Online Appendix:
Hospital Queues, Patient Health and Labor Supply*

Anna Godøy† Venke F. Haaland‡ Ingrid Huitfeldt§
Mark Votruba¶

*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. We are grateful to Martin E. Andresen, Simon Bensnes, David Bishai, David Card, Sverre Kittelsen, Edwin Leuven, Victoria Marone, Joachim Mowinckel, Mari Rege, Kjetil Telle, and participants at the AEA/ASSA meeting (Philadelphia, 2018), NBER Summer Institute (2017), EEA-ESEM (Geneva, 2016) and the Norwegian School of Economics (Bergen, 2018) for valuable comments. Thanks also to Magne Mogstad for guidance and advice; our paper's structure is inspired by Bhuller et al. (2020).

†Norwegian Institute of Public Health and the Department of Health Management and Health Economics, University of Oslo, anna.godoey@gmail.com

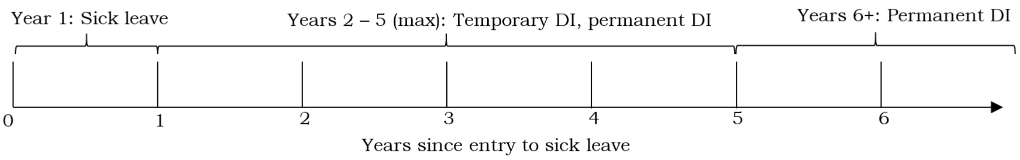
‡University of Stavanger, Norway and Statistics Norway. venke.f.haaland@uis.no

§BI Norwegian Business School and Statistics Norway, Oslo, Norway. ingrid.huitfeldt@gmail.com

¶Economics Department, Weatherhead School of Management, Case Western Reserve University and Center for Health Care Research and Policy at MetroHealth Medical Center. mark.votruba@case.edu

Online Appendix A: Additional results

Figure A1. Stylized benefits timeline



Notes: Figure illustrates time path of health-related benefits for a patient continuously claiming benefits who enters sick leave with full eligibility. See text for details.

Table A1. Bivariate regressions of wait time and congestion on predicted outcomes

Predicted outcomes	(1) Wait time		(2) Congestion	
Predicted Health-related absence t0-t4	-0.024***	(0.004)	0.000	(0.000)
Predicted Permanent DI t4	-0.534***	(0.141)	0.016	(0.016)
Predicted Earnings	-0.001	(0.005)	-0.000	(0.001)
Predicted Employed	0.456***	(0.144)	-0.005	(0.017)
Predicted Earnings if employed	-0.003	(0.005)	-0.000	(0.001)
Predicted Earnings if DI receipt	-0.003	(0.005)	-0.000	(0.001)
Predicted UI days	0.247**	(0.110)	0.013	(0.013)
Predicted GP visits t0-t4	-0.010	(0.053)	0.005	(0.007)
Predicted Hospital days t0-t4	-0.403**	(0.159)	0.008	(0.018)
Predicted Hospital care utilization t0-t4	-0.086***	(0.027)	0.002	(0.003)
Predicted Readmission days t0-t4	-93.163***	(15.685)	0.235	(2.044)
Predicted Emergency admission t0-t4	-4.773**	(2.318)	0.161	(0.281)
Predicted Mortality t4	-0.448***	(0.153)	-0.010	(0.018)
Observations	26,410		26,410	
Dep. mean	190.26		176.55	

Notes: Table shows estimates resulting from bivariate regressions of wait time and congestion on various predicted outcomes. Predicted outcomes are calculated using the covariates in Table ??. All regressions include fixed effects for year-by-referral-month and for hospital-by-procedure. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2. Surgical coding - Chapter N Musculoskeletal system

	(1) Number of observations	(2) Average wait time
Ankle and foot	5,488	256.5
Hip joint and thigh	1,783	194.2
Knee and lower leg	11,291	151.8
Shoulder and upper arm	4,478	183.3
Wrist and hand	3,370	218.5
Total	26,410	190.3

Notes: Surgical procedures included in the estimation sample, using the NCSP coding system.

Table A3. Effects of wait time on health care utilization - standardized variables

	(1) GP visits	(2) GP visits musculoskeletal	(3) Hospital days	(5) Readmission days	(4) Hospital costs	(6) Emergency admissions	(7) Mortality
Panel A: OLS							
Wait time	0.019**	-0.001	0.029***	-0.030***	0.009	-0.006	-0.014**
	(0.008)	(0.009)	(0.009)	(0.010)	(0.008)	(0.008)	(0.006)
Panel B: Reduced form							
Congestion	0.012	0.006	0.014	-0.002	0.008	0.003	-0.001
	(0.011)	(0.010)	(0.010)	(0.012)	(0.009)	(0.010)	(0.011)
Panel C: IV estimates							
Wait time	0.123	0.056	0.141	-0.018	0.076	0.028	-0.006
	(0.111)	(0.103)	(0.097)	(0.119)	(0.092)	(0.102)	(0.106)
Observations	26410	26326	26410	26410	26410	26410	26410

Notes: Table shows the estimated effects of wait time on health outcomes over the five-year window following referral. All outcomes, wait time and congestion are standardized. GP visits indicates the number of visits to the primary care physician. First stage is 0.101 (0.016) ***. GP musculoskeletal is defined as the subset of GP visits that are coded with a musculoskeletal diagnosis code. Hospital days indicates the number of days in hospital, including the day of surgery. Readmission days is the subset of hospital days that are due to visits for the same diagnosis as that for which the patient is awaiting surgery. Hospital costs is the total cost of a patient's hospital utilization measured in Norwegian kroner (NOK). Emergency admissions is the subset of the number of hospital days that are coded as emergency admissions. Mortality is measured as death within five years of referral. All regressions include year-by-referral-month and hospital-by-procedure fixed effects. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A4. Effects of wait time on labor market outcomes - standardized variables

	(1) Health-related absence days	(2) DI receipt	(3) Earnings	(4) Employed	(5) Earnings if employed	(6) Earnings if no DI receipt	(7) UI days
Panel A: OLS							
Wait time	0.008	-0.007	-0.018**	-0.012*	-0.016**	-0.021***	0.001
	(0.008)	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)
Panel B: Reduced form							
Congestion	0.033***	0.034***	-0.017*	-0.013	-0.013	-0.010	0.004
	(0.012)	(0.012)	(0.010)	(0.014)	(0.010)	(0.011)	(0.014)
Panel C: IV estimates							
Wait time	0.322**	0.334**	-0.171*	-0.128	-0.116	-0.101	0.043
	(0.132)	(0.133)	(0.102)	(0.145)	(0.088)	(0.107)	(0.135)
Observations	26410	26410	26410	26410	24266	24989	26410

Notes: Table shows the estimated effects of wait time on labor market outcomes. All outcomes, wait time and congestion are standardized. First stage is 0.101 (0.016) ***. Health-related absence is the total number of health-related absence days (sickness absence, temporary and permanent DI) in the five years following referral, DI receipt is an indicator equal to 100 for patients receiving permanent disability insurance five years after referral, Earnings is earnings in NOK measured five years after referral, Employment is defined as an indicator variable equal to 100 for having positive earnings, UI days is the number of days the individual receives unemployment benefits The sample in column 5 is restricted to patients with positive earnings, and that in column 6 to patients not receiving DI. All regressions include year-by-referral-month and hospital-by-procedure fixed effects. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

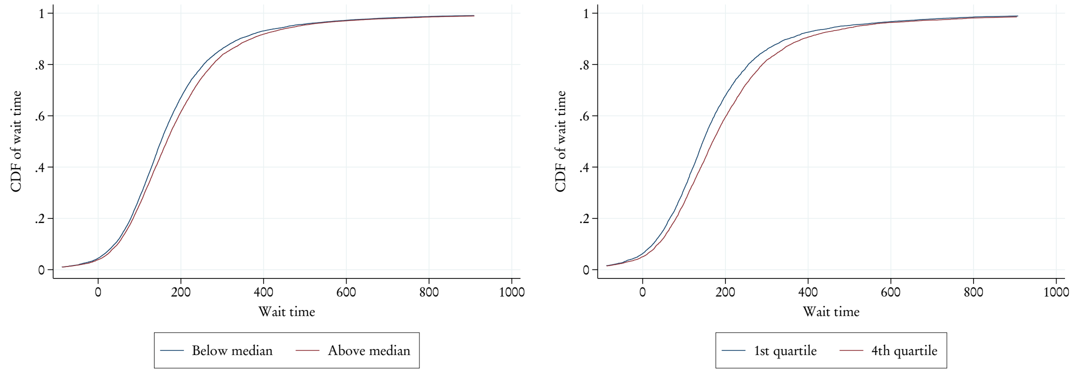
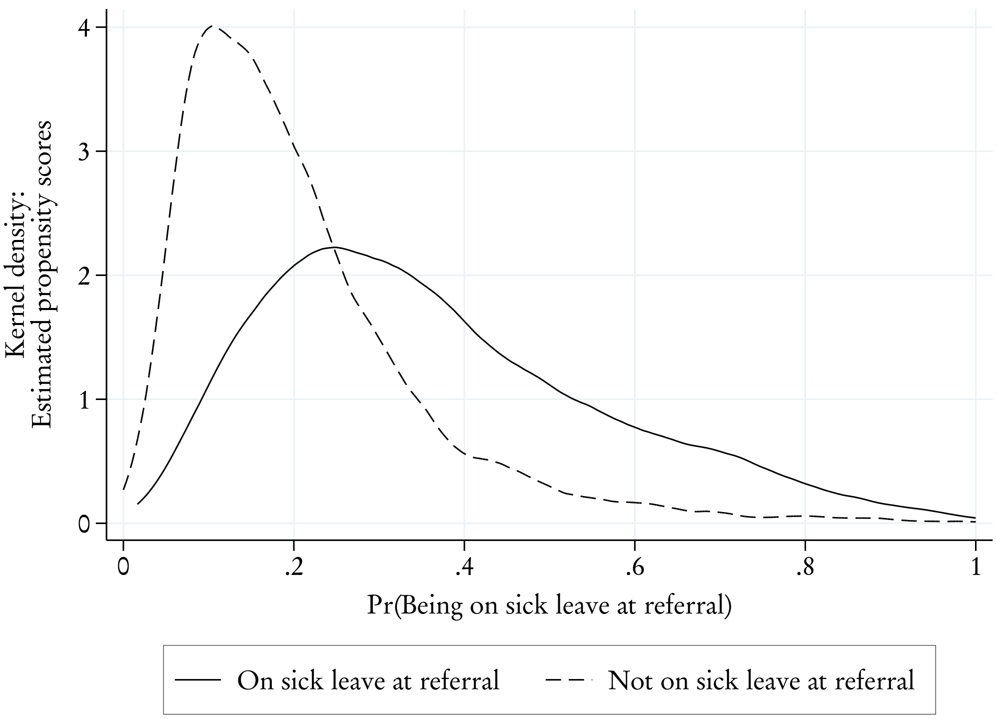
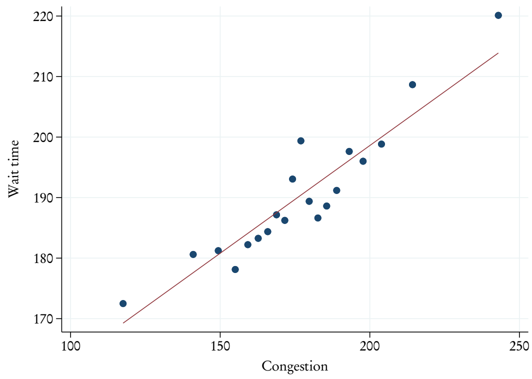


Figure A2. CDFs of wait time by quantile of instrument

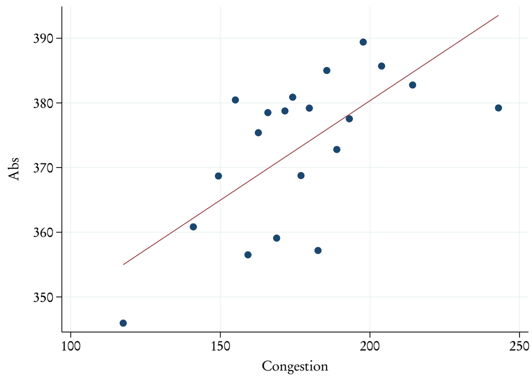
Notes: Kolmogorov–Smirnov tests of equality of distributions yield the following test statistics: (a) Below/above median: $D = 0.058$, $p = 2.5 \times 10^{-19}$, (b) First/fourth quartile: $D = 0.0904$, $p = 3.2 \times 10^{-23}$ Tests of first order stochastic dominance (Kline and Tartari, 2016; Barrett and Donald, 2003) yield the following test statistics: (a) Below/above median: $p = 0.996$, (b) First/fourth quartile: $p = 0.974$

Figure A3. Propensity score by subsamples of sick leave status at referral

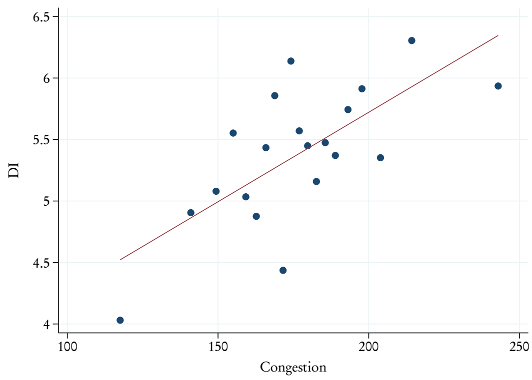




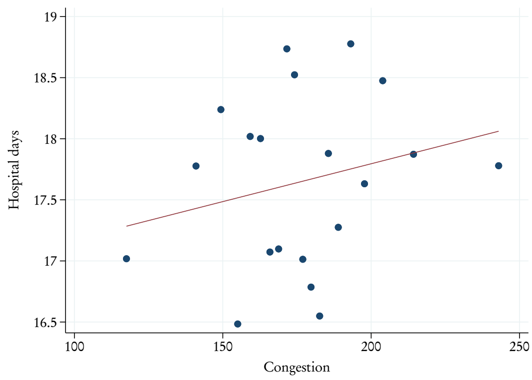
(a) Wait time



(b) Health-related absence



(c) Permanent DI



(d) Hospital days

Figure A4. Binscatter. First stage and reduced form.

Notes: This figure shows binned scatterplots of wait time, congestion and the main outcomes. All variables are purged of time and hospital-procedure fixed effects.

Table A5. Effects by sector and education

	(1)	(2)	(3)	(4)
Sample:	Low education		High education	
	Private	Public	Private	Public
Panel A: First stage				
Congestion	0.352***	0.418***	0.364**	0.299***
	(0.0778)	(0.104)	(0.143)	(0.100)
FS F-stat	20.5	16.1	6.5	8.9
Panel B: Hospital days				
Wait time	0.004	0.028	0.038	0.026
	(0.020)	(0.029)	(0.052)	(0.032)
Dep. mean	16.966	19.693	16.017	18.072
Panel C: Absence days				
Wait time	1.163**	0.702	0.183	-0.339
	(0.575)	(0.742)	(0.626)	(0.623)
Dep. mean	403.919	457.807	188.153	269.559
Panel D: DI year 5				
Wait time	0.039	0.033	0.034	-0.001
	(0.025)	(0.037)	(0.028)	(0.025)
Dep. mean	5.789	7.451	1.648	3.013
Observations	12575	4724	3156	4879

Notes: Table shows the estimated effects of wait time on health (Panel B) and labor market outcomes (Panel C and D) outcomes over the 5-year window following referral. Columns 1 and 2 limit the sample to patients with low education while columns 3 and 4 limit the sample to patients with high education. Additional sample restrictions are applied: in columns 1 and 3 we also restrict the sample to patients working in the private sector, while columns 2 and 4 limit the sample to patients working in the public sector. Low education indicates having high school or less education, high education indicates having longer education than high school. All regressions include year-by-referral-month and hospital-by-procedure fixed effects. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A6. Effects by procedure

	(1)	(2)	(3)	(4)	(5)
Sample:	Shoulder	Hand/wrist	Hip/thigh	Knee	Ankle/foot
Panel A: First stage					
Congestion	0.339***	0.391***	-0.0534	0.327***	0.338***
	(0.106)	(0.134)	(0.237)	(0.0924)	(0.105)
FS F-stat	10.2	8.6	0.1	12.5	10.5
Panel B: Hospital days					
Wait time	-0.0231	0.00200	-0.0177	0.00331	0.0569*
	(0.0253)	(0.0233)	(0.373)	(0.0248)	(0.0299)
Dep. mean	18.497	18.072	26.350	16.095	17.073
Panel C: Health-related absence days					
Wait time	0.453	0.741	-10.79	0.878	1.144
	(0.854)	(0.644)	(49.11)	(0.734)	(0.728)
Dep. mean	536.378	343.464	479.221	331.200	309.931
Panel D: Permanent DI					
Wait time	0.00605	0.0611	-0.893	0.0383	0.0193
	(0.0340)	(0.0414)	(3.972)	(0.0335)	(0.0206)
Dep. mean	8.039	5.964	8.693	4.375	3.845
Observations	4,478	3,370	1,783	11,291	5,488

Notes: Table present the estimated effects of wait time on health and labor market outcomes for subsamples defined by each of the five surgical/medical procedures we study (NCSP/NCMP). All regressions include year-by-referral-month and hospital-by-procedure fixed effects. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01

Table A7. Effects by hospitals with average wait time above/below median

	(1)	(2)
Sample:	Low congestion hospitals	High congestion hospitals
Panel A: First stage		
Congestion	0.240***	0.415***
	(0.0753)	(0.0728)
FS F-stat	10.1	32.5
Panel B: Hospital days		
Wait time	0.0263	0.0166*
	(0.0399)	(0.0101)
Dep. mean	17.187	18.218
Panel C: Health-related absence		
Wait time	-0.351	1.174***
	(0.858)	(0.413)
Dep. mean	359.790	389.505
Panel D: Permanent DI		
Wait time	-0.0135	0.0547***
	(0.0335)	(0.0196)
Dep. mean	5.104	5.720
Observations	14,556	11,854

Notes: Sample is split by the hospital level median wait time. All regressions include year-by-referral-month and hospital-by-procedure fixed effects. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A8. IV estimates for permanent DI in year 5, by graded vs full benefit takeup

	(1) DI	(2) Graded DI	(3) Full DI
Panel A Reduced form			
Congestion	0.015***	0.004	0.012***
	(0.005)	(0.003)	(0.004)
Panel B IV estimates			
Wait time	0.041**	0.012	0.033**
	(0.016)	(0.009)	(0.014)
Observations	26410	26410	26410
Dep. mean	5.381	1.833	3.669

Notes: Table present the estimated effects of wait time on disability pension (graded and full benefit take up). All regressions include year-by-referral-month and hospital-by-procedure fixed effects. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A9. Effects of wait time on health outcomes - with GP fixed effects

	(1) GP visits	(2) GP visits musculoskeletal	(3) Hospital days	(4) Readmission days	(5) Hospital costs	(6) Emergency admissions	(7) Mortality
Panel A: OLS							
Wait time	0.00312*	-0.000	0.00343***	-0.000104***	0.00565	-0.000141	-0.00593**
	(0.00160)	(0.001)	(0.00126)	(0.0000345)	(0.00779)	(0.000149)	(0.00281)
Panel B: Reduced form							
Congestion	0.0126*	0.005	0.00869*	0.0000516	0.0462	0.000525	0.00821
	(0.00659)	(0.004)	(0.00478)	(0.000166)	(0.0335)	(0.000829)	(0.0210)
Panel C: IV estimates							
Wait time	0.0375*	0.014	0.0260*	0.000154	0.138	0.00157	0.0245
	(0.0194)	(0.012)	(0.0141)	(0.000499)	(0.101)	(0.00248)	(0.0624)
Dep. mean	39.362	14.767	17.584	0.328	118.432	1.320	6.709
Observations	25,787	25,787	25,787	25,787	25,787	25,787	25,787

Notes: Table shows the estimated effects of wait time on health outcomes over the five-year window following referral. All regressions include GP fixed effect. First stage F-statistics is 35.9. GP visits indicates the number of visits to the primary care physician. GP musculoskeletal is defined as the subset of GP visits that are coded with a musculoskeletal diagnosis code. Hospital days indicates the number of days in hospital, including the day of surgery. Readmissions days is the subset of hospital days that are due to visits for the same diagnosis as that for which the patient is awaiting surgery. Hospital costs is the total cost of a patient's hospital utilization measured in Norwegian kroner (NOK). Emergency admissions is the subset of the number of hospital days that are coded as emergency admissions. Mortality is measured as death within five years of referral. All regressions include year-by-referral-month and hospital-by-procedure fixed effects. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A10. Effects of wait time on labor market outcomes - with GP fixed effects

	(1) Health-related absence	(2) DI	(3) Earnings	(4) Employed	(5) Earnings if employed	(6) Earnings if no DI receipt	(7) UI days
Panel A: OLS							
Wait time	0.010	-0.001	-0.021	-0.002	-0.015	-0.022	0.001
	(0.022)	(0.001)	(0.015)	(0.001)	(0.015)	(0.016)	(0.003)
Panel B: Reduced form							
Congestion	0.322***	0.011*	-0.133*	-0.007	-0.096	-0.094	0.003
	(0.123)	(0.006)	(0.081)	(0.007)	(0.078)	(0.088)	(0.022)
Panel C: IV estimates							
Wait time	0.961**	0.033*	-0.399	-0.020	-0.265	-0.280	0.009
	(0.408)	(0.019)	(0.251)	(0.022)	(0.219)	(0.267)	(0.067)
Dep. mean	373.127	5.381	471.670	91.899	513.899	494.314	16.341
Observations	25787	25787	25787	25787	23619	24348	25787

Notes: Table shows the estimated effects of wait time on labor market outcomes. All regressions include GP fixed effect. First stage F-statistics is 35.9. Health-related absence is the total number of health-related absence days (sickness absence, temporary and permanent DI) in the five years following referral, DI receipt is an indicator equal to 100 for patients receiving permanent disability insurance five years after referral, Earnings is earnings in NOK measured five years after referral, Employment is defined as an indicator variable equal to 100 for having positive earnings, UI days is the number of days the individual receives unemployment benefits The sample in column 5 is restricted to patients with positive earnings, and that in column 6 to patients not receiving DI. All regressions include year-by-referral-month and hospital-by-procedure fixed effects. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A11. Assessing conditionally random congestion by subsamples of sick leave at referral status

	Not on sick leave at referral				On sick leave at referral			
			Congestion				Congestion	
	mean	sd	b	se	mean	sd	b	se
Age	46.30	9.389	-0.030	0.0262	46.46	9.417	-0.0250	0.039
Female	0.467		-0.349	0.376	0.472		0.203	0.765
Foreign born	0.0842		0.168	0.611	0.107		0.755	0.892
Partner	0.572		1.153**	0.450	0.521		0.170	0.750
Primary education	0.279		0.525	0.655	0.424		0.225	0.984
High school graduate	0.367		-0.200	0.493	0.388		-0.810	1.074
College	0.354		0	.	0.188		0	.
Office job	0.471		0.206	0.455	0.274		-0.487	0.899
Earnings t-2	563.8	316.2	0.001	0.001	464.4	199.1	0.003	0.003
Earnings t-1	560.0	312.2	-0.001	0.001	440.8	208.5	-0.002	0.003
On sick leave at referral	0				1			.
Permanent DI t-1	0				0.448	6.681	0.055	0.044
Health-related absence t-2	16.17	36.02	0.006	0.007	29.57	47.21	0.001	0.012
Health-related absence t-1	19.63	50.38	-0.009*	0.005	118.5	104.7	0.002	0.004
GP visits t-2	5.412	5.642	-0.048	0.049	7.019	6.575	0.023	0.085
GP visits t-1	7.067	6.062	0.037	0.048	12.31	8.028	0.045	0.060
Hospital days t-2	1.757	4.040	-0.061	0.054	1.803	4.189	0.011	0.096
Hospital days t-1	2.023	4.416	0.047	0.060	3.669	7.956	-0.006	0.057
Dep. mean	177.340				174.129			
Observations	19,942				6,468			
F-statistic for joint significance	0.949				0.692			
p-value for joint significance	0.514				0.796			

Notes: The table shows means of observable patient characteristics measured prior to referral, alongside estimates of congestion on these same characteristics, for patients not on sick leave at referral and for patients on sick leave at referral. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A12. Health estimates by sick leave status at referral - reweighted

	(1) Wait time	(2) GP visits	(3) GP visits musculoskeletal	(4) Hospital days	(5) Readmission days	(6) Hospital costs	(7) Emergency admissions	(8) Mortality
Panel A. On sick leave at referral								
Congestion (FS)	0.345***							
	(0.107)							
FS F-stat	10.3							
Wait time		0.008	0.006	-0.011	-0.000	-0.033	-0.001	0.012
		(0.041)	(0.025)	(0.030)	(0.001)	(0.204)	(0.005)	(0.128)
Observations		6468	6468	6468	6468	5351	5565	6468
Dep. mean		46.437	20.728	20.117	0.390	137.152	1.498	8.506
Panel B. Not on sick leave at referral								
Congestion (FS)	0.400***							
	(0.077)							
FS F-stat	26.9							
Wait time		0.041	0.009	0.018	0.000	0.048	0.000	0.014
		(0.039)	(0.015)	(0.022)	(0.001)	(0.116)	(0.003)	(0.076)
Observations		19942	19942	19942	19942	18915	19424	19942
Dep. mean		44.529	15.886	18.860	0.335	124.558	1.423	7.291

Notes: Table shows subsample estimated effects of wait time on labor market outcomes for patients on sick leave at referral (Panel A) and patients not on sick leave at referral (Panel B). We weight each sample by the estimated propensity scores so that they are similar to the opposite group of patients with respect to observable characteristics. Health-related absence is the total number of health-related absence days (sickness absence, temporary and permanent DI) in the five years following referral, DI receipt is an indicator equal to 100 for patients receiving permanent disability insurance five years after referral, Earnings is earnings in NOK measured five years after referral, Employment is defined as an indicator variable equal to 100 for having positive earnings, UI days is the number of days the individual receives unemployment benefits The sample in column 5 is restricted to patients with positive earnings, and that in column 6 to patients not receiving DI. All regressions include year-by-referral-month and hospital-by-procedure fixed effects. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A13. Labor market outcomes by sick leave at referral status - reweighted

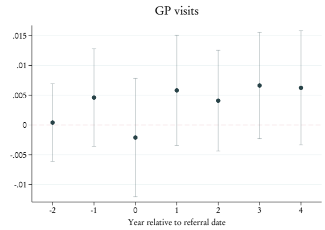
	(1) Wait time	(2) Health-related absence days	(3) DI receipt	(4) Earnings	(5) Employed	(6) Earnings if employed	(7) Earnings if DI receipt	(8) UI days
Panel A. On sick leave at referral								
Congestion (FS)	0.345^{***}							
	(0.107)							
FS F-stat	10.3							
Wait time		2.544^{***}	0.112^{*}	-0.762	-0.115^{*}	-0.269	-0.580	-0.046
		(0.949)	(0.060)	(0.474)	(0.060)	(0.368)	(0.541)	(0.096)
Dep. mean		706.171	11.779	356.534	85.160	419.634	393.942	18.378
Observations		6468	6468	6468	6468	5351	5565	6468
Panel B. Not on sick leave at referral								
Congestion (FS)	0.400^{***}							
	(0.077)							
FS F-stat	26.9							
Wait time		0.489	0.015	-0.149	0.020	-0.210	-0.118	0.003
		(0.394)	(0.020)	(0.162)	(0.027)	(0.146)	(0.170)	(0.063)
Dep. mean		324.081	4.141	423.921	92.596	458.390	439.117	20.041
Observations		19942	19942	19942	19942	18915	19424	19942

Notes: Table shows the estimated effects of wait time on health outcomes for the subsamples patients on sick leave at referral (Panel A) and patients not on sick leave at referral (Panal B). We weight each sample by the estimated propensity scores so that they are similar to the opposite group of patients with respect to observable characteristics. All regressions include year-by-referral-month and hospital-by-procedure fixed effects. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

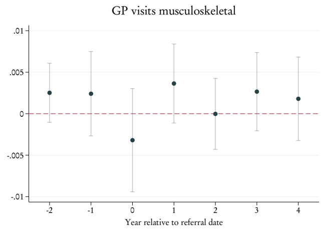
Table A14. Effects by medical history and absence history

	(1)	(2)	(3)	(4)	(5)	(6)
Sample:	High predicted SL	Low predicted SL	High predicted SL		Low predicted SL	
			Not on SL	On SL	Not on SL	On SL
First stage						
Congestion	0.381***	0.330***	0.429***	0.304**	0.311***	0.627***
	(0.068)	(0.071)	(0.088)	(0.122)	(0.067)	(0.211)
FS F-stat	31.6	21.4	23.7	6.2	21.2	8.9
Hospital days						
Wait time	0.012	0.021	0.016	0.019	0.032	-0.045
	(0.017)	(0.017)	(0.021)	(0.034)	(0.020)	(0.045)
Dep. mean	20.131	15.169	19.262	21.522	14.619	19.841
Health-related absence days						
Wait time	1.361**	0.210	0.647	2.619*	-0.119	1.655*
	(0.552)	(0.280)	(0.411)	(1.405)	(0.280)	(0.896)
Dep. mean	541.172	205.083	355.225	838.578	168.481	516.698
DI receipt						
Wait time	0.048**	0.028**	0.010	0.130*	0.013	0.086*
	(0.024)	(0.014)	(0.021)	(0.078)	(0.012)	(0.052)
Dep. mean	9.034	1.727	4.640	16.063	1.193	6.268
Observations	13205	13205	8125	5080	11817	1388

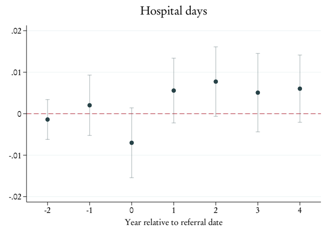
Notes: Table shows the estimated effects of wait time on health and labor market outcomes in the 5-year window following referral. The sample is split according to predicted probability of receiving sickness absence benefits at referral. The predicted sickness absence is the same as that used in Section 6, where absence is predicted from the predetermined observable characteristics in Table **??**, omitting the lagged absence variables. In columns 3-6 we additionally split the sample on patients actually on sick leave at referral and not on sick leave at referral. All regressions include fixed effects for year-by-referral-month and hospital-by-procedure. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$



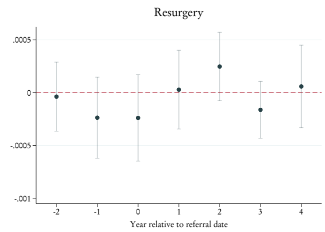
(a) GP visits



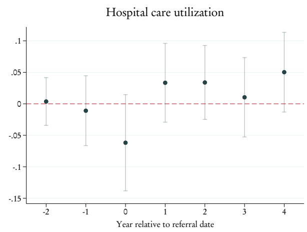
(b) Orthopedics visits to GP



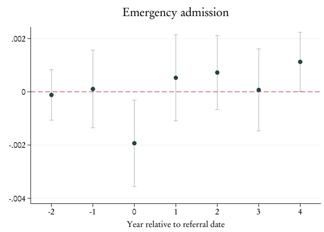
(c) Hospital days



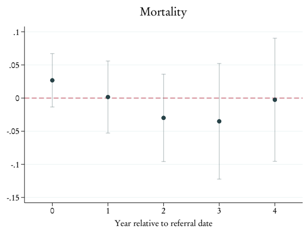
(d) Readmissions



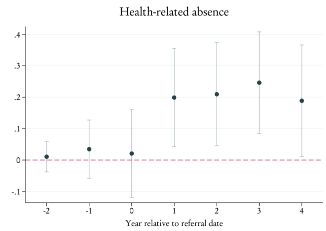
(e) Hospital utilization



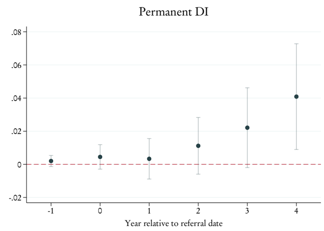
(f) Emergency admissions



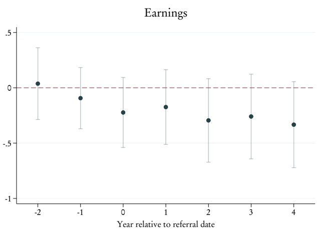
(g) Mortality



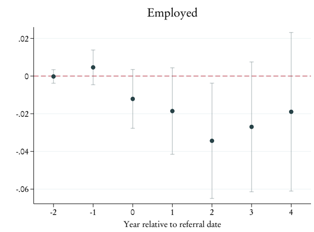
(h) Health-related absence



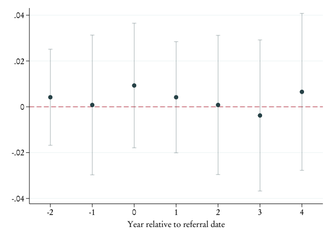
(i) Permanent DI



(j) Earnings



(k) Employed



(l) UI days

Figure A5. Effects of wait time by years since referral

Note: The figure plots the estimated IV effects of wait time for each year relative to the referral year.

Table A15. Characteristics of compliers

		Baseline model		Binary model			
	$pr(X=x)$	FS_x	$\frac{FS_x}{FS}$	FS_x	$\frac{FS_x}{FS} = \frac{pr(X=x\|compl)}{pr(X=x)}$	$\frac{FS_x}{FS} pr(X=x) = pr(X=x\|compl)$	$E(X\|DI)$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
All	1.000	0.356	1.000	0.0010	1.000	1.000	
		(0.056)		(0.0001)			
Education							
Low	0.687	0.372	1.045	0.0011	1.072	0.736	0.769
		(0.063)	(0.0918)	(0.0002)	(0.0691)	(0.0477)	
High	0.313	0.319	0.896	0.0009	0.855	0.268	0.231
		(0.084)	(0.199)	(0.0002)	(0.148)	(0.0466)	
Gender							
Male	0.532	0.335	0.942	0.0011	1.098	0.584	0.398
		(0.051)	(0.143)	(0.0002)	(0.103)	(0.0547)	
Female	0.468	0.362	1.017	0.0009	0.864	0.405	0.602
		(0.087)	(0.148)	(0.0002)	(0.106)	(0.0495)	
Age							
$\geq$45	0.567	0.289	0.812	0.0010	0.971	0.550	0.799
		(0.067)	(0.125)	(0.0002)	(0.0888)	(0.0506)	
<45	0.433	0.443	1.245	0.0011	1.037	0.449	0.201
		(0.085)	(0.161)	(0.0002)	(0.118)	(0.0512)	
Occupation							
Office	0.423	0.298	0.838	0.0009	0.897	0.379	0.302
		(0.089)	(0.168)	(0.0002)	(0.115)	(0.0485)	
Manual	0.249	0.459	1.289	0.0013	1.267	0.316	0.127
		(0.118)	(0.241)	(0.0002)	(0.175)	(0.0434)	

Table shows characteristics of compliers. Population shares are shown in column (1). Column (2) and (3) show group-specific first stages, and the ratio of these to the overall first stage. Columns (4) and (5) show group-specific first stages with binary wait time (wait time above median) and the ratio of these to the overall first stage. Column (6) shows the probability of being in a subgroup conditional on being a complier. Column (7) shows the shares of all workers entering DI for any type of diagnosis, by characteristics.

Online Appendix B: Robustness

Table B1 summarizes results from a set of robustness and specification checks, as discussed in Section **??**. Table B2 presents estimates from a model where patients with very long waits are dropped from the sample.

The stability of results across models with and without additional controls supports the claim that patient characteristics are unrelated to the instrument. To further validate the independence assumption, we also estimate the baseline model with GP fixed effects on all outcomes. This addresses the concern that estimates are biased if GPs help patients choose hospitals based on the expected wait times. Reassuringly, however, these models with GP fixed effects produce results that are very similar to our baseline model.

Appendix Table B3 columns (1) - (3) present models estimated on a sample excluding patients with a history of orthopedic surgery in different windows prior to referral. We might worry that the identifying assumptions of our model are less likely to hold for these patients – for instance, they, or their referring doctors, might have greater access to information as to which hospitals have shorter queues. However, there appears to be no difference between our baseline estimates and estimates from samples which exclude patients with an orthopedic history.

Our baseline estimations use a time frame of 30 days before the referral date of patient i to estimate patient i's average wait time. Figure B1 illustrates the effects of varying this window, plotting IV estimates of the effects on absence days and disability where the instrument is constructed using pre-referral windows of 14 to 50 days. Overall, results are robust to choice of window, though estimated effects tend to be less significant for very short windows (14 days), possibly reflecting increased noise associated with small sample sizes.

Our preferred model defines all outcomes relative to the referral date. In Appendix Table B4, we present alternative models where outcomes are instead defined relative to surgery date. The results generated by this model are consistent with the findings of our preferred model. While longer wait times have no significant effects on post-surgery health care utilization, longer waits do significantly increase post-surgery sick leave, as well as the probability of permanent DI receipt. In a related decomposition exercise, presented in Appendix Table B5 we estimate effects of wait times on pre- and post-surgery absence days and healthcare utilization. These models also indicate that longer wait times significantly increase post-surgery absence, with no significant effects on post-surgery healthcare utilization.

While the F-statistic of our preferred instrument ($F = 40$) is well above the conventional threshold for weak instrument, Lee et al. (2021) find that valid inference requires an F-statistic greater than 100. To assess the robustness of our findings, we have calculated weak instrument-robust confidence intervals. Ap-

Table B1. Robustness

Specification/ sample	(1) Hospital catchment area	(2) Extra controls	(3) Date FE	(4) Max wait time 2 years	(5) Min peers in IV sample 5 peers+	(6) 10 peers+	(7) No labor market restriction	(8) Excluding hips
Panel A: Hospital days								
Wait time	0.042*	0.017	0.011	0.022	0.022*	0.019	0.006	0.018
	(0.021)	(0.012)	(0.012)	(0.015)	(0.013)	(0.015)	(0.017)	(0.011)
Dep. mean	17.650	17.650	17.650	17.575	17.569	17.387	21.659	17.020
Panel B: Health-related absence days								
Wait time	1.317**	0.811**	0.780**	1.165***	0.873**	0.752	0.820**	0.767**
	(0.641)	(0.357)	(0.335)	(0.451)	(0.404)	(0.468)	(0.370)	(0.334)
Dep. mean	373.127	373.127	373.127	373.176	370.412	356.861	626.165	365.446
Panel C: Permanent DI								
Wait time	0.070**	0.041**	0.037**	0.053**	0.035**	0.034*	0.044**	0.032**
	(0.032)	(0.016)	(0.015)	(0.021)	(0.017)	(0.020)	(0.022)	(0.015)
Dep. mean	5.381	5.381	5.381	5.423	5.249	5.032	20.600	5.141
Observations	26,400	26,410	26,410	25,816	24,768	20,728	39,399	24,627

Notes: Table shows the estimated effects of wait time on health (Panel A) and labor market outcomes (Panel B and C) over the 5-year window following referral. All regressions include year-by-referral-month and hospital-by-procedure fixed effects. The specifications mirror the baseline specification reported in Tables **??** and **??**, but in column 1 the instrument is defined using catchment areas based on individuals' places of residence, and in columns 2 and 3 we include additional controls such as: week fixed effects, linear, quadratic and cubic terms for age and earnings, indicators for female, married, foreign-born and education status (column 2) and date fixed effects (column 3). In columns 4-6 and 8 we apply some additional sample restrictions: column 4 excludes patients with longer wait time, columns 5 and 6 exclude hospital-procedure groups in which the number of patients in the referral window dips below 5 or 10 peers at any time, column 8 excludes hip replacement procedures. In column 7 we extend the sample to include patients with no or weak labor market attachment. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B2. Exclude delays from IV

	(1) First stage	(2) Hospital days	(3) Total absence	(4) Permanent DI
Congestion	0.342***			
	(0.0580)			
Wait time		0.00732	0.682**	0.0314**
		(0.0140)	(0.388)	(0.0179)
Observations	16,950	16,950	16,950	16,950
Dep. mean		17.08	366.8	5.286
FS F-stat	35.8			

Notes: All regressions uses our estimation sample defined in Section 3.3.2, but we exclude patients with delayed procedures from the instrument sample defined in Section 3.3.1. All regressions include year-by-referral-month and hospital-by-procedure fixed effects. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

pendix Table B6 presents the 95% Anderson Rubin confidence intervals associated with the IV estimates of effects of wait time on health and labor market outcomes. These results indicate that the effects on five year absence and DI entry remain statistically significant at the five percent level when we adjust for weak instruments.

We examine the robustness of noise to our instrument by employing an alternative "shrinkage-adjusted" instrument, defined as follows:

$$ShrinkIV_{ih} = \left(\widehat{\frac{1}{1 + \tau_h / Npeers_i}}\right)\left(Congestion_{ih} - \overline{Congestion_h}\right) \qquad (1)$$

where τ_h represents the ratio of patient-level variance to congestion-related variance within each procedure group and $Npeers_i$ is the count of peer-patients on which $Congestion_{ih}$ is measured.[1]

The term τ_h expression is often called the "shrinkage factor," as the practical effect of this adjustment is to shrink the original *Congestion* measure towards the mean value for its group.[2]

To estimate τ_h and $Var(\hat{\tau_h})$, we employ the method recommended by Guarino et al. (2015), jointly estimating the relevant variance terms by modeling wait times as a function of three independent terms: a random hospital effect, a random peer

[1]In principle, τ might vary across different hospital-procedure group combinations, not only across procedure groups. However, attempts the estimate values of τ at that level proved too high a demand on our data, producing a handful of outlier τ_{hp} estimates.

[2]An analogous issue exists in the teacher evaluation literature, where empirical Bayes's (EB) estimation of teachers' "value-added" is frequently employed to decrease classification errors across teachers (Guarino et al. 2015). See, e.g., Chetty et al. (2014); Corcoran et al. (2011); Jacob and Lefgren (2008); Kane and Staiger (2008); McCaffrey et al. (2004) for other applications of EB estimation in the teacher value-added literature.

Table B3. Robustness - exclude patients with recent orthopedic surgeries

	(1)	(2)	(3)
Sample:		No orthopedic surgery last	
	180 days	365 days	730 days
Panel A: First stage			
Congestion	0.334***	0.337***	0.331***
	(0.0599)	(0.0615)	(0.0634)
Dep. mean	191	191	190
FS F-stat	31.1	29.9	27.2
Panel B: Hospital days			
Wait time	0.0125	0.00895	0.00865
	(0.0127)	(0.0129)	(0.0139)
Dep. mean	17.471	17.384	17.253
Panel C: Health-related absence days			
Wait time	0.828**	0.871**	0.861**
	(0.385)	(0.384)	(0.391)
Dep. mean	367.888	365.535	366.371
Panel D: Permanent DI			
Wait time	0.0379**	0.0359**	0.0317*
	(0.0184)	(0.0179)	(0.0187)
Dependent mean	5.240	5.134	5.109
Observations	25,459	24,426	23,020

Notes: Three different sample restrictions are applied restricting the sample to patients with no orthopedic surgery the last: 180 days (column 1), 365 days (column 2), and 730 days (column 3). All regressions include year-by-referral-month and hospital-by-procedure fixed effects. Standard errors are clustered at hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B4. Outcomes measured post surgery

	(1) Health-related absence year 1 to 3	(2) Hospital days year 1 to 3	(3) DI year 3
Panel A. Full sample			
Wait time	0.613***	0.011	0.075***
	(0.216)	(0.009)	(0.025)
Dep. mean	227.698	10.249	12.242
Observations	26410	26410	26410
Panel B. On sick leave at referral			
Wait time	1.733**	-0.006	0.253**
	(0.709)	(0.023)	(0.099)
Dep. mean	451.569	12.520	30.226
Observations	6468	6468	6468
Panel C. Not on sick leave at referral			
Wait time	0.252	0.017	0.021
	(0.165)	(0.011)	(0.019)
Dep. mean	155.087	9.512	6.409
Observations	19942	19942	19942

Notes: Table shows the estimated effects of wait time on outcomes measured post surgery; health related absence in year 1 to 3 (column 1), hospital days in year 1 to 3 (column 2) and Disability pension in year 3 (column 3). The sample in Panel B is restricted to patients on sick leave at referral and to patients not on sick leave at referral in Panel C. All regressions include year-by-referral-month and hospital-by-procedure fixed effects. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B5. IV estimates. Outcomes measured over the 5 years following referral, split by before and after surgery

	(1) from entry to 5 years after entry	(2) from entry to surgery	(3) from surgery to 5 years after entry
Panel A: Health-related absence			
Wait time	0.864**	0.193***	0.671**
	(0.353)	(0.060)	(0.324)
Dep. mean	373.383	41.078	332.306
Panel B: Hospitals days			
Wait time	0.017	0.005**	0.011
	(0.012)	(0.002)	(0.011)
Dep. mean	17.713	3.150	14.563
Observations	26410	26410	26410

Notes: Table shows the estimated effects of wait time on pre- and post-surgery absence days (Panel A) and healthcare utilization (Panel B). Outcomes in column 1 are measured in the five years following referral (baseline model). In column 2 outcomes are measured from referral date to surgery date. In column 3 outcomes are measured in the five years following surgery data. All regressions include year-by-referral-month and hospital-by-procedure fixed effects. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

group-specific congestion effect, and random patient-level effects, each assumed to be normally-distributed. I.e., for patient i in peer group g in hospital h, we assume wait time takes the form

$$WaitTime_{igh} = \theta_h + \omega_g + \varepsilon_i \tag{2}$$

with θ_h, ω_g, and ε_i representing the hospital, peer group, and patient-level components, respectively, and $\tau_p = Var(\omega)/Var(\varepsilon)$. Estimation of (2) presents a problem in our context, in that our congestion measure is defined for evolving sets of peer groups, where patients are placed in multiple peer groups. We therefore abstract from that aspect of our setting for the purposes of estimating τ_p, and fit eq. (2) after assigning each patient to fixed 30-day "peer groups" (within each hospital). Eq. (2) was fit via maximum likelihood estimation, with the resulting estimates used to construct the estimate shrinkage factors for subject-specific values of $Npeers$.

This procedure was repeated for each procedure group to produce estimated shrinkage factors for all subjects. Appendix tables B7 and B8 report results analogous to those in Tables 4 and 5, but employing the shrinkage-adjusted IV defined

Table B6. Weak instrument robust confidence intervals

	Point estimate	95% AR confidence interval
Outcome: Health-related absence	0.866**	[.236795, 1.77409]
	(0.353)	
Outcome: DI	0.0409**	[.011919, .082847]
	(0.0163)	
Outcome: Earnings	-0.334*	[-.79746, .050658]
	(0.198)	
Outcome: Employment	-0.0190	[-.06915, .022693]
	(0.0215)	
Outcome: Earnings if employed	-0.217	[-.56141, .114258]
	(0.164)	
Outcome: Earnings if DI	-0.194	[-.640711, .235724]
	(0.205)	
Outcome: UI days	0.0171	[-.112744, .117167]
	(0.0538)	
Outcome: GP visits	0.0206	[-.019905, .059674]
	(0.0186)	
Outcome: GP visits musculoskeletal	0.00489	[-.014541, .023613]
	(0.00892)	
Outcome: Hospital days	0.0174	[-.007604, .043272]
	(0.0119)	
Outcome: Hospital costs	0.0660	[-.101956, .24038]
	(0.0801)	
Outcome: Emergency admissions	0.000486	[-.003538, .004364]
	(0.00185)	
Outcome: Mortality	-0.00246	[-.10198, .100812]
	(0.0474)	

Note: Table presents IV estimates of effects of wait time on health and labor market outcomes, together with Anderson-Rubin weak instrument robust confidence intervals. All regressions include year-by-referral-month and hospital-by-procedure fixed effects. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

by eq. (1).

As Ballou et al. (2004) and others have shown, if τ_h was a known parameter, $ShrinkIV_{ih}$ would represent the best linear unbiased predictor of the mean (true) congestion faced by i's peer patients. Therefore, we should anticipate that our shrinkage-adjusted instrument would improve the power of our first-stage model and the precision of estimated effects of wait time. In general, we would not anticipate a dramatic change in the magnitude of estimated effects of wait time, as the deficiency the $ShrinkIV$ instrument corrects is (in principle) one of efficiency rather than bias. If effects of wait time are either homogeneous or uncorrelated with the patient shrinkage factors, we would expect no change in the IV estimates, as the shrinkage adjustment would have proportional effects on both the first-stage and the reduced-form estimates. Small changes in the IV estimates are perhaps more likely, as the $ShrinkIV$ instrument alters the weight contributed by different individuals to the estimated LATE, reducing the effective contribution of patients with higher values of $\tau_h/Npeers_i$.

The estimation results arrived at by means of our shrinkage-adjusted instrument, presented in Appendix Tables B7 and B8, largely confirm these expectations. Judging by the F-statistic, the first-stage is about 15% more powerful when the shrinkage-adjusted instrument is used.

The IV estimates are generally robust to the alternative instrument, although the estimated wait time effect on DI receipt is somewhat smaller, and the estimated effect on hospital days is now modestly larger and significant at the 10% level.

In order for our IV strategy to be valid, the exclusion restriction must hold. Violations of the exclusion restriction could occur when hospitals face higher than normal capacity constraints, if this results both in patients waiting longer for surgery (longer wait times for planned procedures) and higher volumes of surgery being performed, possibly reducing the quality of each procedure (if there is a quantity-quality trade-off).[3] To examine this, we construct an auxiliary dataset containing all orthopedic procedures performed during the years 2010-2011. This dataset includes emergency admissions and patients who are referred for 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 procedures in each time period (week/month). We then estimate a set of models for studying the sickness absence of patients undergoing emergency (unplanned) surgery. These patients have, by definition, not spent time in a queue awaiting treatment. As a consequence, the outcomes for this group

[3] Wait times could also be positively correlated with the quality of treatment, for instance if patients are willing to accept longer waits for treatment of higher perceived quality. To the extent that quality is fixed over time, this will be absorbed by the procedure by group fixed effects. If patients respond to time-varying changes in (perceived) quality of care, our estimates could reflect changes in treatment quality as well as effects of wait time., however the results from our hospital catchment area models suggest this is not the main driver.

Table B7. Effects of wait time on health outcomes - shrunk estimates

	(1) Wait time	(2) GP visits	(3) GP visits musculoskeletal	(4) Hospital days	(5) Readmission days	(6) Hospital costs	(7) Emergency admissions	(8) Mortality
Panel A: Reduced form								
ShrinkIV		0.018	0.002	0.019*	-0.000	0.055	0.001	0.019
		(0.016)	(0.009)	(0.010)	(0.000)	(0.075)	(0.002)	(0.048)
Panel B: IV estimates								
ShrinkIV (FS)	0.773***							
	(0.113)							
FS F-stat	46.7							
Wait time		0.023	0.003	0.025*	-0.000	0.072	0.001	0.024
		(0.021)	(0.012)	(0.013)	(0.001)	(0.095)	(0.002)	(0.062)
Dep. mean	190	39.345	14.753	17.650	0.329	119.006	1.330	6.778
Observations	26,410	26,410	26,410	26,410	26,410	26,410	26,410	26,410

Notes: Table shows the estimated effects of wait time on health outcomes over the five-year window following referral when using the shrunk IV. All regressions include GP fixed effect. First stage F-statistics is 35.9. GP visits indicates the number of visits to the primary care physician. GP musculoskeletal is defined as the subset of GP visits that are coded with a musculoskeletal diagnosis code. Hospital days indicates the number of days in hospital, including the day of surgery. Readmissions days is the subset of hospital days that are due to visits for the same diagnosis as that for which the patient is awaiting surgery. Hospital costs is the total cost of a patient's hospital utilization measured in Norwegian kroner (NOK). Emergency admissions is the subset of the number of hospital days that are coded as emergency admissions. Mortality is measured as death within five years of referral. All regressions include year-by-referral-month and hospital-by-procedure fixed effects. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B8. Effects of wait time on labor market outcomes - shrunk estimates

	(1) Wait time	(2) Health-related absence	(3) DI receipt	(4) Earnings	(5) Employed	(6) Earnings if employed	(7) Earnings if no DI receipt	(8) UI days
Panel A: Reduced form								
ShrinkIV		0.666**	0.0259**	-0.260	-0.019	-0.152	-0.177	0.00849
		(0.294)	(0.0128)	(0.186)	(0.019)	(0.169)	(0.195)	(0.0442)
Panel B: IV estimates								
ShrinkIV (FS)	0.773***							
	(0.113)							
FS F-stat	46.7							
Wait time		0.861**	0.0336**	-0.337	-0.025	-0.190	-0.224	0.0110
		(0.404)	(0.0171)	(0.249)	(0.026)	(0.211)	(0.251)	(0.0564)
Dep. mean	190	373.127	5.381	471.670	91.882	513.899	494.314	16.341
Observations	26,410	26410	26410	26410	26410	24266	24989	26410

Notes: Table shows the estimated effects of wait time on labor market outcomes when using the shrunk IV. Health-related absence is the total number of health-related absence days (sickness absence, temporary and permanent DI) in the five years following referral, DI receipt is an indicator equal to 100 for patients receiving permanent disability insurance five years after referral, Earnings is earnings in NOK measured five years after referral, Employment is defined as an indicator variable equal to 100 for having positive earnings, UI days is the number of days the individual receives unemployment benefits The sample in column 5 is restricted to patients with positive earnings, and that in column 6 to patients not receiving DI. All regressions include year-by-referral-month and hospital-by-procedure fixed effects. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

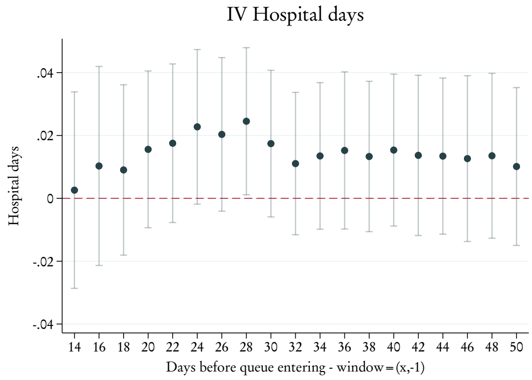
Table B9. Absence, emergency patients

	(1) Hospital days	(2) Health-related absence	(3) DI receipt
Congestion	0.0117	-0.0167	0.0121
	(0.00895)	(0.132)	(0.0286)
Dep. mean	31.12	280.7	53.79
Observations	62,723	62,723	62,723

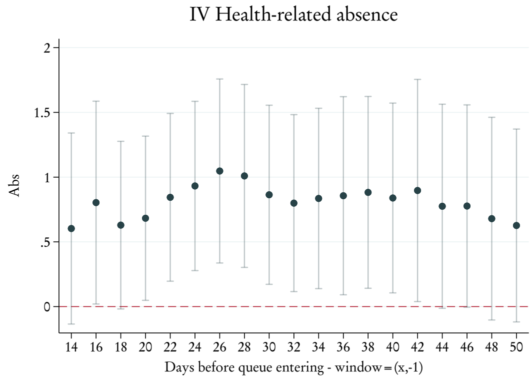
Note: Table shows models estimated on a sample of patients admitted for emergency orthopedic surgery. In these models, congestion refers to the average wait time of non-emergency patients in the hospital-by-month group. All regressions include year-by-referral-month and hospital-by-procedure fixed effects. Standard errors are clustered at the hospital-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

can be used to estimate placebo models. Specifically, we estimate regressions of five-year absence and DI on the wait times of scheduled patients, controlling for calendar time and hospital-by-procedure fixed effects. If the exclusion restriction holds, we would expect to find zero congestion effects for this group. Conversely, a positive relationship between congestion and later sickness absence would indicate that congestion influences outcomes through channels other than individual wait times, which would violate the exclusion restriction.

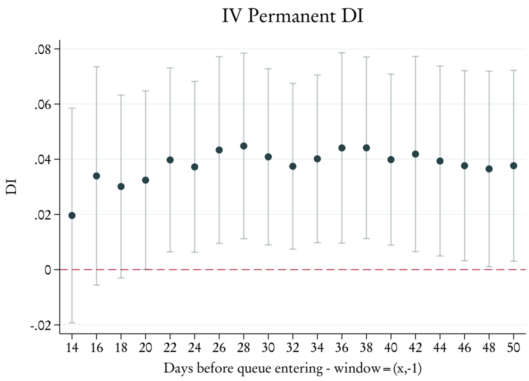
The results of this exercise are shown in Table B9. The model finds no significant congestion effects on absence or DI for patients undergoing unplanned surgeries. Moreover, the standard errors of these estimates allow us to rule out substantial increases in post-surgery absence for these patients. The 95% confidence intervals indicate that ten days additional mean wait time for planned surgeries decreases cumulative absence days by 1.7 days over the five-year period; the likelihood of 5-year DI receipt increases by no more than 0.1 percentage points. 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 hospitals).



(a) Hospital days



(b) Health-related absence days



(c) DI receipt

Figure B1. Varying window used to define the instrument.

Notes: Each point represents coefficients (95% CI) from separate IV estimations of hospital days (panel a), health-related absence (panel b), and DI (panel c) on wait time, with varying length of the pre-referral window used to define the instrument (congestion). All regressions include fixed effects for hospital-by-procedure and year-by-referral-month.

Online Appendix C: Cost-benefit analysis

Part 1: We first want to estimate the fiscal savings expected to accrue from reducing a worker's wait time by one day. Consider a representative worker, listed at the start of year 1, who waits T days for surgery instead of $T-1$ days. Using the average age (46.3) of the sample, we project estimated changes in benefit payments through year 20.

For each of the first 5 years after referral, we estimate IV models of annual benefit payments. Table C1 shows the results generated by these models.

Our data do not allow us to estimate effects beyond $y = 5$. Instead, our baseline calculations assume that the effects estimated for year 5 continue through retirement. That is, assume $\beta dy = \beta d5$ for $y = 6, \ldots, 20$.

We then take all these parts and calculate the PDV of the stream of yearly effects. Let $\phi < 1$ represent the annual discount factor. The estimated change in the PDV of benefit payments can be expressed:

$$P\hat{D}V = \sum_{y=1}^{20} \phi^y \beta_y$$

This $P\hat{D}V$ gives the PDV for the expected "fiscal savings" resulting from reducing by one day the wait time of a single representative worker.

Part 2: We use this result to infer the fiscal savings that would accrue if an additional procedure was added, using the marginal procedure to take out the patient at the back of the line. We note that this calculation yields a somewhat conservative estimate, as it completely ignoring any spillover effects on later entrants. In this exercise, we use the sample average wait time (168 days) when calculating total savings as follows:

$$P\tilde{D}V = P\hat{D}V \times 168$$

This $P\tilde{D}V$ gives the PDV for the expected "fiscal savings" arising from the insertion of one additional procedure into the system.

Part 3: While we do not have data on the cost needed to insert one additional procedure into the system, data on average costs may be suggestive. That is, we use the DRG payment levels as an approximate measure of the marginal cost hospitals incur for each procedure they perform under current capacity constraints and throughput levels. Presumably, it would cost more than this to insert an additional procedure into a hospital's operations.

Taking the average spending per procedure (NOK 33,150) as a benchmark, we construct some hypotheticals. For instance, suppose the cost of inserting an additional procedure into the system is twice that amount, because (say) the system needs to pay generous overtime wages to those contributing the extra work.

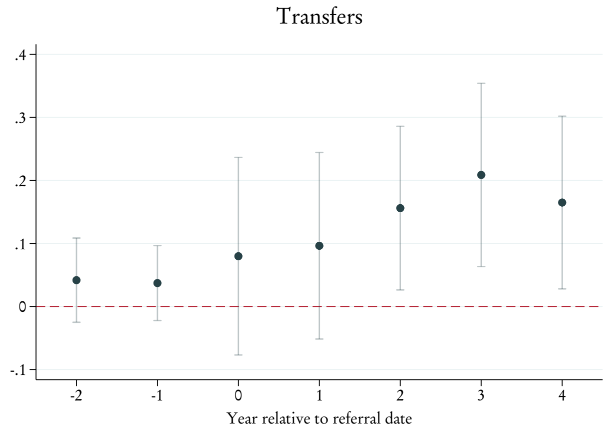


Figure C1. IV estimates for transfers.

notes: The figure plots the estimated effects of wait time on the number of hospital days and days of health-related absence relative to referral year.

Then we could compare $P\tilde{D}V$ to NOK 66,300 to determine whether inserting the additional procedure would yield a net cost reduction for the government. More generally, we can pose the policy question in the following way. How much more than its normal DRG rate should the system be willing to pay to insert an additional procedure into the system? The answer to this is given by $P\tilde{D}V/33,150$.

These calculations rely on admittedly strong assumptions.[4] For years 6-20, we assume that the effect of wait time on transfer is equal to the estimated effect in year 5. This may not hold true: effects in later years could be larger or smaller. In particular, it could be the case that longer wait times shift the timing of DI receipt forward - patients with shorter wait times may still access DI in later years, in which case the effect would diminish over time. Panel B of Table C1 illustrates how the calculations change when we assume that effects on transfers fall by 10% each year starting in year 6.

Rows (2) and (3) of Table C1 present cost-benefit calculations calculated separately by sick leave status at referral. Patients on sick leave who are assigned short wait times and do not enter DI could have increasing rates of sick pay over time. In that case, the net effect of wait time on transfers would erode over time, and our exercise would overstate the fiscal gains from shorter waits. With that caveat, our most conservative estimates imply a cost saving ratio of 34 for patients on sick leave, roughly four times the figure for the pooled sample.

[4]Moreover, we do not take account of the deadweight loss of distortionary taxes (i.e. we understate costs), though we can think of this as entering into the multiplicative factor linking marginal cost to average cost. We also do not take account of lost income tax revenue (understate benefits).

Table C1. Cost-benefit calculations

		(1) Pooled	(2) On SL	(3) Not on SL
Panel A: Baseline				
A.	PDV years 1-20	2339	9129	1186
B.	Days	168	168	168
C.	Total saving = A x B	391 707	1 529 108	198 655
D.	Cost saving ratio: C/33,150	11.82	46.13	5.99
Panel B: Convergence				
A.	PDV years 1-20	1495	6761	854
B.	Days	168	168	168
C.	Total saving = A x B	250 339	1 132 459	143 036
D.	Cost saving ratio: C/33,150	7.55	34.16	4.31

Notes: Table shows calculations of the NPV value of reducing waiting times by inserting 1 additional procedure (rows A-D) together with the cost-saving ratio defined as the largest ratio of marginal to average cost of surgery where the expected benefit exceeds the marginal cost.

Online Appendix D: Context on orthopedic surgeries and conditions leading to absences and DI receipt

In this paper, we explore effects of wait time on health and labor market outcomes by focusing on waits for orthopedic procedures. In this appendix, we provide a discussion of this choice, together with a discussion of if and how our findings might generalize. Orthopedic procedures are of particular interest because they are so tightly connected to musculoskeletal conditions, a major diagnosis group leading to long term disability. Moreover, by analyzing elective surgeries where there is potentially some slack in scheduling times, our empirical strategy is able to pin down causal effects of longer waits in the presence of endogenous wait times. Below, we discuss if and how these results may extend to other diagnoses and surgical procedures.

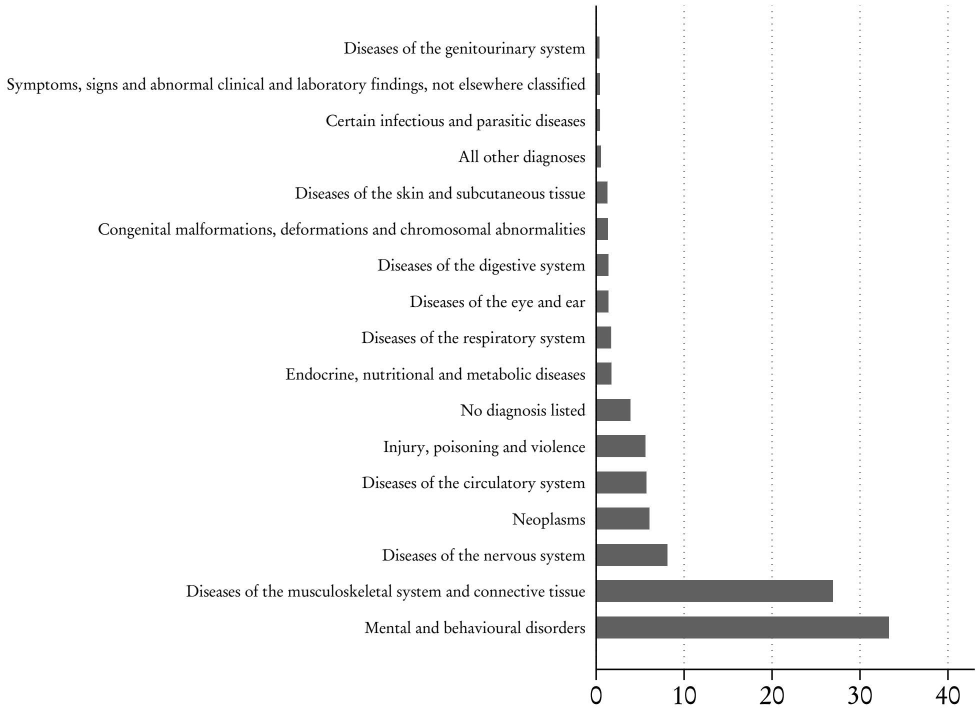
Other diagnoses

During the five-year period from 2012 to 2016, approximately 29,000 people enrolled in permanent disability insurance each year. Permanent disability benefit requires a permanent loss in earnings capacity that results from illness or injury. A wide variety of diagnoses could qualify for eligibility. Figure D1 shows the distribution of new DI entrants by diagnosis. The two most common diagnoses are mental and behavioral disorders (33%) or musculoskeletal conditions (27%), followed by diseases of the nervous system (8%, primarily MS, ME/CFS, and epilepsy), neoplasms/cancer (6%) diseases of the circulatory system (6%), and accidents/external causes (6%).

These diagnoses vary in their etiology as well as in the availability of effective treatments. In principle, we would like to know more about how timely access to treatment affects DI entry across these diagnosis categories. To illustrate, mental and behavioral disorders is the largest single diagnosis category for new DI entrants. Estimating effects of timely access to psychiatric treatment on later health and labor market outcomes is clearly policy relevant.

At the same time, our empirical strategy in this paper requires us to order events along a stylized timeline where wait times are defined relative to a single focal treatment date. While this arguably is a useful framework for surgical interventions, it is less suited for psychiatric care, where treatments are likely to be ongoing over a longer duration, involving substitution between levels of care (inpatient/outpatient/primary care). For other diagnoses (e.g. cancer), while there could be delays in diagnosis, it is generally unlikely that there will be significant delays in treatment once patients have been correctly diagnosed.

Figure D1. New DI entrants by diagnosis, 2012-2016



Notes: Numbers from the Norwegian Labour and Welfare Administration https://www.nav.no/no/nav-og-samfunn/statistikk/aap-nedsatt-arbeidsevne-og-uforetrygd-statistikk/uforetrygd/diagnoser-uforetrygd

Other surgeries

Orthopedic surgeries - surgeries related to the musculoskeletal system - are of particular interest as they address one of the leading diagnoses that lead to DI entry. However, other surgical procedures are more prevalent in the years directly preceding DI entry. Figure D2 shows the share of DI entrants who have had one or more elective surgeries, grouped by NCSP chapters, in the five year period immediately preceding DI entry. The most common type of procedure are transluminal endoscopies (NCSP chapter U), followed by minor surgical procedures (T), musculoskeletal (N - the focus of the current paper), and skin (Q).

We have implemented our preferred IV model to estimate effects of waiting for the four largest surgical procedure categories. In all these models, the estimation samples are constructed analogously to our preferred estimation sample.

Table D1 presents summary statistics for workers referred to these four most common surgical procedure groups. On average, waits are shorter for the other procedures, especially for endoscopies. Workers who undergo endoscopies and skin

Table D1. Summary statistics - workers referred to other elective surgeries

	N mean/sd	Q mean/sd	T mean/sd	U mean/sd
Wait time	190.3	149.4	135.0	104.6
	(184.3)	(190.3)	(209.6)	(147.3)
Congestion	176.6	111.0	104.4	86.3
	(52.3)	(40.6)	(46.7)	(41.1)
On sick leave at referral	0.24	0.11	0.27	0.15
Diagnosis last pre-referral consult				
Musculoskeletal	0.64	0.32	0.41	0.25
Digestive	0.034	0.060	0.055	0.20
Psychological	0.084	0.14	0.10	0.13
Skin	0.022	0.12	0.041	0.025
Cardivascular	0.031	0.041	0.054	0.049
General symptoms	0.055	0.078	0.090	0.084
Observations	26410	13541	16553	51540

Notes: summary statistics for workers referred to transluminal endoscopies (NCSP chapter U), minor surgical procedures (T), musculoskeletal (N), and skin (Q) surgical procedures. Data on diagnosis collected from ICPC-2 chapter from patients' last pre-referral GP consultation if records indicates that a sickness absence certificate was issued.

procedures are less likely to be on sick leave at referral. For patients who undergo orthopedic surgeries, a large majority (64%) have their most recent pre-referral absence certificate issued for musculoskeletal conditions. For the other surgical procedure groups, the pattern is less clear; overall, there does not appear to be a single dominant diagnosis category for either of the three groups. To illustrate, only 12% of patients undergoing skin surgery have their most recent absence certificate due to skin conditions.

To summarize, patients undergoing the other three most common surgical procedure groups are less likely to be on sick leave at referral, and the surgeries may be less likely to target the cause of sickness absence for patients who are on sick leave.

Table D2 and D3 present IV estimates of effects of wait time by surgical procedure group. For the full sample (Table D2), the instrument binds for all procedure groups except for skin surgeries, that is, the F statistic of the excluded instrument is above conventional thresholds for minor surgeries and transluminal endoscopies. Longer waits have no significant effects on absence or DI enrollment for either of these groups.

When the estimation sample is restricted to only patients on sick leave (Table

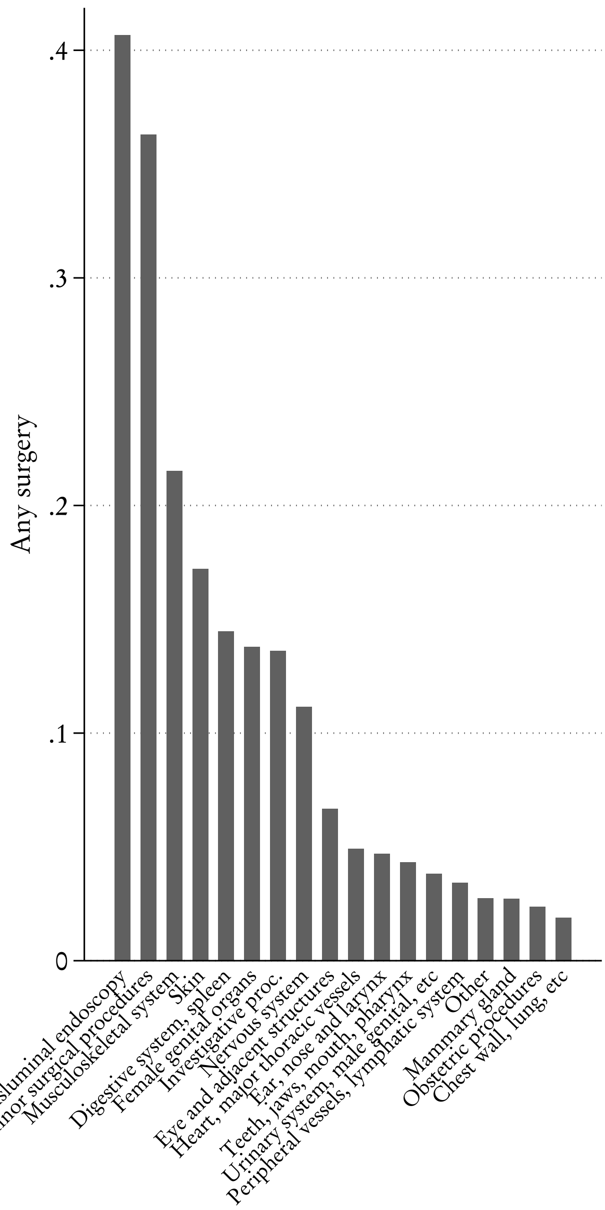
D3) , the instrument binds only for transluminal endoscopies. Again, we find no effect on DI or hospital days, however we do find a marginally significant increase in days lost to sickness absence.

Table D2. Effects of wait time by surgical procedure group - full sample

	(1) All	(2) N	(3) Q	(4) T	(5) U
First stage					
Congestion	0.508***	0.356***	0.172**	0.431***	0.894***
	(0.0806)	(0.0563)	(0.0749)	(0.102)	(0.141)
Outcome: DI receipt					
Wait time	0.00422	0.0409**	-0.0100	-0.0209	-0.000636
	(0.00578)	(0.0163)	(0.0356)	(0.0200)	(0.00527)
Dependent mean	5.142	5.381	3.493	8.047	4.521
Outcome: Health-related absence					
Wait time	0.114	0.866**	-0.407	-0.319	0.0299
	(0.131)	(0.353)	(0.767)	(0.384)	(0.124)
Dependent mean	319.5	373.1	224.0	403.9	290.0
Outcome: Hospital days					
Wait time	-0.00537	0.0174	-0.0746	-0.00588	-0.00674
	(0.00784)	(0.0119)	(0.0532)	(0.0284)	(0.00965)
Dependent mean	19.78	17.65	16.63	27.85	19.11
F-stat	39.80	39.97	5.284	17.88	40.21
N	108044	26410	13541	16553	51540

Note: Table presents IV estimates of effects of wait time on health and labor market outcomes for workers referred to transluminal endoscopies (NCSP chapter U), minor surgical procedures (T), musculoskeletal (N), and skin (Q) surgical procedures. DI receipt is an indicator equal to 100 for patients receiving permanent disability insurance five years after referral. Health-related absence is the total number of health-related absence days (sickness absence, temporary and permanent DI) in the five years following referral. Hospital days indicates the number of days in hospital, including the day of surgery. All regressions include year-by-referral-month and hospital-by-procedure fixed effects. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure D2. Surgeries in the 5-year period before DI receipt



Notes: Figure shows share of DI recipients who have one or more surgeries in each category in the five year period preceding DI entry. Categories based on 1-digit NCSP chapters.

Table D3. Effects of wait time by surgical procedure group - on sick leave at referral

	(1) All	(2) N	(3) Q	(4) T	(5) U
First stage					
Congestion	0.423***	0.345***	0.316*	0.160	0.843***
	(0.0829)	(0.111)	(0.179)	(0.155)	(0.170)
Outcome: DI					
Wait time	0.0329	0.135**	0.0137	-0.238	0.0283
	(0.0272)	(0.0656)	(0.126)	(0.271)	(0.0255)
Dependent mean	16.83	13.96	18.73	20.00	17.07
Outcome: Absence 1-5					
Wait time	1.236**	2.719**	1.981	-1.227	0.848*
	(0.490)	(1.063)	(1.784)	(2.943)	(0.446)
Dependent mean	790.5	769.5	778.0	847.6	778.4
Outcome: Hospital days					
Wait time	-0.0290	0.00575	-0.175	-0.175	-0.00653
	(0.0295)	(0.0295)	(0.190)	(0.283)	(0.0377)
Dependent mean	29.26	21.16	31.74	39.33	29.84
F stat	26.06	9.765	3.148	1.070	24.73
N	20274	6468	1431	4394	7982

Note: Table presents IV estimates of effects of wait time on health and labor market outcomes for workers referred to transluminal endoscopies (NCSP chapter U), minor surgical procedures (T), musculoskeletal (N), and skin (Q) surgical procedures. DI receipt is an indicator equal to 100 for patients receiving permanent disability insurance five years after referral. Health-related absence is the total number of health-related absence days (sickness absence, temporary and permanent DI) in the five years following referral. Hospital days indicates the number of days in hospital, including the day of surgery. All regressions include year-by-referral-month and hospital-by-procedure fixed effects. Standard errors are clustered at the hospital-by-procedure level. Stars indicate significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Online Appendix E: Mechanisms - comparing magnitudes and correlated outcomes

In Section 6 of the paper, we discuss possible mechanisms, including health, human capital depreciation, and preferences. In this appendix, we present two additional elements in this discussion. First, we use published estimates from empirical studies of work absence, productivity, job loss and DI receipt to, comparing the magnitudes of their findings to the results of our preferred specifications. Second, we present results from exercises analyzing pairwise correlation of effects on different outcomes across subgroups.

Extrapolation from the literature

The literature on human capital deterioration suggests that this channel offers at most, a partial explanation for the labor supply effects we estimate. Existing research suggests that the average annual rate at which human capital depreciates during separations from work is less than 2-6 percent (Arrazola and Hevia, 2004; Weber, 2014; Dinerstein et al., 2020), which implies that 10 days of increased workplace absence translates into an expected productivity loss of around 0.1 percent. It seems unlikely that productivity losses of this magnitude could induce a 0.4 percentage point increase in DI receipt.

Rege et al. (2009) find that worker exposure to a plant closure event raises DI receipt by 1.1 percentage points. This paper may provide an especially relevant comparison, as the population and outcome variables are defined in a manner fairly similar to our own paper.[5] Extrapolating from our IV estimates, we can infer that 30 days of additional wait time has roughly the same impact on DI receipt as when a worker's plant closes. In order for the human capital channel to explain our results, we would thus need to argue that 30 days of additional wait time has the same impact on future earnings/employment prospects as experiencing a plant closure. An effect of this magnitude seems implausible, in particular given our finding that longer wait times do not have any significant impacts on the probability of receipt of unemployment insurance.[6]

Dahl et al. (2014)'s study of intergenerational effects of DI receipt provides evidence of the importance of preferences, broadly defined, in determining DI rates. Using a judge fixed effects design, the authors find that parental DI receipt significantly increases the probability of children entering DI programs. In addition to being statistically significant, the estimated effects are also large: parental DI

[5] Models estimated on a sample of Norwegian workers, DI entry measured five years after baseline.

[6] Moreover, we note that the income effects of plant closure reported in Rege et al. (2009) are quite large. Excluding subjects who entrolled for DI, Rege et al. estimate that exposure to plant closure reduced the likelihood of full-time employment by 6.6 percentage points and reduced worker incomes by 9.3%. In contrast, our analyses found no significant income losses among workers who do not enter DI (see Table **??** column 6).

receipt raises children's DI entry by 6.1 percentage points. In our setting this would correspond to 149 days of waiting. A natural question then is whether we think that waiting an additional 20 weeks for surgery could yield similar changes in patients' willingness to enroll for DI, as a result of changes in knowledge, attitudes or preferences. In our view such shifts in preferences, broadly defined, would be more likely to occur for patients who are on sick leave at the time of referral. For this group, longer wait times are likely to increase total time spent in a work-disabled state; these patients might also be more likely to receive information about temporary or permanent DI from their physician as sick pay benefits approach exhaustion.

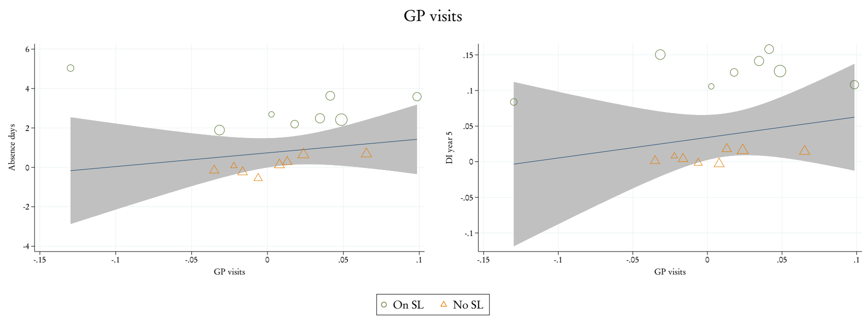
Subsample analyses across outcomes

As discussed throughout the paper, we think of health care utilization as a proxy for underlying health changes. Similarly, we think of unemployment insurance (UI) receipt as a proxy for human capital loss, broadly defined. Our analyses found no effects of wait time on either of these outcomes, on average or across subgroups. This is a first indication that effects on health and or human capital loss are unlikely mechanisms. To analyze these possible channels further, we estimate a set of models to see whether estimated effects on health care utilization and UI across subgroups correlate with effects on absence from work and DI. The intuition here is simple: if effects are driven by health, then the estimated effects of wait time on health care utilization should be correlated with the effects on absence rates. We split the sample by gender, age (above/below 45), education (college/no college) and occupational category (office/manual). For each of these categories, we then estimate effects separately according to sick leave status at referral.[7]

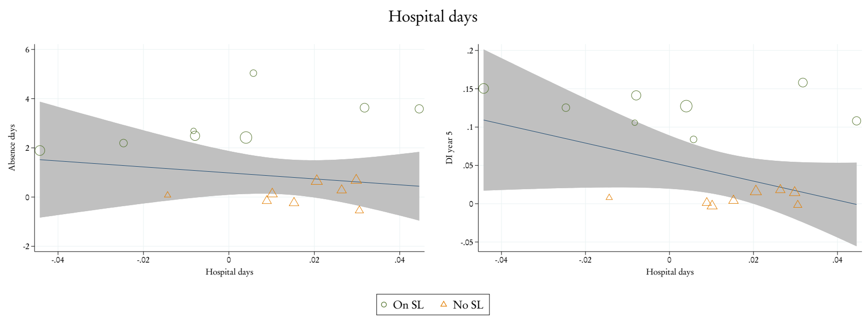
In Figure E1, we then plot the estimated effects of wait time on health care utilization and unemployment insurance against the estimated effects of wait time on total absence days and disability insurance receipt. The top panel plots the effects on the total number of GP visits on the x axis, and the effects on absence days (left panel) and DI receipt (right panel) on the y-axis. If the increase in absence and DI were driven by adverse impacts on health, we would expect these estimates to be positively correlated, i.e. the line of best fit should slope upward. The figures give no indication that this is the case. The effects on absence days are close to zero for all the subsamples that were not on sick leave at referral, while estimated effects are positive for patients on sick leave at referral across subsamples.[8] As

[7]I.e., we run a total of 16 regressions for each of the following outcomes: total number of GP visits, total number of hospital days, total number of days received unemployment insurance, total number of health-related absence days and a dummy for DI receipt in year 5.

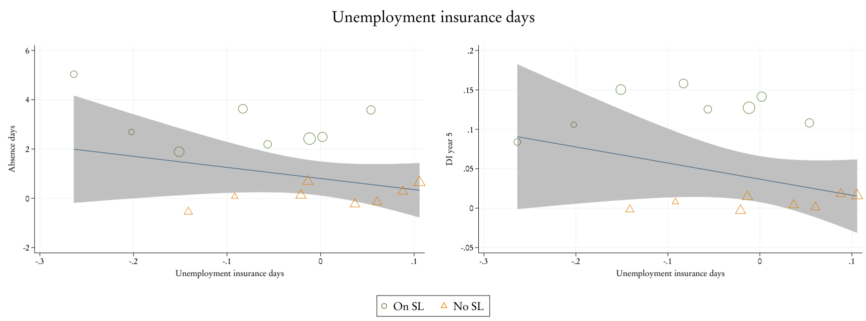
[8]The fact that the positive effects of wait time on absence and DI for patients on sick leave are present across subgroups gives a further indication that the differential effect by sick leave at referral is not driven by differences in observable characteristics.



(a) Effects on health-related absence and DI receipt vs. effects on GP visits



(b) Effects on health-related absence and DI receipt vs. effects on hospital days



(c) Effects on health-related absence and DI receipt vs. effects on unemployment insurance

Figure E1. Correlated effects across outcomes.

Notes: Each point represents one pair of IV regressions estimated on a subsample defined by sick leave status at referral and each of the following four binary categories: gender, age (above/below 45), education (college/no college) and occupation (manual/office). All regressions include fixed effects for hospital-by-procedure and year-by-referral-month.

before, the effects on GP visits are scattered around zero. Crucially, there is no correlation between the two. That is, there is no systematic pattern where groups that happen to have large estimated increases in GP visits have large increases in absence days and DI receipt; the slope of the line of best fit is not significantly different from zero. The middle panel of Figure E1 shows the corresponding analysis for total number of hospital days. Again, we find a similar pattern: while the estimated effects of wait time on hospital days are scattered around zero, we find no indication that these coefficients are positively correlated with the estimated effects on absence rates or DI.

The bottom panel correlates the effects of UI receipt with effects on DI and absence days. Rates of human capital depreciation likely vary across occupations and across workers of different skill levels and types. For instance, Görlich and De Grip (2008) find evidence that higher-skilled women self-select into occupations for which the wage penalties due to career interruptions are smaller, while Weber (2014) finds evidence of higher depreciation rates among less educated workers.[9] It is possible that the subsample of workers receiving disability insurance benefits in the aftermath of longer waits are those for whom the human capital depreciation implications of work interruptions are especially severe. Our finding that the labor supply implications of waiting are concentrated among lower-educated workers is broadly consistent with Weber (2014) in that respect. However, we fail to find evidence that less educated workers have a correspondingly high risk of involuntary job loss (as measured by UI receipt). More generally, we find no indication that the subsamples with increases in UI receipt have larger labor supply responses; the line of best fit is not statistically different from zero.

[9]In contrast to the Weber (2014) result, Edin and Gustavsson (2008) find no evidence of differential skill deterioration among less educated workers when analyzing an explicit measure of literacy skills.

References

Arrazola, M. and J. d. Hevia (2004). More on the estimation of the human capital depreciation rate. *Applied Economics Letters 11*(3), 145–148. (document)

Ballou, D., W. Sanders, and P. Wright (2004). Controlling for student background in value-added assessment of teachers. *Journal of educational and behavioral statistics 29*(1), 37–65. (document)

Barrett, G. F. and S. G. Donald (2003). Consistent tests for stochastic dominance. *Econometrica 71*(1), 71–104. A2

Bhuller, M., G. B. Dahl, K. V. Løken, and M. Mogstad (2020). Incarceration, recidivism, and employment. *Journal of Political Economy 128*(4), 1269–1324. *

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014). Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American Economic Review 104*(9), 2593–2632. 2

Corcoran, S. P., J. L. Jennings, and A. A. Beveridge (2011). Teacher effectiveness on high-and low-stakes tests. *Society for Research on Educational Effectiveness*. 2

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

Dinerstein, M., R. Megalokonomou, and C. Yannelis (2020). Human capital depreciation. Working paper, National Bureau of Economic Research. (document)

Edin, P.-A. and M. Gustavsson (2008). Time out of work and skill depreciation. *ILR Review 61*(2), 163–180. 9

Görlich, D. and A. De Grip (2008). Human capital depreciation during hometime. *Oxford Economic Papers 61*(suppl_1), i98–i121. (document)

Guarino, C. M., M. Maxfield, M. D. Reckase, P. N. Thompson, and J. M. Wooldridge (2015). An evaluation of empirical bayes's estimation of value-added teacher performance measures. *Journal of Educational and Behavioral Statistics 40*(2), 190–222. (document)

Jacob, B. A. and L. Lefgren (2008). Can principals identify effective teachers? evidence on subjective performance evaluation in education. *Journal of labor Economics 26*(1), 101–136. 2

Kane, T. J. and D. O. Staiger (2008). Estimating teacher impacts on student achievement: An experimental evaluation. nber working paper no. 14607. *National Bureau of Economic Research*. 2

Kline, P. and M. Tartari (2016). Bounding the labor supply responses to a randomized welfare experiment: A revealed preference approach. *American Economic Review 106*(4), 972–1014. A2

Lee, D. S., J. McCrary, M. J. Moreira, and J. R. Porter (2021). Valid t-ratio inference for iv. Technical report, National Bureau of Economic Research. (document)

McCaffrey, D. F., J. Lockwood, D. Koretz, T. A. Louis, and L. Hamilton (2004). Models for value-added modeling of teacher effects. *Journal of educational and behavioral statistics 29*(1), 67–101. 2

Rege, M., K. Telle, and M. Votruba (2009). The effect of plant downsizing on disability pension utilization. *7*(4), 754–785. _eprint: https://academic.oup.com/jeea/article-pdf/7/4/754/10313311/jeea0754.pdf. (document)

Weber, S. (2014). Human capital depreciation and education level. *International Journal of Manpower 35*(5), 613–642. (document), 9