Waiting for Surgery: Effects on Health and Labor Supply*

Anna Godøy[†] Venke Furre Haaland[‡] Ingrid Huitfeldt[§] Mark Votruba[¶]

July 12, 2017

∼ Preliminary and incomplete – comments welcome ∼

Abstract

In universal health care systems, patients often face significant wait times for treatment when capacity constraints are binding. In this paper, we estimate the effects of wait time for orthopedic surgery (days from referral to surgery) on health and labor market outcomes, using microdata covering all publicly funded orthopedic surgeries in Norway referred in 2010 and 2011. As the system assigns higher priority to more urgent cases, naive OLS estimates linking observed wait times to individual patient outcomes could reflect selection bias. Our identification strategy exploits quasi-random variation in wait times for surgery generated by the idiosyncratic variation in system congestion at the time of a specific patient's entry into the queue. Precisely, we instrument a patient's wait time by the average wait time of other patients queued for the same procedure at the same hospital around the same time. We find that longer wait times for surgery significantly increase health related work absence: For every 10 days spent waiting for surgery, sick leave in the two years following referral increases by about 2.6 days. Moreover, longer wait times do not appear to have any lasting health implications.

JEL codes: I120, J320

Keywords: Wait time, queues, hospital treatment, health outcomes, labor market attachment.

*This research has received support from the Research Council of Norway (grants #214338 and #227117); the Western Norway Regional Health Authority; and the research department at Stavanger University Hospital. Data made available by Statistics Norway and the Norwegian Patient Register have been essential for the project. We are grateful to Sverre AC Kittelsen, Edwin Leuven, Mari Rege and Kjetil Telle for helpful discussions, suggestions and comments at various stages of this project. Comments from conference and seminar participants at EEA-ESEM (Geneva, 2016) are gratefully acknowledged.

[†]Institute for Research on Labor and Employment, University of California, Berkeley and Institute for Social Research, Norway. E-mail: anna.godoy@berkeley.edu

[‡]University of Stavanger, Norway. E-mail: venke.f.haaland@uis.no

[§]Statistics Norway and the Frisch Centre, Oslo, Norway. E-mail: ish@ssb.no

[¶]Case Western Reserve University and University of Stavanger. E-mail: mark.votruba@case.edu

1 Introduction

Queues are a ubiquitous feature of universal health care systems, and an issue of persistent public concern. Universal systems employ queues to handle excess patient demand under capacity constraints, frequently resulting in significant wait times for surgeries and medical procedures. For example, average wait time for hip replacement in 2014 was 91 days in the UK and 152 days in Norway.[1] Waiting imposes welfare costs on the patients seeking treatment, at minimum, through the reduced utility associated with the delayed care. Patients suffering a longer period in need of treatment may further experience health deterioration, possibly resulting in lower efficacy from treatment.[2] Policymakers allocating resources have to balance the costs of increasing capacity against the expected health gains from faster treatment. There might, however, be additional fiscal effects to consider: long wait times could negatively impact labor supply. A negative effect on labor supply could happen both in the short run through an extended period of being unfit for work while waiting for treatment, and in the long run due to permanent health effects, behavioral responses or persisting detachment to the labor market. Long-term absence from work, perhaps in combination with the occurring sickness and associated contact with the health care system, may spur a process of marginalization from the labor market (Hultin et al., 2012). Possible mechanisms are physical inactivity leading to slower recovery (Malmivaara et al., 1995); and human capital depreciation, including loss of network and lower productivity (Mincer, 1974; Rees, 1966; Calvó-Armengol and Jackson, 2004; Becker, 1991).

This paper investigates the health and labor supply implications of waiting for orthopedic surgeries. There are (at least) two arguments for focusing on orthopedics. First, orthopedic conditions are seldom life threatening, and procedures are often not time-sensitive, allowing hospitals greater freedom in delaying surgery. As a consequence, individual wait time is more likely to be driven, in part, by random fluctuations in queue length. Second, musculoskeletal conditions are the leading causes of health-related work absence, constituting about 40% of all sick leave spells in Norway (Brage et al., 2013). Orthopedic surgeries may therefore be particularly relevant for studying downstream labor supply responses.

Identifying a causal relationship of wait time for hospital treatment on health outcomes and labor market attachment is challenging, as wait time is presumably correlated with unobservable individual characteristics, such as health and preferences for work, which affect both health outcomes and work force attachment. In particular, the healthiest patients will typically be assigned longest wait times, as more serious cases are given priority, giving rise to a spurious positive correlation between wait time and labor market

[1] Figures are obtained from *OECD.Stat* at `https://data.oecd.org/health.htm` Health Care Utilisation/Waiting times (date retrieved 10/28/2016).

[2] Empirically, evidence on the effect of wait times on health is mixed. Some observational and randomized studies find no evidence that longer wait times reduce pre-treatment health status (Derrett et al., 1999; Hirvonen et al., 2007) or in-hospital mortality (Hamilton and Bramley-Harker, 1999; Moscelli et al., 2016). Moscelli et al. (2016) find no evidence of wait times for coronary bypass being associated with higher in-hospital mortality and only weak association with emergency readmissions. Other studies suggest that long wait time may also lead to psychological problems, such as depression and anxiety (Underwood et al., 1993; Brownlow et al., 2001).

attachment. On the other hand, patients with more resources might be more skilled at navigating the health care system, enabling some degree of queue jumping even within a public system. This channel could introduce a negative bias in the relationship between wait time and later employment.

We address these endogeneity issues by employing an instrumental variable (IV) approach, utilizing idiosyncratic variation in system "congestion" at the time of a specific patient's entry into the queue. Specifically, we instrument patient i's wait time by the average wait time of other patients queued for the same procedure at the same hospital around the same time as patient i. This empirical approach is enabled by rich administrative data covering the entire population of Norway, matched with unique individual patient data comprising all visits to general practitioners (GPs) and publicly funded specialists and hospitals.

The crucial identifying assumption for our instrument variable approach is that the patients referred to treatment during short wait times are not systematically different from the patients referred during long wait times. Wait time could operate as a rationing device causing some people to forego care or opting for private options when queues are long. If so, this is a potential problem for our IV model through self-selection bias.[3] Our rich data allow us to carefully investigate the plausibility of this assumption by exploring the correlation between our instrument, average wait time, and a battery of observable individual characteristics such as age, education, income, prior labor market attachment and health care history. We find no evidence that patients referred to treatment during periods of long expected wait time are different from people referred to treatment in periods of short expected wait time, suggesting any negative bias introduced by high socioeconomic status individuals opting for privately funded care as a response to long expected wait time is small to negligible.

Our study contributes to the literature by, to the best of our knowledge, being the first to credibly identify labor supply responses from hospital wait time. Aakvik et al. (2015) analyze the effect on sickness absence of being exposed to a reform in Norway that aimed at reducing wait time. They do not, however, explicitly estimate effects of wait time, but rather identify a reform effect.[4] Moreover, their sample includes only people who are on sick leave before admission to the hospital, and define wait time as days from the first day of the absence spell until treatment. Our approach exploits the exact date of referral to the hospital, and we can therefore additionally include people who are not on sick leave at referral. In addition to estimating causal effects of wait time on health and labor market attachment, our rich data allow us to investigate mechanisms through which labor supply might be affected.

Our results indicate a significant negative effect of waiting for orthopedic surgery on

[3]Empirically, there is mixed evidence on the effectiveness of queues as a rationing device in health care. Martin and Smith (1999) find that demand for treatment is relatively inelastic with respect to wait times, while Martin and Smith (2003) find demand elasticities for elective surgery between negative .1 and .2 (-0.07 for orthopedics)

[4]The reform, 'Faster Return to Work', is explained in Section 2

labor supply. In the two years following referral, each additional day spent waiting for treatment increases sickness absence by 0.26 days. The increase in sick leave is likely not only explained by increased absence while waiting for surgery; there are indications that post-surgery absence increases as well. However, longer wait times do not appear to have any serious lasting impacts on health, nor on disability benefit uptake within two years of referral.

2 Institutional Setting

Hospitals

Somatic specialist health care in Norway is funded primarily through taxes and transfers from the national government. Access to hospital services is either via emergency admissions or through referrals from general practitioners acting as gatekeepers, who are responsible for all initial assessment, examinations and treatment of patients. Patients pay a very low or zero price for using hospital services.[5] In addition to explicit rationing by gatekeepers, utilization is rationed by wait times, aiming at prioritizing patients based on their medical need for health care. After an individual has been referred for a specialist health treatment, the patient is assigned either a priority status or a non-priority status. The patients with priority status receive an assigned 'time limit' denoting the time by which the patient should receive treatment. The time limit is assigned by health professionals based on the patient's medical condition. The patient then enters a wait list, depending solely on his or her priority status and time-limit.

Until 2007, a patient's priority was strictly based on his or her health condition and expected efficacy of treatment. After 2007, however, the priority status was also based on his or her labor market attachment. This was the consequences of a 'Faster Return' reform (FRW), which had the purpose of decreasing the wait time for those who were on sick leave while waiting for treatment, urging a faster return to work. The reform allocated the hospitals additional resources to provide individuals on sick leave fast treatment, while, ostensibly, not affecting the wait time of other patients without FRW status.

In general, patients in need of specialist health care are typically assigned a hospital based on their home address. Since 2001 patients who are referred to specialist health care have had the right to choose the hospital at which they want to receive treatment. Patients may choose to be treated at hospitals outside of their referral area; either at another health trust within their region or in another region, but the latter is infrequently observed.[6]

[5] Patients' health care expenses are mainly subsidized by national insurance schemes. Some services, such as outpatient visits and visits to primary care physicians are subject to small co-payment rates. In 2015, the out-of-pocket payment rate for an outpatient procedure was 320NOK (40USD). However, once a patient's yearly total out-of-pocket health care expenditures exceed about 2100NOK (260USD) all further expenses within that calendar year are reimbursed.

[6] 90% of elective surgeries are performed within the patients' own region, and 22% chooses another hospital of which they do not belong within their own region (own calculations). An information service called Free Hospital Choice facilitates the option to freely choose hospital by making quality indicators such as expected wait time publicly available.

Health-related benefits: Sickness absence and disability insurance

Employees usually receive sick pay equivalent to their regular salary from the first day of sickness absence. Expenses during the first 16 days are covered by the employer, while the Norwegian Labour and Welfare Service (NAV) takes over the responsibility on the 17th day of sick leave. The wage replacement ratio for sick pay is 100% and benefits can be maintained for up to 12 months.[7] Persons who are still unable to work after one year of sickness may apply for rehabilitation or disability benefits. Disability insurance benefits amount to 66% of the applicants' wage.[8] All health-related benefits must be certified by a physician.

3 Data and Descriptives

3.1 Data Sources

The empirical analysis is based on data that combine several administrative registers obtained from Statistics Norway (SSB), the Norwegian Patient Register (NPR) and the Control and Payment of Health Reimbursement (KUHR). A unique personal identifier is provided every Norwegian resident at birth or upon immigration, enabling us to match the wait list records with administrative data of the entire resident population of Norway. Data provided by SSB contain birth and death dates, sex, district and municipality of residence, country of origin, education, occupation, annual earnings and health-related benefits. Information on sickness absence and disability insurance uptake comes from social security registers that contain complete records for all individuals. As employers are responsible for the initial period of sickness related absence, administrative social security data only identify sick leave spells lasting at least 17 days.

NPR data contain complete patient level observations for all somatic public hospitals and private hospitals contracting with regional health authorities in Norway from 2008. Records include hospital identifiers, patient identifiers, patient municipality, age, sex, main diagnoses (ICD10), comorbidities (ICD10), surgical/medical procedures (NCSP/NCMP),[9] time of deaths in/out of hospital, exact time, date and place of admissions and discharges, length of stay, DRG groups and DRG cost weight.[10] In addition, starting in 2010, NPR

[7]Benefits are capped at higher earnings; in 2015, the benefit cap was approximately 540,000 NOK or around 68,000 USD. However, all public sector workers and many private sector workers are covered by employer provided top-up insurance.

[8]DI benefits are calculated based on the three best years among the 5 latest years before sickness. Benefits are capped at about 540,000 NOK or around 68,000 USD

[9]Surgical procedures are coded according to the NOMESCO Classification of Surgical Procedures (NCSP). Medical procedures are classified according to NCMP - Norwegian classification of medical procedures.

[10]Each patient discharged at a somatic hospital is assigned a DRG group that uniquely determines the reimbursement rate. Patients within the same DRG group are ostensibly homogeneous with respect to both medical criteria and financial costs of treatment. Main diagnosis, comorbidities, medical and surgical procedures, age, and resource consumption, are crucial components when allocating patients to a particular group.

data include information about date of referral (i.e. the date at which the hospital received the referral).

The KUHR database contains all visits in the primary care, as well as visits to specialists, since year 2001. Data include patient identifier, date of visit, diagnosis group, reimbursement code and size of patient deductible.

3.2 Sample

Our linked data include Norwegian residents in 2005-2013, as well as NPR and KUHR data from 2008-2013. Our instrument is constructed on a sample consisting of 69,257 individuals referred to an orthopedic surgery (NCSP-code starting with "N") in 2010-2011, and treated within two years of referral.[11] For each procedure at every hospital, we require there to be at least 10 referrals per month throughout 2010-2011.[12] The instrument is calculated as the (leave-out) average wait time within each combination of procedure×hospital×time period, where the time period is defined as 14 days before and after the referral of each individual. We require there to be at least 5 individuals within each combination. In Section 6 we show that results are robust to varying both the window size and the queue size cutoff.

In our estimation sample, patients referred to two different surgeries at the same day are dropped. Finally, we retain only patients with a likely attachment to the labor market, as defined by being aged between 25 and 60 and earning more than two times the substantial gainful activity level in the year before referral (about $2 \times 11,000$USD in 2016).[13] This leaves us with a sample of 25,958 individuals. There are 22 hospitals in our sample, including data on 5 distinct orthopedic procedures (see Table A1 for description and volume of procedures included). In total, this amounts to 70 groups of hospitals×procedures, as not all procedures are performed at all hospitals.

3.3 Variable definitions

Our dependent variables capture both health outcomes and labor market attachment, as presented in Table 1. Health outcomes are measured as episodes occurring between date of

[11]This includes all planned admissions with non-missing date of referral. Observations with wait time longer than two years are dropped, as these likely represent error records. 2010 is the earliest year in which date of referral is reliably defined in the hospital data; and we include only two years of hospital admissions/visits to allow a subsequent two-year period for studying outcomes. Unfortunately, we do not observe referrals that are not ending in any hospital visits.

[12](i) Procedures are coded using Classification of Surgical Procedures (NCSP). We use only the two first letters of the code. See Appendix A for NCSP codes included. (ii) We exclude the restriction of 10 referrals in July which is a low-activity month.

[13]The substantial gainful activity level ('basic amount') corresponds to approximately 11,000USD in 2016. The 'basic amount' is used by the Norwegian Social Insurance Scheme to determine eligibility for and magnitude of benefits like old age pension, disability pension, and unemployment compensation. The 'basic amount' is adjusted annually by the Norwegian Parliament to account for inflation and general wage growth. 1 basic amounts equals 92,576NOK per May 1, 2016, and we assume a currency exchange rate NOK/USD=8. We follow previous studies (Havnes and Mogstad, 2011a,b) and refer to an individual as full-time employed in a given year if he or she is earning more than four 'basic amounts,' and as employed (part-time or full-time) if he or she is earning more than two 'basic amounts'.

referral and within two years, and defined as (i) number of inpatient days; (ii) number of hospital outpatient visits; (iii) number of visits to the hospital within the same diagnosis group for which the patient is waiting; (iv) the dollar value of any utilization at GPs, specialists or hospitals;[14] and (v) mortality rates, measured as deaths within three years of referral.

Labor market outcomes are variables counting the number of sick leave days in the first and second year from referral date, and the sum of the two years; a binary indicator for receiving disability (DI) benefits in the second year after referral; and annual indexed earnings (one unit corresponds to the substantial gainful activity level amount of about 11,000USD). Note that our earnings variable includes sickness absence benefits.

Sickness absence and visits to the hospital and GP are measured using exact dates, such that year one corresponds to 1-365 days after referral, and year two refers to 366-730 days after referral. Our measure of disability insurance lacks precise dates of spells. DI benefits in year two is therefore equal to one if the patient is receiving any welfare benefits in the second calendar year following referral. Earnings are also measured per calendar year and therefore subject to the same constraints.

The key explanatory variable is the number of days spent waiting from the referral date to the first observed treatment date. All models include fixed effects for hospital×procedure, year and month of referral. In models where we include additional control variables (measured in the year prior to referral), these are: week fixed effects, linear, quadratic and cubic terms for age, earnings and indicators for female, married, foreign born and education status (high school dropout, high school graduate, college). All variables are summarized in Table 1.

From Table 1 we see that mean wait time is 173 days with standard deviation 133 days. The distribution of this variable is depicted in Figure A1. There are slightly more males than females. The share of people with primary education; high school graduation and college education is about one third for all groups. About one fourth of the patients are on sick leave at referral date. Average sick leave duration is higher in the first year following referral date compared to the second year. This is likely because treatment is most often undergone in the first year, hence yielding higher sick leave periods due to recovery.

[14] To construct the health care utilization measure, we apply the nationally set DRG-specific weights for all hospital stays (see https://helsedirektoratet.no/finansieringsordninger/innsatsstyrt-finansiering-isf-og-drg-systemet/innsatsstyrt-finansiering-isf). For visits to GP or specialists outside of the hospital, we sum over all fee-for-service reimbursement rates using the prices set nationally, following 'Fastlegetariffen': http://normaltariffen.legeforeningen.no/pdf/Fastlegetariff_2016.pdf. The utilization measure consists of all these expenditures combined.

Table 1: Descriptive Statistcs

Variable	Mean	SD
Background and explanatory variables		
Age	46.25	(9.166)
Female	0.463	(0.499)
Primary education	0.325	(0.468)
High school graduate	0.366	(0.482)
College	0.309	(0.462)
Earnings	5.975	(3.289)
On sick leave at referral	0.256	(0.437)
Wait time	173.4	(133.3)
AWT	178.4	(50.24)
Outcome variables		
Earnings t1-t2	9.870	(7.878)
Sick leave t1	85.21	(99.09)
Sick leave t2	41.14	(81.65)
Sick leave t1-t2	126.4	(135.4)
Disability insurance benefits t2	0.165	(0.371)
Health care utilization ($) t1-t2	7389.2	(8744.8)
Admissions to same procedure t1-t2	1.085	(0.304)
Outpatient visits t1-t2	5.930	(6.661)
Inpatient days t1-t2	1.907	(5.616)
Mortality (per 1,000) within 3 years of referral	3.005	(54.74)
Observations	25,958	

Notes: Summary statistics for various outcomes and background variables. Background variables are measured the year prior to referral. Outcomes are shown for years relative to year of referral: for sick leave and health measures, year 1 is measured from date of referral until day 365; year 2 is from day 266-730. For DI benefits and earnings, year 2 refers to second calendar year after referral. Earnings are indexed; one unit corresponds to about $11,000.

4 Identifying effects of wait time for hospital treatment

4.1 Threats to identification

Waiting for hospital treatment may affect health outcomes, as well as incidence and duration of sickness leaves. Identifying a causal relationship of wait time for hospital treatment on health outcomes and labor market attachment is challenging, as wait time is presumably correlated with unobservable individual characteristics, such as health and preferences for work, which affect both health outcomes and labor supply. To frame ideas, consider the following linear regression model, estimated in a sample of individuals who receive a specialist referral:

$$Y_{ihp} = \beta_0 + \beta_1 X_i + \delta WT_i + u_{ihp}, \tag{4.1}$$

where Y_{ihp} is a measure of health or labor market outcomes for patient i undergoing procedure p at hospital h; X_i is a vector of control variables; and WT_{ihp} represents days from referral to treatment, i.e the wait time. In this model, our coefficient of interest, δ, provides an unbiased estimate of the effect of wait time on sick leave under the assumption that variation in wait time (conditional on X_i) is uncorrelated with unobservable determinants of sick leave duration or health outcomes.

There are several reasons why the exogeneity assumption is unlikely to hold. First, patients with the greatest need are given priority in the allocation of treatment slots. As a result, healthier patients typically have longer wait times than patients with a more urgent need for medical care. While the prioritization mechanism ensures that healthy people are subject to longer wait times, healthy people are also less likely to have long absence spells, possibly biasing our estimate of wait time. Moreover, after the FRW reform was passed in 2007, hospitals are allowed to give priority to patients who are on sick leave or at high risk of entering sick leave. This scheme could also lead to an association between short wait times and high incidence of work absence.

Finally, observed wait time may to some extent be determined by individual behaviors that are correlated with health outcomes. For example, patients with better knowledge of the health care system may be able to jump ahead in the queue. These persons may also be more likely to have a fast recovery and lower sick leave duration regardless of wait time. Such selection effects would cause patients with better prospects for recovery to have shorter wait time, biasing our estimate of δ towards finding negative effects from longer wait times.

To summarize, OLS estimates of Equation (4.1) are likely to suffer from omitted variable bias, yielding a biased estimate of δ. To address concerns of omitted variable bias and endogeneity arising in estimation of Equation (4.1), we will instrument for individual time to hospital treatment (WT) with a constructed measure for average wait time (AWT).

4.2 Instrument: Average wait time

In our empirical strategy we exploit variation in wait times that arise because the degree of system congestion fluctuates over time. As a result, otherwise similar patients have different expected wait times based on the date they enter the queue. By using observed wait times of other patients referred to treatment around the same time, we are able to recover a consistent measure of average wait time (AWT) that is independent of the patient's own pre-referral health and labor market attachment. This measure, AWT, is then linked to individual health and labor market data to identify causal effects of wait times on post-referral outcomes.

The measure of average wait time is constructed using a sample of patients who undergo surgical treatment for orthopedic conditions at Norwegian hospitals. In our baseline specification, the average wait time facing patient i is constructed by calculating the average observed wait times of all other patients at the same hospital h and procedure p whose referral dates fall within a four week-window of patient i's referral date (two weeks before and two weeks after).[15]

In order for the identification strategy to be valid, the independence assumption should hold: AWT should be as good as random, conditional on calendar time and hospital×procedure. That is, it should be uncorrelated with patients' observed and unobserved pre-referral characteristics. If this assumption holds, simple models linking individual outcomes to AWT will give consistent unbiased estimates of the (reduced form) causal effects of average wait time.

Referral to specialist health care is based on a medical evaluation, leaving little scope for patients to time referrals to periods of low wait times. Moreover, as there are no direct costs of being on the wait list, there is no incentive for patients or primary care providers to delay referral once the decision has been made that a surgical procedure is the best treatment. As the instrument is constructed using wait times of other patients only, AWT is not determined by i's own underlying health, priority status, or previous labor market attachment. While hospitals with long wait times may be different from hospitals with shorter wait times, our 2SLS model fully controls for time-invariant hospital characteristics by including hospital×procedure fixed effects. Year and month fixed effects are included to control for seasonality and general time effects.

The independence assumption may be violated if some patients respond to long average wait times by seeking out treatment at private hospitals operating outside the public health care system. While a large majority of orthopedic surgeries are performed in public hospitals or private hospitals contracting with the government, there is a small and growing market for privately funded hospitals that perform certain surgical operations. The costs of these procedures are not reimbursed by the government, but are paid for by the patients themselves or through employer-sponsored private health insurance. Thus high income patients might opt out of public health care when wait times are long, resulting in a negative correlation between socioeconomic status and average wait time.

Whether or not high socioeconomic status patients choose private health care options when wait times are long cannot be tested directly, as privately funded procedures are not included in the patient register data. However, the dataset does include a large set of observable characteristics that are correlated with health and labor market outcomes, including age, education and previous earnings, as well as proxies for pre-referral health status such as visits to GP and hospital, and time spent on sick leave in the years prior to referral.

Following Dahl et al. (2014), Table 2 gives empirical support for the claim that average

[15]This is the leave-out mean, defined as

$$AWT_{ihp} = \left(\frac{1}{N_{hp(i)} - 1}\right) \times \left(\sum_{k \neq i}^{N_{hp(i)}-1} WT_k\right) \tag{4.2}$$

where N_{hp} is the total number of people referred to the same hospital-procedure as person i, and within the same time frame (i.e. 14 days before or after person i's referral date). WT_k is the observed wait time of any other person k entering the queue at the same time as person i.

wait time is random-like conditional on hospital×procedure, year and month. Precisely, Table 2 shows estimates from separate regressions of a number of variables capturing patients' demographic, work and health related variables on wait time and AWT, respectively. (Hospital-procedure and time fixed effects are also included in these models). All covariates are measured in the year prior to referral.

The first column documents that these control variables are predictive of patients' *individual* wait time. Recalling the discussion on threats to identification, our fears that individual wait times are correlated with unobserved determinants of health appear to be justified. In particular, patients with higher education or income experience significantly shorter wait times compared people with lower education and earnings, respectively, consistent with a scenario where better knowledge of the health care system facilitates some degree of "queue jumping". Being on sick leave is also associated with significantly shorter wait times - consistent with the health care system giving priority to patients with more serious health problems. Though the significant associations between background variables and individual wait time are interesting alone, they pose no threat to our identification strategy. Importantly, these same characteristics are generally not statistically related to *average* wait time (AWT). Only one of the variables is individually statistically significant at the 10%-level: prior health care utilization. However, the estimate is extremely low; -0.0001 from a mean of 1164USD. Testing for joint significance on all historic variables yields a p-value of 0.099. The joint significance at the 10%-level is likely due to the contribution from health care utilization. To illustrate, when not controlling for health care utilization, the joint p-value is 0.22. When testing for multiple hypotheses as we do in this table, the risk of getting one significant variable by pure chance is high. Altogether, we argue the estimated models presented in Table 2 support the claim that AWT is random-like, conditional on hospital, procedure and calendar time.

4.3 Instrumental variable model

Our empirical model can be described by the following two-equation system:

$$WT_{ihp} = \alpha_0 + \alpha_1 X_{ihp} + \alpha_2 AWT_{ihp} + \varepsilon_{ihp}, \tag{4.3}$$

$$Y_{ihp} = \beta_0 + \beta_1 X_{ihp} + \delta WT_{ihp} + \nu_{ihp}, \tag{4.4}$$

where Y_{ihp} is a measure of patient i's health and labor market outcomes; X_{ihp} is a vector of control variables (hospital×procedure, year and month of referral); AWT_i denotes average wait time; and ε_{ihp} and u_{ihp} are error terms. This specification controls for any differences over time across hospitals or procedures in the quality of hospital or health of patients. We perform two-stage least squares (2SLS) estimations with Equation (4.3) as the first stage and Equation (4.4) as the second stage.

The coefficient of interest, δ, represents the effect of wait time for hospital treatment

Table 2: Testing for random-like average wait time

	(1) Wait time		(2) AWT	
Age	-0.0669	(0.0807)	-0.0056	(0.0172)
Female	4.8219**	(2.2074)	-0.1064	(0.3080)
Foreign born	16.6598***	(2.5116)	-0.2901	(0.4363)
Partner	-0.1336	(1.4142)	0.4613	(0.3096)
Primary education	3.7936**	(1.6175)	-0.1370	(0.3376)
High school graduates	-2.0832	(1.6210)	0.0752	(0.4511)
College	-1.6764	(1.9034)	0.0605	(0.4726)
Earnings	-1.0888***	(0.3499)	-0.0551	(0.0434)
On sick leave at referral	-32.6147***	(4.0289)	0.3106	(0.4134)
Sick leave days	-0.0151	(0.0169)	0.0001	(0.0017)
Health care utilization ($)	-0.0008**	(0.0003)	-0.0001*	(0.0001)
Observations	25,958		25,958	
Dep. mean	173.4		178.4	
F-statistic for joint significance	17.2		1.9	
Joint p-value	0.000		0.099	
R-squared	0.024		0.014	

Notes: All variables are measured in year prior to referral. Robust standard errors clustered at hospital level in parentheses; *p<0.1, **p<0.05, ***p<0.01. AWT is calculated as the average wait time of other patients entering the queue at the same time (±14days) as patient i. Columns 1 and 2 show results for separate regressions of individual wait time and AWT, respectively, on patient characteristics, while also controlling for hospital×procedure, year and month fixed effects.

on the outcome variable. While the independence assumption is sufficient for a causal interpretation of reduced form estimates, additional assumptions are required for our IV model to produce a causal effect of δ.

First, the instrument should be relevant: the instrumental variable, average wait time, should be correlated with the endogenous regressor, individual wait time. In order for this to hold, individual wait times should be determined in part by local fluctuations in excess demand and capacity constraints within each hospital. As discussed in the introduction, our paper's focus on orthopedic surgeries implies that this assumption is more likely to hold: orthopedic conditions are rarely life threatening, leaving hospitals with considerable discretion in delaying surgeries when excess demand is high. The relevance assumption can be tested directly by examining the first stage of the 2SLS estimation results.

Second, the instrument should affect the outcome only through its effect on individual wait time. This exclusion restriction would be violated if, say, increased average wait times affected health outcomes through lower quality caused by congestion in the hospital unit. The exclusion restriction cannot be tested directly, however, we can examine whether there are signs of congestion effects in the data by looking at the correlation between average wait times and the volume of orthopedic surgeries. Moreover, patients who are admitted for immediate surgery (emergency admissions) may provide a useful control group, as they are treated by the same medical teams without being subject to a waiting period. If the exclusion restriction holds, then average wait times should have no effects on outcomes for this group. This is exactly what we find in our robustness test presented in Appendix B.

In the presence of heterogeneous treatment effects, 2SLS regressions identify a local average treatment effect (LATE) for patients whose wait time for surgery is affected by the queue length at the time of referral. Our identification strategy hinges on there being a significant number of patients who, had they entered the queue at another point in time, would have been subject to a different wait time. However, for some groups of patients, the wait time for surgery may never be affected by the concurrent average wait time, e.g. patients with particularly serious injuries may have a more urgent need for surgery, allowing them to skip ahead of the queue. The local nature of IV estimates means that the estimated effects may not be informative about impacts on this group.

With heterogeneous treatment effects, the instrument should additionally satisfy monotonicity in order for our estimates to have the LATE interpretation. The monotonicity assumption states that the instrument should affect the variable being instrumented in only one direction: Longer AWT at the time of referral should always lead to individual wait times remaining unchanged or increasing. The monotonicity assumption would be violated if there exist some subset of patients who experience shorter wait times as a result of longer average wait times. Given the institutional context, it is difficult to think of scenarios where the monotonicity assumption would be violated. Again, the monotonicity assumption cannot be tested directly, but it does have a testable implication: The estimated first stages should be positive across subgroups with varying average wait times. We return to this in Section 6.

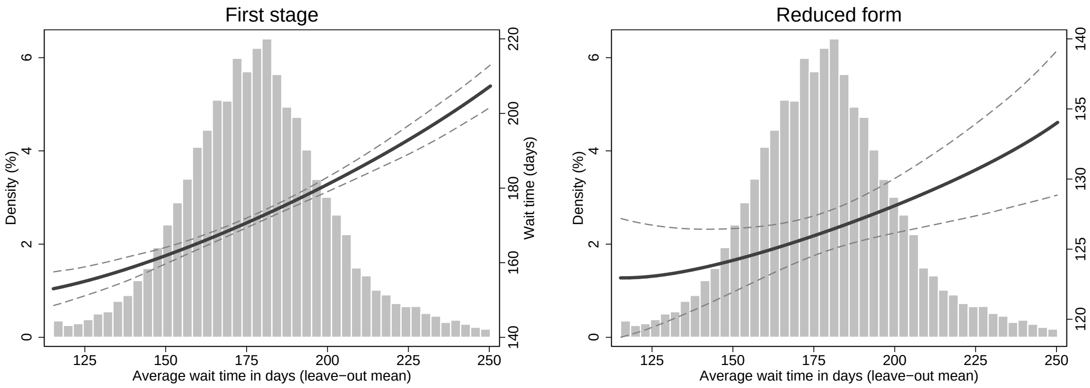
5 Results

5.1 Graphical Evidence

We begin our presentation of results by providing a graphical representation of the IV approach in Figure 1. A histogram for the density of AWT is depicted in both panels, and captures the distribution of (leave-out) average wait time for all patients entering the queue. In these figures, AWT is included as the residual from a regression of (leave-out) average wait time on fixed effects for hospital×procedure, year and month of referral, then rescaled to the instrument mean.[16]

Panel (A) illustrates the relationship between AWT and individual wait times, corresponding to the first stage equation (4.3). The graph plots a local linear regression of individual wait time against AWT. Individual wait time is monotonically increasing in average wait time, and is close to linear. This provides some evidence that the monotonicity assumption may be satisfied. An increase in average wait time by 100 days is associated with an approximate increase in individual wait time of about 40 days. Panel (B) plots

[16]Figure A1 depicts both the residualized instrument (i.e the residual from a regression of AWT on fixed effects for hospital×procedure, year and month); and the 'raw-instrument' (AWT)



(a) First stage

(b) Reduced form

Figure 1: Effect of average wait time (AWT) on individual wait time (first stage) and sickness absence days (reduced form)

Notes: Panel (a) illustrates the first stage. Solid line is a local linear regression of individual wait time on average wait time. Panel (b) shows the reduced form equivalent. Solid line is a local linear regression of sickness absence days within two years following referral date on average wait time. In both figures, average wait time is included as the residual from a regression of (leave-out) average wait time on fixed effects for hospital×procedure dummies, year and month of referral. Histogram of (residualized) average wait time is shown in the background of both figures (top and bottom 1% excluded from the graph). Dashed lines represent 90% CI.

the reduced form effect of average wait time against sickness absence days within two years following the referral date, again using a local linear regression. Sickness absence is monotonically increasing in average wait time; an increase in average wait time by 100 days predicts about seven more days of sickness absence within the two years following referral.

5.2 Regression Estimates

This section discusses the effects of wait time on health outcomes and labor market attachment. First, we present our baseline IV estimates on health and labor market outcomes during the two year period following referral to treatment. Next, extended models are estimated to shed further light on the underlying mechanisms. These models aim to investigate whether effects primarily occur while patients are awaiting treatment, or whether effects persist after surgery.

The first set of models estimates effects on health outcomes and health care utilization measures. Table 3 presents results from estimation of Equations (4.3) and (4.4); the corresponding OLS estimates are included for reference. All models shown in this table include dummies for hospital×procedure, month and year of referral.

OLS estimates (panel A) appear to find some statistically significant correlations between wait time and health outcomes: longer wait time is positively correlated with the number of outpatient visits and total health care utilization, while being negatively correlated with mortality and total number of same-condition hospital visits. However, interpreting these correlations is complicated by the likely non-random nature of individual wait time, as discussed in previous sections.

Moving on to the IV estimates (panel B), column (1) indicates that our first stage is positive and strongly significant (F-value 44.1). When average wait time increases by one day, individual wait time increases by approximately 0.38 days. IV estimates of effects on outpatient visits, inpatient days and visits to the same procedure for which the patient is waiting, are all small and nonsignificant. These estimates are fairly precise, and we are able to rule out any considerable effects. Similarly, we find no effect on mortality (the effect is rather imprecisely estimated, likely due to the low sample average) or on total utilization of health care services. Next, table 4 presents effects of wait time on labor market outcomes: sickness absence, disability insurance receipt and earnings (including sick leave benefits). OLS results shown in Panel A show that the association between individual wait time and total sick leave in year 1 and 2 is positive and statistically significant but small in magnitude. These estimates are likely to reflect a combination of selection effects brought about by the non-random assignment of wait time, as well as any causal effects of wait time on health and labor market outcomes.[17] In either case, interpreting the OLS estimates is complicated, and the estimates reported should not be given a causal interpretation.

Panel B shows the IV estimates for labor market outcomes. The effect of wait time on sick leave is positive both in year one (column 2) and year two (column 3). When looking at effects for each year separately, it is worth keeping in mind that wait time has an effect on the timing of surgery and the associated recovery period. Longer average wait times make it more likely that surgery and recovery take place during the second year after referral, in turn affecting the timing of sickness absence. For completeness, column (4) then shows the effect of wait time on total sickness absence in year 1 and 2: One more day of waiting increases sick leave in the two years following referral by about 0.26 days.

We find no effects of wait time on earnings, however, one should keep in mind that our earnings measure includes income from sick pay. In particular, declines in market productivity associated with sick leave absences are not fully captured in our earnings measure because of the 100% wage replacement rate. We might therefore expect effects on earnings to be limited until patients have exhausted their sick pay benefits (after which they may enter DI, where the wage replacement rate is about 66%). Since a potential loss of earnings may to some extent be compensated through the sickness benefit scheme, any potential changes in earnings will mostly reflect forgone career/wage increase opportunities due to the absence; or employment effects, e.g. changing jobs. In any case, longer wait times do not appear to lead to any immediate earnings penalty.

The final outcome studied is disability insurance (DI) receipt. When considering fiscal spillovers from longer wait times, any effects on DI are particularly interesting as DI tends to be a more permanent state, with low rates of recipients returning to work. The model finds no effect of wait time on the probability of receiving DI within two years of referral date. From a policy perspective, this is reassuring, as it suggests that longer wait times

[17]In particular, an intentional policy of prioritizing patients in need of immediate treatment is likely to introduce negative selection bias, possibly accounting for the negative association between wait time and sick leave in year 1.

do not lead to permanent withdrawal from the labor market.

However, by looking at DI receipt only in the first two years after referral, we risk understating the true effects of wait times on long term labor market attachment. Transitions to DI may be delayed if patients need time to learn about how their medical condition affects their ability to do their work. Delays could also occur as a result of the institutional context: People are only eligible for DI once they have exhausted their sick pay benefits, which typically happens after 12 months of uninterrupted absence from work.

To further investigate the possibility of long term effects of wait times, we estimate an extended set of models examining the timing of effects in more detail. Conceptually, long wait time may increase the absence after referral in two distinct ways: First, longer wait times could lead to increased absence from work while waiting for treatment; for example, a patient may be unable to work until she gets a specific surgery - delaying the surgery will increase her absence rates only for as long as she is waiting for treatment. In this case, we would not expect to see long term effects.

Second, longer wait times could lead to increased absence post-surgery. This could occur through two channels: first, some surgical procedures may be time sensitive, meaning longer wait times could potentially lead to permanently reduced functional capacity in the affected patients. Second, there could be behavioral effects, such as habit formation; the longer you are absent from work while waiting for treatment, the lower is the cost of remaining on sickness leave also post-recovery. In either case, positive effects of wait times on post-surgery absence may be an early sign that wait times could impact long term labor supply - in particular if these effects do not seem to be fading over time.

Models estimating effects on absence before and after surgery are not likely to be informative: People with long wait times spend more time at risk for pre-surgery absence, and less time at risk for post-surgery absence, leading to spurious correlations and biased estimates. Rather, we estimate additional models of effects relative to referral as well as for fixed post-surgery time windows.

First, the model is estimated by quarter relative to referral date. By construction, patients in our sample wait for a maximum of two years (see Section 3), and only about 10% of all patients are still waiting at the beginning of year 2 (see Figure A1). If effects are driven primarily by pre-surgery absence, we would expect the largest effects in the first year after referral. On the other hand, if effects are larger in the second year after referral compared to the first year, this could indicate that there is some persistence in effects.

Results are presented in Figure 2. Panel (a) shows effects on sickness absence, while panel (b) shows effects on total health care utilization. The positive effect on sickness absence extends even throughout the fourth quarter of the second year since referral, suggesting that sickness absence effects are long-lived, extending well beyond surgery. Consistent with previous findings, there does not appear to be corresponding significant effects on total health care utilization (at least over our limited followup time). Nonetheless, longer waits had the predictable effect of pushing health care utilization back in time.[18]

[18]As there is a nonzero share of patients with very long wait times, we cannot exclude that any effect

In the next set of models, we re-estimate the baseline model, this time defining the outcome as the number of absence days in each of the first four thirty-day bins following surgery. Figure 3 shows results for the full sample. The point estimates are fairly big, and quantitatively similar to those found in Figure 4 (on average about 0.01 per month compared to 0.26 over 24 months). There seems to be an elevated effect of wait time on sickness absence in the first month after surgery, after which the effect fades out over time. In comparison, we still find no effect on post-surgery health care utilization.[19]

To summarize, longer wait times for surgery do not have any negative effects on the health outcomes that are included in our dataset. Meanwhile, longer wait time significantly increases sickness absence in the two years after referral; the model predicts 2.6 additional absence days for every 10 days spent waiting for surgery. Extended models indicate that effects on sickness absence are persistent beyond the period spent waiting for treatment and the immediate recovery period. The effect on sick leave is likely not only operating through an increase in absence while waiting for treatment. There are indications that post-surgery absence increases as well. We find no effect on the probability of entering DI within two years of referral. However, the persistent effect on absence suggests there may still be long term effects on labor supply, either due to adverse effects on health outcomes not captured in our dataset (e.g. pain and minor physical disabilities), or through behavioral effects/habit formation.

observed in the end of year two is because some patients just recently received treatment or are still waiting for treatment. To explore if the effects observed in Figure 2 reflect effects on sick leave that go beyond the waiting and recovery period, we exclude patients waiting more than 365 days and replicate the regressions from Figure 2. Appendix Figure A3 shows results from these regressions. Importantly, the estimated coefficients produced by this model need to be interpreted with caution, as we are conditioning the sample on an endogenous variable (wt>365 days). Nonetheless, comparing Figure 2 with Figure A3 there are still indications that restricting the sample to shorter wait times reveals the same persisting long-term effect on sick leave. (The estimated coefficients produced when conditioning the model on $wt > 365$ might be biased. For example, outliers with particularly long wait time, for reasons having nothing to do with system congestion, might strongly affect the measure of AWT. Additionally, we might have clusters of outliers in some periods. We explore the consequences of estimating on a sample of people with $wt > 365$ in Appendix Figure A2. Here, we plot the effect of AWT from separate regressions of a binary indicator for the probability of waiting more than X days, where X varies from 10 to 710 days (i.e. $Y = P(WT > X) = \beta AWT + X'\alpha$). From this figure we observe that the effect of AWT is smoothly decreasing in X and approaching zero, suggesting that restricting the sample to individuals with less than 365 wait days is likely not problematic.)

[19] The effects of wait time on post-surgery sick leave are biased upwards by the fact that, conditional on age at referral, post-surgery age increases mechanically with wait time. The size of the bias depends on the extent that sick leave use increases with age. To this end, the post-surgery model is estimated separately with no age controls; additional controls for age at referral; and age at surgery, respectively. This produces almost identical figures, and we therefore confine the presented results to the model with no additional controls for age (only fixed effects for hospital×procedures, year and month).

Table 3: Estimates of wait time on health outcomes

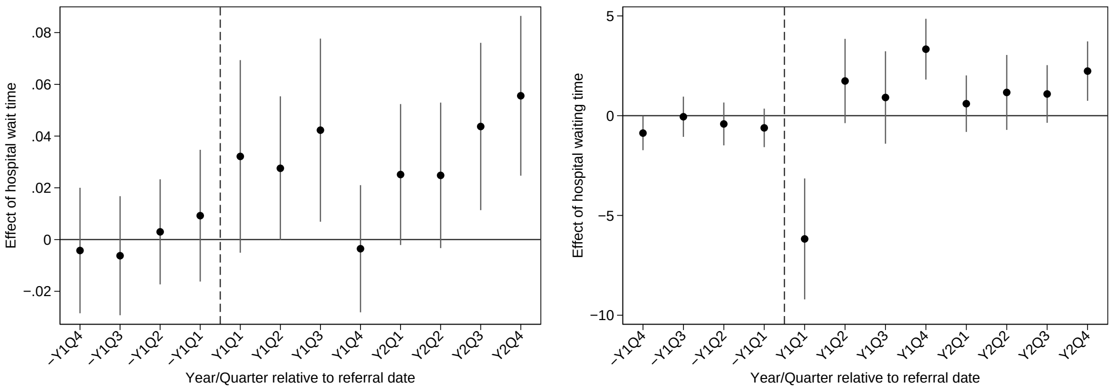
	(1) First stage	(2) Outpatient visits	(3) Inpatient days	(4) Re-surgery	(5) Mortality	(6) Total utilization($)
Panel A. OLS						
Wait time		0.00168***	0.0000332	-0.000154***	-0.00824**	1.196*
		(0.000572)	(0.000401)	(0.0000174)	(0.00299)	(0.673)
Panel B. IV						
AWT	0.376***					
	(0.0566)					
Wait time		0.00589	0.00556	0.00000102	0.0206	6.667
		(0.00455)	(0.00397)	(0.000158)	(0.0325)	(5.011)
Observations	25,958	25,958	25,958	25,958	25,958	25,958
Controls	✓	✓	✓	✓	✓	✓
Dep. mean	178	5.9	1.9	1.1	3.0	7389.2
FS F-stat	44.1					

Notes: Effects of wait time on health outcomes. All outcomes except mortality are measured in the two years following referral. The same procedure admission outcome measures the number of visits/admissions to hospital for the same diagnosis for which the patient is waiting in queue. Mortality rate (per 1,000) is within three years of referral. All models include dummies for hospital×procedure, month and year of referral. Robust standard errors clustered at hospital level in parentheses; *p<0.1, **p<0.05, ***p<0.01.

Table 4: Estimates of wait time on labor market outcomes

	(1) First stage	(2) Sick leave in year 1	(3) Sick leave in year 2	(4) Sick leave in year 1 and 2	(5) Earnings in year 1 and 2	(6) DI within year 2
Panel A. OLS						
Wait time		-0.0820***	0.107***	0.0250**	0.000270	0.0000952***
		(0.00912)	(0.00895)	(0.0106)	(0.000247)	(0.0000203)
Panel B. IV						
AWT	0.376***					
	(0.0566)					
Wait time		0.108**	0.148**	0.257***	0.000387	-0.000158
		(0.0507)	(0.0613)	(0.0880)	(0.00493)	(0.000192)
Observations	25,958	25,958	25,958	25,958	25,958	25,958
Controls	✓	✓	✓	✓	✓	✓
Dep. mean	178	85.2	41.1	126.4	9.9	0.2
FS F-stat	44.1					

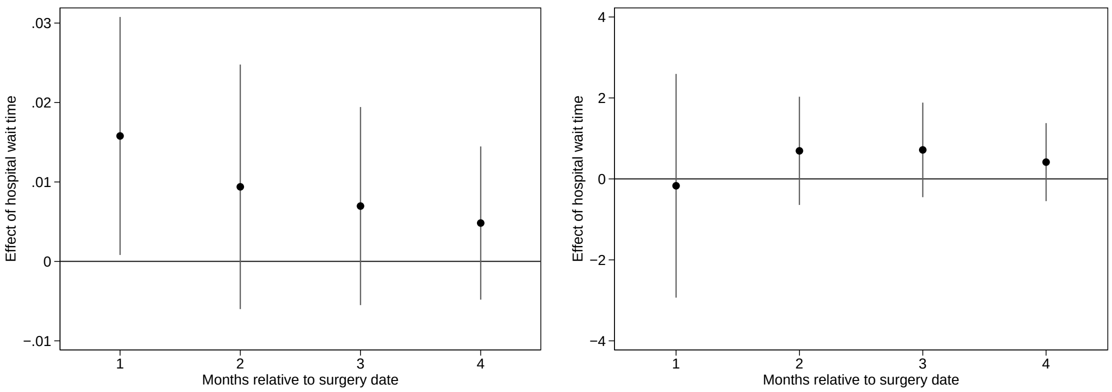
Notes: Effects of wait time on labor supply. Sick leave variables count days of absence in the specific year. Earnings are measured as the total earnings in the first and second calendar year after referral. DI is measured as any health-related DI in the second year after referral. All models include dummies for hospital×procedure, month and year of referral. Sick leave t1 counts the number of sick leave days in the first year after referral date. Robust standard errors clustered at hospital level in parentheses; *p<0.1, **p<0.05, ***p<0.01.



(a) Sickness absence

(b) Total health care utilization in USD

Figure 2: Outcome estimates by quarter relative to referral.
Notes: Each point in the figures represents coefficients (90% CI) from separate IV estimations of wait days on quarterly sick leave days and health outcomes. All regressions include controls for hospital-procedure, year and month. Y1Q1 is the quarter starting with the date of referral.



(a) Sickness absence

(b) Total health care utilization in USD

Figure 3: Outcomes per month relative to surgery.
Notes: Each point in the figures represents coefficients (90% CI) from separate IV estimations of wait days on sick leave days after surgery per month after surgery. All regressions include controls for hospital-procedure, year and month.

5.3 Heterogeneity

Results indicated that when wait times increase by 10 days, sickness absence increases by almost 3 days in the two years following referral, while no health effects were found. These (local) average results may mask heterogeneous responses. In table 5 we explore if the effects of longer wait time on sick leave days in the two years following referral date are different based on patients' characteristics. One motivation for these subsample analyses is that the marginal cost of being absent from work might be perceived differently for different individuals. In Norway, sick leave compensation rates are 100% for most workers, however, some higher paid private sector workers may incur substantial earnings losses if their incomes are above the benefit cap and they do not have access to top-up insurance. In

Table 5: Heterogeneity in sickness absence

	Outcome: Sick leave in year 1&2 after referral					
	(1) Low educ	(2) High educ	(3) Male	(4) Female	(5) Age<45	(6) Age>45
First stage	0.354***	0.432***	0.401***	0.347***	0.458***	0.324***
	(0.0658)	(0.0577)	(0.0701)	(0.0696)	(0.0630)	(0.0605)
IV	0.258*	0.213**	0.175*	0.368**	0.140	0.364***
	(0.134)	(0.102)	(0.0907)	(0.169)	(0.115)	(0.134)
Observations	17,949	8,009	13,948	12,010	10,876	15,082
Controls	×	×	×	×	×	×
Dep. mean	140.7	94.1	113.8	141.0	116.5	133.5
FS F-stat	29.0	56.2	32.7	24.9	52.8	28.7

Notes: Subsample analyses. All model includes dummies for hospital×procedure, month and year of referral, gender, and linear, quadratic and cubic terms for age. Robust standard errors clustered at hospital level in parentheses; *p<0.1, **p<0.05, ***p<0.01.

addition, individuals might expect that sick leave absence affects job security, future wage growth or promotion opportunities (Ichino and Riphahn, 2005). If the negative impacts of sickness absence on future career prospects are more important in high skill jobs, highly educated patients may face a higher marginal cost of absence even when replacement rates are the same. Moreover, low-skilled jobs may require the worker to be more physically fit, rendering a stronger effect of increased wait time on sickness-related absence.

Table 5, column (1) and (2), shows that the effect of wait time on sick leave days in the first two years following referral date is positive and significant both for people with high education (college degree) and for people with low education. The effect for low education individuals is slightly higher, however, the difference is not significant.

In column (3)-(4) and (5)-(6) we study differentiated effects across gender and age, respectively. The effect of wait time appears to have higher impact to the subsequent sick leave for women compared to men, and for senior patients (age>45) compared to younger patients, however, neither of the differences are statistically significant.

Table 6 shows that the effect of wait time on health care utilization also seems to differ by patient characteristics, however, again very few of the estimates are significantly different. Only when comparing men and women, we find significant differences. One more day of waiting increases the total health care utilization for men by 15 dollars. This is significantly higher than the effect found for women, which is negative yet nonsignificant.

Overall, the effects on total health utilization across different subsamples add additional support that the increase in sick leave is not working through a health channel. The subsamples where the health effects are larger are generally the ones where sick leave effects are smaller.

Table 6: Heterogeneity in health care utilization

	Outcome: Health care utilization ($) in year 1&2 after referral					
	(1) Low educ	(2) High educ	(3) Male	(4) Female	(5) Age<45	(6) Age>45
First stage	0.354***	0.432***	0.401***	0.347***	0.458***	0.324***
	(0.0658)	(0.0577)	(0.0701)	(0.0696)	(0.0630)	(0.0605)
IV	8.210	3.629	14.49*	-2.639	6.286	5.054
	(7.524)	(7.258)	(8.112)	(5.880)	(5.716)	(9.998)
Observations	17,949	8,009	13,948	12,010	10,876	15,082
Controls	×	×	×	×	×	×
Dep. mean	7523.2	7088.8	7035.1	7800.4	5818.4	8521.8
FS F-stat	29.0	56.2	32.7	24.9	52.8	28.7

Notes: Subsample analyses. All models include dummies for hospital×procedure, month and year of referral. Robust standard errors clustered at hospital level in parentheses; *p<0.1, **p<0.05, ***p<0.01.

6 Robustness tests

We perform several tests to ensure internal validity. In order for average wait time to be a valid instrument, it must be uncorrelated with patient characteristics. Though this requirement cannot be empirically confirmed, Table 2 supported the validity of this claim by regressing the instrument on several observable patient characteristics (conditional on dummies for hospital×procedure, year and month of referral), finding no meaningful significant relationships. As a second test, all results are presented with and without a large set of control variables added to the baseline regressions. These are measured the year before referral and include: week fixed effects, linear, quadratic and cubic terms for age, earnings and indicators for female, married, foreign born, education status (high school dropout, high school graduate, college). While all results presented thus far do not contain additional controls, results with additional controls are presented in Tables A2 and A3. The stability of results across models with and without additional controls supports the claim that patient characteristics are unrelated to average wait time.

Figure 1 showed that wait time increases monotonically in average wait time, providing some evidence that the monotonicity assumption may be satisfied. Moreover, in Section 5.3 we split the sample based on characteristics of the patient such as age, sex and education. First stage estimates for each of these subsamples were consistently positive and of considerable size, in line with the monotonicity assumption.

In Table 7 we redo the analysis with different versions of the instrument; by changing the queue window; changing the queue size; and trimming the instrument for extreme values. Our baseline estimations use a time frame of 14 days before and after the referral date of patient i to estimate patient i's average wait time. We further require each combination of hospital×procedure×time period to have at least 5 observations. Varying these arbitrarily set boundaries would reveal if e.g. individual i's own wait time affects the average wait time, or if our results are simply a result of these boundaries.

Our baseline result is presented at the top of Table 7 for comparison; note that this specification yields the strongest first stage F-statistics. In model (1) through (5), we maintain the requirement of 5 observations per combination of hospital-procedure-time period, but vary the time period. Model (1) narrows the queue window to 7 days before and after referral; model (2) expands to 21 days before and after referral; model (3) uses 7 days before referral and no days after; and model (4) and (5) uses 14 and 21 days before referral, respectively, and no days after. Models (6) and (7) retain the window of 14 days before and after, but change the number of observations required for each hospital-procedure-time period to be included in the estimation. Model (6) drops no patients (i.e. $N \geq 1$), while model (7) requires at least 10 patients. Both the first stage and IV estimates are somewhat sensitive to the instrument specification; however importantly, IV-estimates are quantitatively very similar across all models. None of the alternative specifications yield a point estimate below the 95% confidence interval of our baseline specification. Only one of the specifications (model 3) give a point estimate slightly higher than the baseline model's confidence interval; however, the confidence interval of this specification overlaps with the baseline CI.

As a last specification check, we use a trimmed version of the instrument, as shown in model (8) and (9). In these specifications, we recode observations of AWT above the 99th (95th) percentile, using the instrument value at the 99th (95th) percentile as the maximum; similarly for instrument values below the 1st (5th) percentile, we truncate the instrument such that all values below this bound is equal to the instrument value at the 1st (5th) percentile. Both the first stages and IV estimates are almost identical to the baseline model.

Table 7: Specification Checks for Wait Time Estimates

	Sickness absence days 1 and 2 years after referral		
Specification	First stage	IV	FS F-stat/Obs.
(Baseline) ±14 days, N≥ 5	0.379***	0.248***	47.6
	(0.055)	(0.0814)	25,958
Alternative queue window (given N≥ 5)			
(1) ± 7 days	0.243***	0.265**	32.7
	(0.0424)	(0.104)	26,359
(2) ± 21 days	0.471***	0.207**	43.1
	(0.0718)	(0.085)	25,454
(3) 7 days before	0.134***	0.473**	14.5
	(0.035)	(0.199)	25,589
(4) 14 days before	0.228***	0.291**	24.4
	(0.0461)	(0.129)	26,146
(5) 21 days before	0.288***	0.273**	30.0
	(0.0527)	(0.118)	25,998
Alternative queue size (given window size ± 14 days)			
(6) $N \geq 1$	0.379***	0.246***	47.3
	(0.055)	(0.0814)	25,960
(7) $N \geq 10$	0.382***	0.228***	46.5
	(0.056)	(0.0804)	25,915
Trimmed instrument (given window size ± 14 days and N≥ 5)			
(8) 1%/99%	0.391***	0.242***	39.0
	(0.0625)	(0.0874)	25,958
(9) 5%/95%	0.383***	0.254**	31.2
	(0.0685)	(0.0987)	25,958

Notes: Table shows alternative specifications of the instrument. Models (1) through (5) vary queue window, while keeping queue size (i.e. persons per hospital-procedure-window) fixed at $N \geq 5$. Models (6) and (7) vary the queue size for window size ± 14 days. Models (8) and (9) show results when the instrument is trimmed by truncating the instrument at 1%(5%) and 99%(95%). Values of AWT<1%(5%) are set equal to AWT at 1%(5%); values where AWT>99%(95%) = AWT at 99%(95%). All models includes year, month, and procedure×hospital fixed effects. Robust standard errors clustered at hospital level in parentheses; *p<0.1, **p<0.05, ***p<0.01.

7 Conclusions

This paper examines effects of wait time for orthopedic surgeries on patients' health and labor market outcomes by exploiting variation in wait times generated by the idiosyncratic variation in system congestion at the time of a specific patient's entry into the queue. We find that longer wait times for surgery do not have any negative effects on the health outcomes that are included in our dataset. Meanwhile, longer wait times significantly increase sickness absence in the years after referral: the model predicts 2.6 additional absence days for every 10 days spent waiting for surgery. Extended models indicate that effects on sickness absence are persistent beyond the period spent waiting for treatment and the immediate recovery period; there are indications that post-surgery absence increases as well. Our model finds no effect on the probability of entering DI within two years of referral. However, the persistent effect on absence suggests there may still be long term effects on labor supply, either due to adverse effects on health outcomes not captured in our dataset (e.g. pain and minor physical disabilities), or through behavioral effects/habit formation.

In general, long wait times for medical treatments will arise when hospitals' capacity for performing surgeries - staff, equipment etc - is insufficient to match the inflow of new referrals. To completely eliminate wait times, hospitals would have to operate at high levels of standby capacity. When allocating resources to the health care sector, policymakers have to balance costs of shorter waitlists against the gains from faster treatment. Typically, a primary concern in these policy discussions are the potential long medical consequences of delaying time-sensitive treatment.

In the present paper, we have considered an additional effect of longer waitlists: reduced wait times may have fiscal spillovers through effects on labor supply. Our models indicate that this is indeed the case, at least in the short to medium run. Shorter wait times significantly reduce the number of days spent on sickness absence during the first two years after referral. Reduced absence decreases the direct costs of social insurance programs, and increases government revenue from income and payroll taxes.

To be clear, our analysis does not allow us to give a definitive answer to the questions of whether current wait times are too long, or what the optimal wait time should be for any given procedure. The analysis provides only a partial picture of the costs and benefits we would need to analyze to draw welfare conclusions: While our results show that long wait times lead to more sickness absence, we are not able to measure the direct reduction in a patient's utility from delayed recovery. As our health measures are all utilization based, we may fail to pick up effects on subjective well-being. Moreover, (absent) changes in earnings potential might understate true productivity effects as sickness absence is subject to 100% wage replacement. Finally, we lack credible estimates on the costs of reducing wait times for surgery, making a complete cost-benefit analysis impracticable.[20]

It is important to emphasize the local nature of our findings. Our results hold in a setting of universal health care where average wait times are fairly high compared to other OECD countries, and sick leaves are generously compensated. The IV estimates represent a local average treatment effect (LATE) for patients whose wait time for orthopedic surgery is affected by the queue length at the time of referral. This means we need to be cautious in extrapolating the causal effects we estimate to other medical fields or countries with different institutional systems. In particular, while finding no effect of wait time on health outcomes for orthopedic patients, we may expect that delaying treatment could seriously deteriorate health for other more critical health conditions, such as for example cancer.

[20]The costs of decreasing wait times are not easy to measure. Though standby capacity is intuitively costly, Siciliani et al. (2009) and Iversen (1993) argue that there may also be costs to very long waitlists, resulting in a U-shaped relationship between wait times and cost. Siciliani et al. (2009) find in their sample from the English National Health Service that the level of wait time which minimizes total costs is below ten days.

Moreover, average wait times for a particular orthopedic procedure differ substantially across countries. The effects of increased wait time on health outcomes and sick leave are specific for the baseline wait times in our sample, and may not necessarily easily translate to other countries where baseline wait times are different.

Nonetheless, understanding how wait time affects labor market attachment is important for countries considering policies with the intention to decrease hospital queues. There has been an enormous rise in sick leave and disability pension utilization in many countries. It is therefore increasingly important for policy-makers to understand the determinants of participation in health-related social security programs. Waiting for hospital treatment might be a trigger for patients onto such security programs.

References

Aakvik, A., T. H. Holmås, and E. Kjerstad (2015). Prioritization and the elusive effect on welfare – a norwegian health care reform revisited. *Social Science & Medicine 128*, 290 – 300.

Becker, G. S. (1991). *A treatise on the family.* Cambridge, Mass.: Harvard University Press. Enl. ed.

Brage, S., J. Nossen, and O. Thune (2013). Flere blir sykmeldt for symptomer og plager. *Arbeid og velferd 2*, 76–86.

Brownlow, H., S. Benjamin, J. Andrew, and P. Kay (2001). Disability and mental health of patients waiting for total hip replacement. *Annals of the Royal College of Surgeons of England 83*(2), 128.

Calvó-Armengol, A. and M. O. Jackson (2004). The effects of social networks on employment and inequality. *The American Economic Review 94*(3), pp. 426–454.

Dahl, G., A. Kostøl, and M. Mogstad (2014). Family welfare cultures. *Quarterly Journal of Economics*.

Derrett, S., C. Paul, and J. M. Morris (1999). Waiting for elective surgery: effects on health-related quality of life. *International Journal for Quality in Health Care 11*(1), 47–57.

Hamilton, B. H. and R. E. Bramley-Harker (1999). The impact of the nhs reforms on queues and surgical outcomes in england: evidence from hip fracture patients. *The Economic Journal 109*(457), 437–462.

Havnes, T. and M. Mogstad (2011a). Money for nothing? Universal child care and maternal employment. *Journal of Public Economics 95*(11-12), 1455–1465.

Havnes, T. and M. Mogstad (2011b). No child left behind: Subsidized child care and children's long-run outcomes. *American Economic Journal: Economic Policy 3*(2), 97–129.

Hirvonen, J., M. Blom, U. Tuominen, S. Seitsalo, M. Lehto, P. Paavolainen, K. Hietaniemi, P. Rissanen, and H. Sintonen (2007). Evaluating waiting time effect on health outcomes at admission: a prospective randomized study on patients with osteoarthritis of the knee joint. *Journal of evaluation in clinical practice 13*(5), 728–733.

Hultin, H., C. Lindholm, and J. Möller (2012). Is there an association between long-term sick leave and disability pension and unemployment beyond the effect of health status? - a cohort study. *PLoS ONE 7*(4), 1 – 7.

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

Iversen, T. (1993). A theory of hospital waiting lists. *Journal of Health Economics 12*(1), 55 – 71.

Malmivaara, A., U. Hakkinen, T. Aro, M.-L. Heinrichs, L. Koskenniemi, E. Kuosma, S. Lappi, R. Paloheimo, C. Servo, V. Vaaranen, and S. Hernberg (1995, Feb 09). The treatment of acute low back pain–bed rest, exercises or ordinary activity? *The New England journal of medicine 332*(6), 351–355.

Martin, S. and P. C. Smith (1999). Rationing by waiting lists: an empirical investigation. *Journal of Public Economics 71*(1), 141–164.

Martin, S. and P. C. Smith (2003). Using panel methods to model waiting times for national health service surgery. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 166*(3), 369–387.

Mincer, J., S. P. (1974). Family investments in human capital: Earnings of women. *Journal of Political Economy 82(2)*, 76–108.

Moscelli, G., L. Siciliani, and V. Tonei (2016). Do waiting times affect health outcomes? evidence from coronary bypass. *Social Science & Medicine*.

Rees, A. (1966). Information networks in labor markets. *The American Economic Review 56*(1/2), 559–566.

Siciliani, L., A. Stanciole, and R. Jacobs (2009). Do waiting times reduce hospital costs? *Journal of Health Economics 28*(4), 771 – 780.

Underwood, M., R. Firmin, and D. Jehu (1993). Aspects of psychological and social morbidity in patients awaiting coronary artery bypass grafting. *British Heart Journal 69*(5), 382–384.

Appendix A

Table A1: NCSP surgical coding - Chapter N Musculoskeletal system

Surgical procedure	Volume
NB Shoulder and upper arm	4,343
ND Wrist and hand	2,841
NF Hip joint and thigh	1,625
NG Knee and lower leg	12,172
NH Ankle and foot	4,977

Notes: Surgical procedures included in the estimation sample.

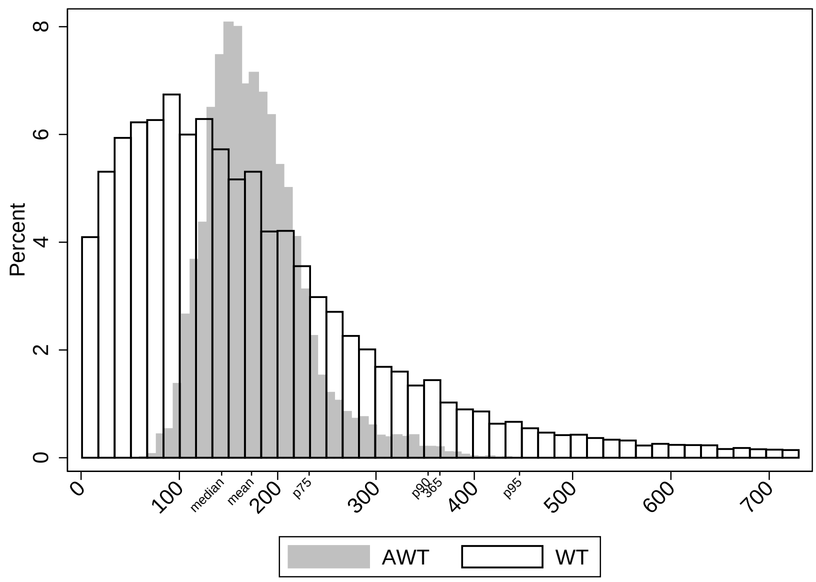


Figure A1: Distribution of Individual Wait Time and Average Wait Time
Notes: Average wait time is a leave-out mean calculated over people referred to the same hospital-procedure within a four-weeks window around person *i*'s referral (gray histogram). Histogram for individual wait time in white. Labeled ticks on the x-axis refers to values on the distribution of individual wait time.

Table A2: Estimates of wait time on labor market outcomes - additional controls

	(1) First stage	(2) Sick leave in year 1	(3) Sick leave in year 2	(4) Sick leave in year 1 and 2	(5) Earnings in year 1 and 2	(6) DI within year 2
Panel A. OLS						
Wait time		-0.0759***	0.109***	0.0333***	-0.00109**	0.000118***
		(0.00935)	(0.00886)	(0.0110)	(0.000503)	(0.0000233)
Panel B. IV						
AWT	0.379***					
	(0.0550)					
Wait time		0.0984*	0.149**	0.248***	-0.00274	-0.000167
		(0.0503)	(0.0604)	(0.0814)	(0.00567)	(0.000217)
Observations	25,958	25,958	25,958	25,958	25,958	25,958
Controls	×	×	×	×	×	×
Dep. mean	178	85.2	41.1	126.4	9.9	0.2
FS F-stat	47.6					

Notes: Labor outcomes estimated with additional control variables. These are: week fixed effects, linear, quadratic and cubic terms for age, earnings and indicators for female, married, foreign born, education status (high school dropout, high school graduate, college). All models includes dummies for hospital×procedure, month and year of referral. Sick leave outcomes are measured in days per year. Robust standard errors clustered at hospital level in parentheses; *p<0.1, **p<0.05, ***p<0.01.

Table A3: Estimates of wait time on health outcomes - additional control variables

	(1) First stage	(2) Outpatient visits	(3) Inpatient days	(4) Same procedure adm	(5) Mortality	(6) Total utilization ($)
Panel A. OLS						
Wait time		0.00179***	0.0000537	-0.000152***	-0.00794**	1.302*
		(0.000576)	(0.000417)	(0.0000169)	(0.00294)	(0.717)
Panel B. IV						
AWT	0.379***					
	(0.0550)					
Wait time		0.00556	0.00513	0.00000651	0.0200	6.793
		(0.00446)	(0.00396)	(0.000155)	(0.0324)	(5.355)
Observations	25,958	25,958	25,958	25,958	25,958	25,958
Controls	×	×	×	×	×	×
Dep. mean	178	5.9	1.9	1.1	3.0	7389.2
FS F-stat	47.6					

Notes: Health outcomes estimated with additional control variables. These are: week fixed effects, linear, quadratic and cubic terms for age, earnings and indicators for female, married, foreign born, education status (high school dropout, high school graduate, college). All models includes dummies for hospital×procedure, month and year of referral. All outcomes except mortality is measured within the two years following referral. Mortality rate is within three years of referral. All models include dummies for hospital×procedure, month and year of referral. Robust standard errors clustered at hospital level in parentheses; *p<0.1, **p<0.05, ***p<0.01.

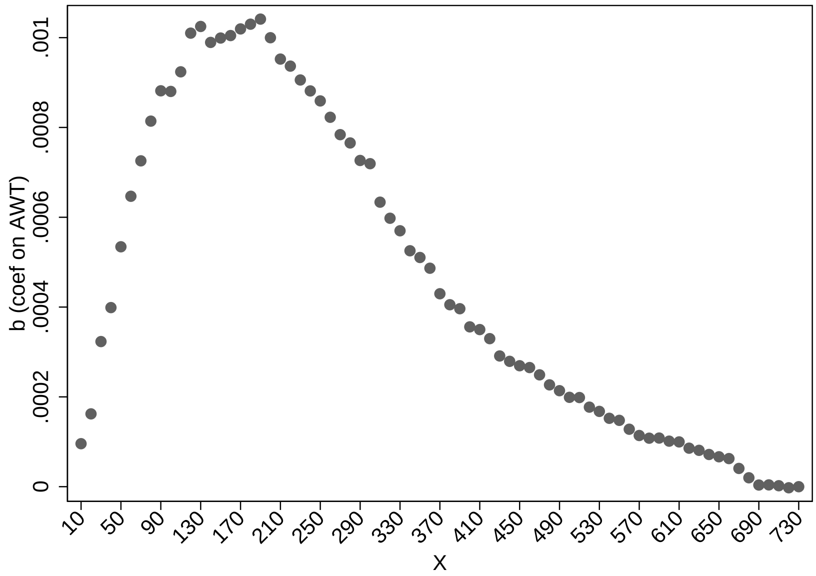
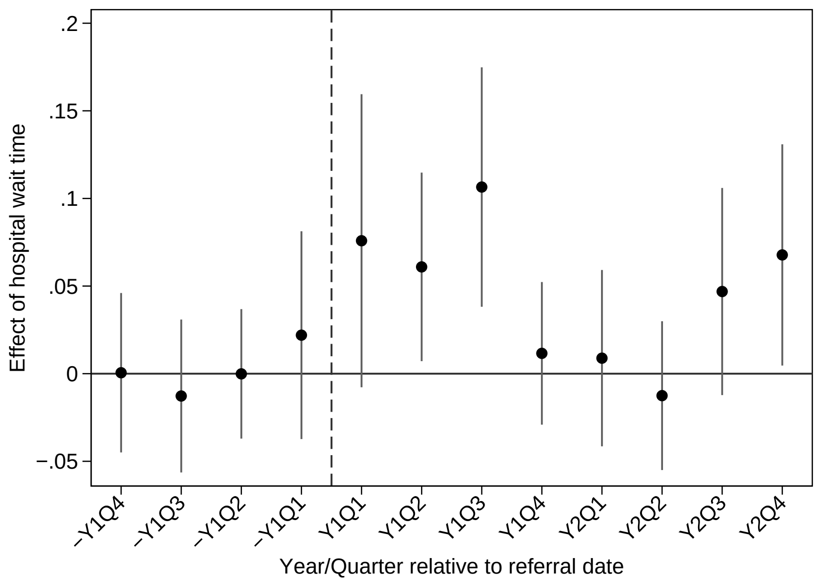


Figure A2: The effect of AWT on a binary indicator for the probability of wait time to be higher than X days.
Notes: Figure shows separate estimations of the effect of AWT on P(WT>X)



(a) All workers

Figure A3: Admission estimates by quarter for patients waiting less than 365 days.
Notes: Each point in the figure represent the beta coefficient from separate regressions of a binary indicator for the probability of waiting more than X days, where X varies from 10 to 710 days, on average wait time (i.e. $Y = P(WT > X) = \beta AWT + X'\alpha$). Regressions include fixed effects for hospital-procedure, year and month. Y1Q1 is the quarter starting with the date of referral

Table B1: Volume and wait time

	Volume of patients per time unit	
	(1)	(2)
	Monthly	Weekly
Wait time	0.0830*	0.00340
	(0.0461)	(0.00723)
Observations	1,949	7,627
Dependent mean	238.4	60.52

Notes: Effects of wait time on volume. Robust standard errors clustered at hospitals in parentheses; *p<0.1, **p<0.05, ***p<0.01.

Appendix B

The exclusion restriction

In order for the IV estimation strategy to be valid, the instrument must satisfy the exclusion restriction. The instrument, AWT, should affect our outcomes only through increased wait times. The exclusion restriction would be violated if, for instance, AWT was correlated with the quality of treatment, as this would open up a second causal channel.

While the exclusion restriction cannot be tested directly, we can examine the data for signs that it may be violated. Our worry is that, when hospitals face higher than normal capacity constraints, this results both in patients waiting longer for surgery (longer wait times for planned procedures) and higher volume of surgeries being performed, possibly reducing the quality of each procedure (if there is a quantity-quality trade-off). In a first step, we examine whether wait times are correlated with total surgery volumes.

To do this, we construct an auxiliary dataset containing all orthopedic procedures performed during the years 2010-2013. This dataset includes emergency admissions and patients who are referred to several procedures in the same referral period. This sample is used to construct datasets containing average wait times for scheduled patients, as well as counts of the total number of surgeries in each time period (week/month). We then estimate the following equation

$$OP_{ht} = \beta \bar{WT}_{ht} + \theta_t + \theta_h + \varepsilon_{ht} \tag{7.1}$$

where OP_{ht} is the total surgery volume at hospital h in period t, $\bar{WT}_{ht}$ is the average wait times of scheduled surgeries performed at hospital h in period t, and θ_t and θ_h are fixed effects for time periods and hospitals.

Results are presented in table B1. When aggregating to the monthly level, there is a small positive effect, which is significant at the 10% level. Quantitatively, the effect is small: 1 day longer average wait times among scheduled patients predicts 0.08 more surgeries per month, or an increase of 0.035% relative to the mean. When aggregating to a weekly level, the estimated effect disappears.

Table B2: Sickness absence for emergency patients

	(1) Sickness absence in year 1 after surgery	(2) Sickness absence in year 1 and 2 after surgery
Wait time	-0.0101	-0.00693
	(0.00853)	(0.0114)
Observations	63,528	63,528
Dependent mean	31.22	38.19

Notes: Effects of wait time on sickness absence. Robust standard errors clustered at hospitals in parentheses; *p<0.1, **p<0.05, ***p<0.01.

Longer average wait times appear to be weakly associated with higher total volume of surgeries. The size of the estimated effects is small, suggesting that this is not likely to impact treatment outcomes and later absence patterns. However, to test this more directly, we estimate a second set of models, studying sickness absence of patients undergoing emergency (unplanned) surgeries.

These patients have, by definition, not spent time in a queue awaiting treatment. As a consequence, the outcomes of this group can be used to estimate placebo models. If the exclusion restriction holds, we would expect to find zero effects of average wait times for this group. Conversely, a positive relationship between long average wait times and later sickness absence would indicate that average wait times influence outcomes through channels other than individual wait times, which would violate the exclusion restriction.

We estimate the following model:

$$Y_{iht} = \beta \bar{WT}_{ht} + \theta_t + \theta_h + \varepsilon_{ht} \tag{7.2}$$

where Y_{iht} is i's sickness absence in the first 12/24 months after surgery, and $\bar{WT}_{ht}$ is average wait times for patients undergoing scheduled surgeries during month t. The estimation sample consists of all unplanned orthopedic surgeries in 2010 and 2011.[21] Results from this exercise are shown in table B2.

The model finds no significant effects of average wait times on sickness absence for unplanned surgeries. This is in line with what we would expect if the exclusion restriction holds. To summarize, we find no evidence that longer wait times have an independent effect on treatment quality (e.g. through congestion effects at the hospital).

[21] We exclude later operations to ensure that we have a full two years of data on sickness absence.