0	0	0	0	0	0
Kätilöopisto	0	<10	0	0	<10	0	0	0
Mänttä Hospital	<10	12	21	40	38	51	41	36
Oulu University Hospital	0	0	0	0	0	0	<10	0
Raahe Health Center	0	0	0	<10	0	0	0	0
Rauma Hospital	<10	0	0	0	0	0	0	0
Riihimäki Hospital	0	0	<10	0	0	0	0	0
Salo Hospital	0	<10	<10	0	0	<10	0	0
Satakunta Central Hospital	0	0	0	0	<10	0	0	0
Selkämeri Hospital	0	0	0	0	0	<10	0	0
Tampere University Hospital	40	26	49	42	39	35	42	57
Turku University Hospital	0	<10	0	0	<10	0	0	0
Vaasa Central Hospital	0	0	0	<10	<10	<10	0	0
Vammala Hospital	0	0	0	0	0	0	<10	0

B Average births in wards by group

Figure 5 depicts the average yearly births in the wards the mothers from different groups give birth in. Average of the ward size for the control group patients is shown in a dashed line and shows little change over the years studied. The short and long dashed line indicates the average ward size of patients in the inflow ward catchment areas. A small increase around the closure years can be observed. The dotted and dashed line indicates the average ward size for patients in closure ward catchment areas, which seems to also increase after closures. This indicates that patients from closures areas travel to go give birth in larger hospitals.

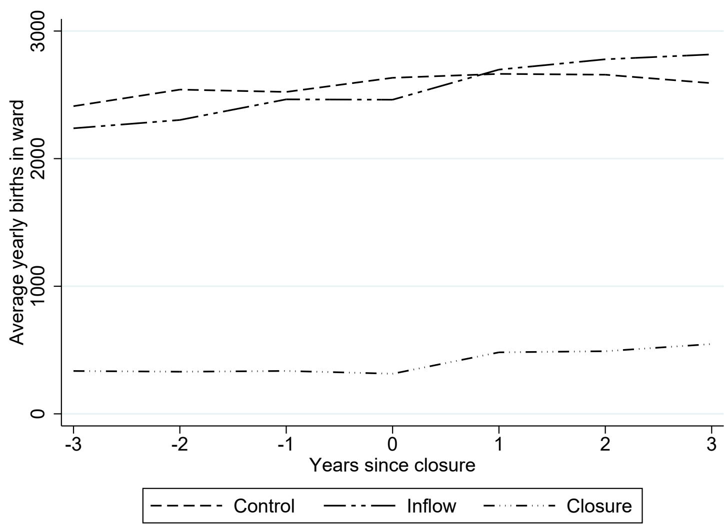


Figure 5: Average ward size by group

C Distance

Figure 6 depicts the average distance to the ward by treatment groups in an event study. Control group is shown in a dashed line and shows relatively small change in average distance over the time period studied. The short and long dashed line indicates average distances travelled by patients in the inflow ward catchment areas. The dotted and dashed line indicates average distances for patients in closure ward catchment areas. For the closure group, it seems the average distance increases slightly after the closures occur. This observation is aligned with the rationale for ward closures. The closures studied are inflicted on small wards in small communities, that only cater to relatively few mothers. The mothers have to on average travel longer distances after closures, but the distance still remains very moderate.

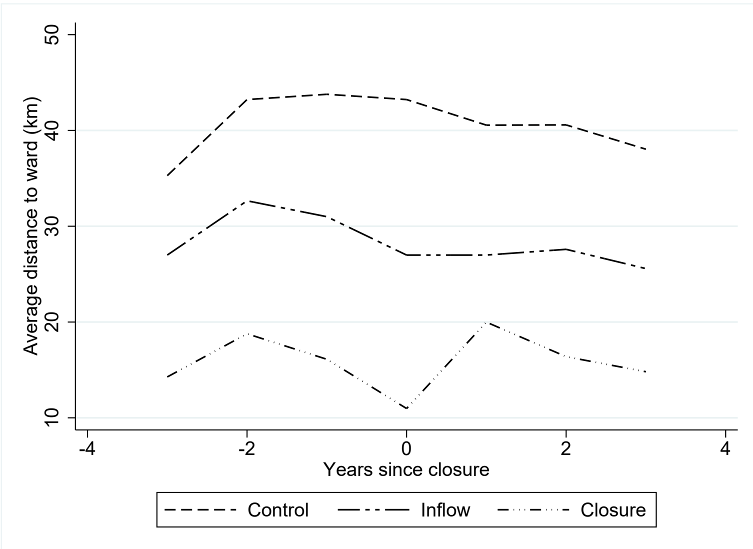


Figure 6: Average distance of patients to maternity ward by group

D Additional event studies

D.1 Health outcomes: perinatal mortality

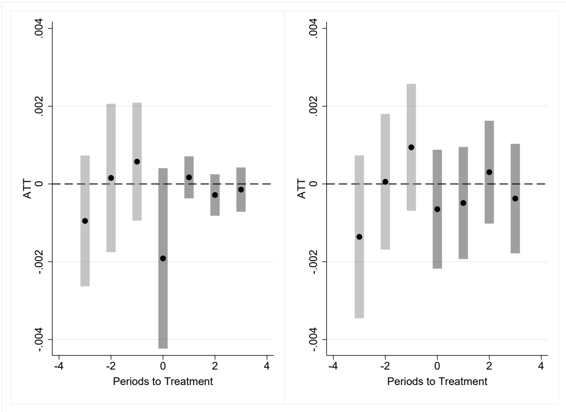


Figure 7: Perinatal mortality in closure and inflow areas

D.2 Educational outcomes

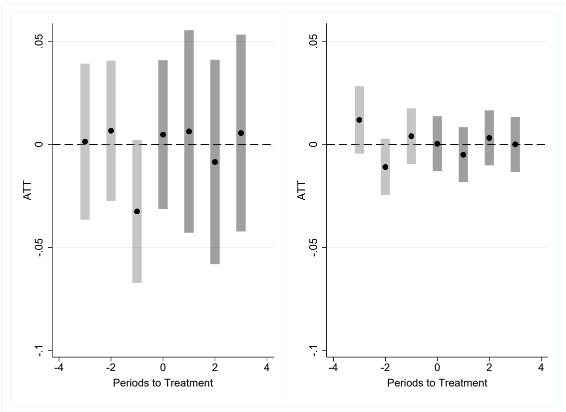
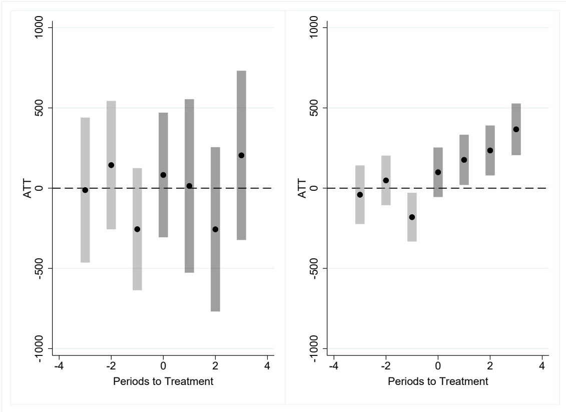


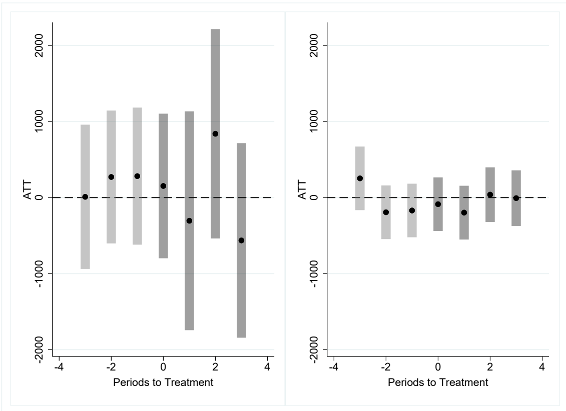
Figure 8: Upper secondary general education diploma in closure and inflow areas

D.3 Labor market outcomes

a



b



c

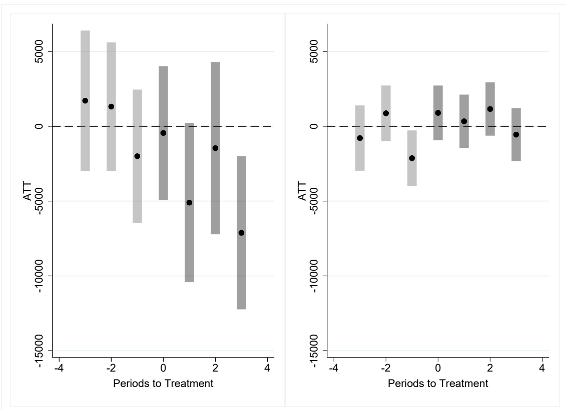


Figure 9: Aggregated event ATTs for closure and inflow areas **Panel a**: disposable income at 20, **panel b**: total taxable income at 25, **panel c**: total debt at 25