

This is a repository copy of *The impact of extending nurse working hours on staff sickness absence:evidence from a large mental health hospital in England*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/159867/>

Version: Published Version

Article:

Rodriguez Santana, Idaira orcid.org/0000-0003-2022-3239, Anaya Montes, Misael, Chalkley, Martin John orcid.org/0000-0002-1091-8259 et al. (3 more authors) (2020) The impact of extending nurse working hours on staff sickness absence:evidence from a large mental health hospital in England. *International Journal of Nursing Studies*. 103611. ISSN 0020-7489

<https://doi.org/10.1016/j.ijnurstu.2020.103611>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

This is a repository copy of *The impact of extending nurse working hours on staff sickness absence : evidence from a large mental health hospital in England*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/159867/>

Version: Accepted Version

Article:

Rodriguez Santana, Idaira orcid.org/0000-0003-2022-3239, Anaya Montes, Misael, Chalkley, Martin John orcid.org/0000-0002-1091-8259 et al. (3 more authors) (2020) The impact of extending nurse working hours on staff sickness absence : evidence from a large mental health hospital in England. *International Journal of Nursing Studies*. ISSN 0020-7489

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Elsevier Editorial System(tm) for
International Journal of Nursing Studies
Manuscript Draft

Manuscript Number: IJNS-D-19-01798R2

Title: The impact of extending nurse working hours on staff sickness absence: Evidence from a large mental health hospital in England

Article Type: SI: Shift work in nursing

Keywords: Key Words: shift patterns, mental health providers, nurses, sickness absence, health workforce, England

Corresponding Author: Dr. Idaira Rodriguez Santana, Ph.D

Corresponding Author's Institution: Centre for Health Economics

First Author: Idaira Rodriguez Santana, Ph.D

Order of Authors: Idaira Rodriguez Santana, Ph.D; Misael Anaya Montes, Ph.D.; Martin J Chalkley, Professor; Rowena Jacobs, Professor; Tina H Kowalski, Ph.D; Jane E Suter, Ph.D

Manuscript Region of Origin: UNITED KINGDOM

Abstract: Background: A pressing international concern is the issue of mental health workforce capacity, which is also of concern in England where staff attrition rates are significantly higher than in physical health services. Increasing demand for mental health services has led to severe financial pressures resulting in staff shortages, increased workloads, and work-related stress, with health care providers testing new models of care to reduce cost. Previous evidence suggests shift work can negatively affect health and wellbeing (increased accidents, fatigue, absenteeism) but can be perceived as beneficial by both employers and employees (fewer handovers, less overtime, cost savings).

Objective: This study reports an evaluation of the impact of extending the shifts of nurses and health care assistants from 8 to 12 hours. Using data before and after the policy change, the effect of extended working hours on short term sickness (< 7 days) on staff is examined.

Setting: The setting is six inpatient wards within a large mental health hospital in England where the shift extension took place between June and October 2017. The Data come from wards administrative records and the analysis is performed using weekly data (N=463).

Methods: Causal inference methods (Interrupted Time Series and Difference-in-Difference) are used to compare staff sickness rates before and after the implementation, where the outcome variable is defined as the ratio of total sickness hours over the total scheduled working hours (FTE) in a given week. Patient casemix, staff demographics, ward and time variables are included as controls.

Results: Estimation results establish that the extended shifts are associated with an increased percentage of sickness hours per week of between 0.73% and 0.98%, the equivalent of a complete shift per week per ward.

Conclusion: This is the first study to use causal inference to measure the impact of longer shifts on sickness absences for mental health

workforce. The analysis is relevant to other providers which may increasingly look towards these shift patterns as a means of cost saving.

Dear Prof. Griffiths,

Many thanks for the acceptance of our manuscript: "The impact of extending nurse working hours on staff sickness absence: Evidence from a large mental health hospital in England" for publication. We appreciate the careful review and constructive suggestions from the two reviewers. It is our belief that the manuscript is substantially improved after making the suggested edits.

As requested, we have removed all the abbreviations. Please let us know if there are any other outstanding issues with the manuscript.

Yours sincerely,

Idaira Rodríguez Santana

The impact of extending nurse working hours on staff sickness absence: Evidence from a large mental health hospital in England

Idaira Rodriguez Santana^{1,*}, Misael Anaya Montes¹, Martin Chalkley¹, Rowena Jacobs¹, Tina Kowalski² and Jane Suter²

¹ Centre for Health Economics, University of York, Heslington, York, YO10 5DD, United Kingdom

² The York Management School, University of York, Freboys Lane, Heslington, York, YO10 5GD, United Kingdom

*Corresponding author: Idaira.rodriquezsantana@york.ac.uk; Tel: +44 (0)1904 321408

Twitter Handles: @Idalo_Ip (Idaira Rodriguez Santana); @Jacobs__R (Rowena Jacobs); @CHEyork (Centre for Health Economics); @UoY_TYMS (The York Management School)

IJNS AUTHOR CHECKLIST

You will need to submit a completed version this checklist plus the checklist from the any relevant reporting guideline along with your paper. This checklist addresses a number of important issues. It is intended to help you to make sure your manuscript meets some basic requirements. It should be read in conjunction with the guide for authors, and is not a replacement for it. We have prepared a template that may help you to structure your paper (see guide for authors)

PART 1 Basic requirements - For the items below, please tick or the relevant page number in the right hand column to confirm you have included/addressed the items in your manuscript. For more detail please consult the guide for authors:		Insert a tick or page number(s)
Ethical approval and informed consent	For all research papers <i>only</i> , please ensure that your manuscript includes details of the ethical approval granted including the body that granted it and any reference number. If ethical approval was not required, give a clear statement of the basis on which this assessment was made, with reference to the ICMJE requirements. This should include confirmation of informed consent by participants. Place this at the end of you methods section .	N/A Check pg 21
Study registration	Give any study registration number (e.g. ISRCTN) in the abstract and in the body of the paper. For clinical trials (as defined by the ICMJE), the abstract should include the registration date and the date of first recruitment. [not applicable to letters / editorials]	N/A
Funding sources	State sources of funding and the role of funders in the conduct of the research or include a statement 'no external funding' at the end of the paper .	Pg 21
Conflict of interests	State any actual or potential conflicts of interest in a section at the end of the paper . If there are none, include a statement "Conflicts of interest: none". The substance of this declaration should match details provided in file(s) uploaded at submission.	Pg 21
Title	The title is in the format 'Topic / question: design/type of paper' [not applicable to letters / editorials]	Pg 1 (Title Page)
Abstract	A structured abstract of no more than 400 words appropriate to the design of the study (and as directed by relevant reporting guidelines) is included at the beginning of your paper. No references are cited in the abstract. [not applicable to letters / editorials]	Pg 1
	You may include a final section to their structured abstract with an additional final section: "Tweetable abstract" summarising a key message in no more than 140 characters. [not applicable to letters / editorials]	
	No abbreviations (other than SI units) or references are to be used in the title or the abstract of the paper	
Key words	Give between four and ten key words, which accurately identify the paper's subject, purpose, method and focus. Use the Medical Subject Headings (MeSH®) thesaurus or Cumulative Index to Nursing and Allied Health (CINAHL) headings where possible (see http://www.nlm.nih.gov/mesh/meshhome.html).	Pg 1
Contribution of the Paper statements	After the abstract under the headings "What is already known about the topic?" and "What this paper adds" give 2-3 single sentence bullet points (each) summarising key contributions. [not applicable to letters / editorials]	Pg 2
Abbreviations	The paper does not contain any abbreviations, acronyms or "initialisms" other than the limited exceptions noted in the guide for authors.	
Other Published accounts	Other published and in press accounts of the study from which data in this paper originate are referred to in the paper and the relationship between this and other publications from the same study is made clear in the paper. [not applicable to editorials or letters unless reporting analysis / data]	

Please provide below full references to ALL other publications from this study and explain the relationship to the current paper. To assist editors upload copies of papers where the abstract / full text is not readily available (including those under review elsewhere, which will be treated in strict confidence).

PART 2 Standards of reporting	The editors require that manuscripts adhere to recognized reporting guidelines relevant to the research design used. Guidelines endorsed by the IJNS are listed below. These and others can be found at http://www.equator-network.org/ . As a separate file, we require you to submit a completed checklist detailing how and where the matters detailed in the guideline are addressed in your paper. Do NOT submit the guideline itself. Indicate below what guideline you have used. [please note and use the appropriate extensions – eg. CONSORT extension for cluster trials]	Checklist submitted**
Randomised (and quasi-randomised) controlled trial	CONSORT – Consolidated Standards of Reporting Trials	
Qualitative studies	COREQ: Consolidated criteria for reporting qualitative research	
Systematic Review of Controlled Trials	PRISMA - Preferred Reporting Items for Systematic Reviews and Meta-Analyses	
Study of Diagnostic accuracy / assessment scale	STARD Standards for the Reporting of Diagnostic Accuracy studies	
Observational cohort, case control and cross sectional studies	STROBE S trengthening the R eporting of O bservational Studies in E pidemiology	x
Quasi experimental / non-randomized evaluations	TREND - Transparent Reporting of Evaluations with Non-randomized Designs	
Other (please name / give source)		
Not applicable (please elaborate)	<i>If there is no applicable guideline, upload a blank file with the words 'not applicable' when requested at submission.</i>	

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	Page No
Title and abstract	1	(a) Indicate the study’s design with a commonly used term in the title or the abstract	1
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	1
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	2,3
Objectives	3	State specific objectives, including any prespecified hypotheses	4,5
Methods			
Study design	4	Present key elements of study design early in the paper	4,5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6,7
Participants	6	(a) Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up Case-control study—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls Cross-sectional study—Give the eligibility criteria, and the sources and methods of selection of participants	6
		(b) Cohort study—For matched studies, give matching criteria and number of exposed and unexposed Case-control study—For matched studies, give matching criteria and the number of controls per case	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7,8
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7,8
Bias	9	Describe any efforts to address potential sources of bias	7,8
Study size	10	Explain how the study size was arrived at	6,7
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	6,7
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	8,9,10
		(b) Describe any methods used to examine subgroups and interactions	8,9,10
		(c) Explain how missing data were addressed	N/A
		(d) Cohort study—If applicable, explain how loss to follow-up was addressed Case-control study—If applicable, explain how matching of cases and controls was addressed Cross-sectional study—If applicable, describe analytical methods taking account of sampling strategy	N/A
		(e) Describe any sensitivity analyses	10,11

Continued on next page

Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	6,7
		(b) Give reasons for non-participation at each stage	6,7
		(c) Consider use of a flow diagram	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	12,13,14,15
		(b) Indicate number of participants with missing data for each variable of interest	N/A
		(c) Cohort study—Summarise follow-up time (eg, average and total amount)	N/A
Outcome data	15*	Cohort study—Report numbers of outcome events or summary measures over time	N/A
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	15,17
		Cross-sectional study—Report numbers of outcome events or summary measures	15,17
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	8,16,17
		(b) Report category boundaries when continuous variables were categorized	N/A
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	17/18
Discussion			
Key results	18	Summarise key results with reference to study objectives	18,19
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	19,20
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	20
Generalisability	21	Discuss the generalisability (external validity) of the study results	20,21
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	21

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at www.strobe-statement.org.

1 Abstract

2 **Background:** A pressing international concern is the issue of mental health workforce capacity,
3 which is also of concern in England where staff attrition rates are significantly higher than in physical
4 health services. Increasing demand for mental health services has led to severe financial pressures
5 resulting in staff shortages, increased workloads, and work-related stress, with health care providers
6 testing new models of care to reduce cost. Previous evidence suggests shift work can negatively affect
7 health and wellbeing (increased accidents, fatigue, absenteeism) but can be perceived as beneficial by
8 both employers and employees (fewer handovers, less overtime, cost savings).

9 **Objective:** This study reports an evaluation of the impact of extending the shifts of nurses and health
10 care assistants from 8 to 12 hours. Using data before and after the policy change, the effect of
11 extended working hours on short term sickness (< 7 days) on staff is examined.

12 **Setting:** The setting is six inpatient wards within a large mental health hospital in England where the
13 shift extension took place between June and October 2017. The Data come from wards administrative
14 records and the analysis is performed using weekly data (N=463).

15 **Methods:** Causal inference methods (Interrupted Time Series and Difference-in-Difference) are used
16 to compare staff sickness rates before and after the implementation, where the outcome variable is
17 defined as the ratio of total sickness hours over the total scheduled working hours (full time
18 equivalents) in a given week. Patient casemix, staff demographics, ward and time variables are
19 included as controls.

20 **Results:** Estimation results establish that the extended shifts are associated with an increased
21 percentage of sickness hours per week of between 0.73% and 0.98%, the equivalent of a complete
22 shift per week per ward.

23 **Conclusion:** This is the first study to use causal inference to measure the impact of longer shifts on
24 sickness absences for mental health workforce. The analysis is relevant to other providers which may
25 increasingly look towards these shift patterns as a means of cost saving.

26 **Key Words: shift patterns, mental health providers, nurses, sickness absence, health workforce,**
27 **England**

28 Contribution of the paper

29 What is already known about the topic?

- 30 • There is a growing concern worldwide around mental health workforce capacity,
31 characterized by staff shortages and high attrition rates and it **has led providers to test new**
32 **shift patterns** to reduce costs.
- 33 • There is uncertainty regarding the benefits of the 12-hour shift system with perceived
34 efficiencies from an employer perspective, which may be offset by adverse effects on
35 employees' health and wellbeing.

36 What this paper adds?

- 37 • It is the first to analyse the impact of 12-hour shifts on sickness absences for mental health
38 wards in England.
- 39 • Relative to 8-hours, 12-hour shifts increase sickness absence between 0.73% and 0.98%
- 40 • The methodology employed allows identifying the effect of the 12-hour shifts on sickness
41 absences from other factors, and the data used overcomes limitations of self-selection from
42 previous studies where the policy has been optional.

43 1. Introduction

44
45 The mental health workforce faces all sorts of stressors; many of those are common to all health care
46 staff, such as limited resources or overcrowded wards, whilst others are specific to the mental health
47 care setting. Examples of the latter are: dealing with patients' physical and/or verbal violent behaviour
48 (Holloway et al., 2000, Renwick et al., 2016), the use of coercive measures such as restraint and
49 detention of patients (Bonner et al., 2002), continuous monitoring of patients at risk of self-harm
50 (Hagen et al., 2017) and in the most extreme cases with patients' suicide (Fothergill et al., 2004) and
51 with the public inquiries associated with these deaths, that often allocate blame to staff members
52 (Holloway, Szmukler et al. 2000). According to O'Connor et al. (2018), the mental health workforce
53 reports higher emotional exhaustion than emergency nurses and equal burnout to cancer professionals.
54 The described stressors can explain, at least partially, the difficulties in recruiting and retaining mental
55 health workforce. According to the World Health Organization (2007), there is a growing concern
56 worldwide around mental health workforce capacity, characterized by staff shortages and high
57 attrition rates. With excessive stress stemming from intense patient interactions over extended periods

1
2 58 of time exposure to this working environment over a longer shift may have a greater impact on
3
4
5 59 employee wellbeing (Edwards and Burnard, 2003).
6

7
8 60 Historically in England, mental health services have been chronically underfunded and, despite
9
10 61 commitments to increase funding, providers are currently under financial constraints (British
11
12 62 Medical Association, 2018). The latter, in addition to the described workforce pressures, have led
13
14 63 mental health providers to test new models of delivering care to reduce costs (Kings Fund, 2015)
15
16 64 (NHS Evidence, 2010). Increasing nurses' shift length from 8 to 12 hours is becoming common
17
18 65 practice in North America and the United Kingdom to address staff and financial pressures (Harris et
19
20 66 al., 2015). The most recent estimates from Ball et al. (2015), based on Royal College of Nurses survey
21
22 67 data, shows a substantial increase in the proportion of NHS nurses working 12-hour shifts; 31% in
23
24 68 2005 compared with 52% in 2009. The latter figure contrasts with the average for 12 European
25
26 69 countries, also based on survey data, where only 14% of acute nurses were working 12 or more
27
28 70 hours per shift (Griffiths et al., 2014).
29
30

31
32 71 There is uncertainty regarding the benefits of the 12-hour shift system with perceived efficiencies
33
34 72 from an employer perspective, such as fewer handovers and less overlap between shifts (Dall' Ora et
35
36 73 al., 2015); handovers and overlaps are regarded as unproductive and associated with a larger
37
38 74 number of errors due to discontinuity of care (NHS Evidence, 2010). There are potential benefits to
39
40 75 employees such as less travel time and longer periods between shifts (Knauth, 2007). However,
41
42 76 these benefits may be potentially offset by stress and burnout (Stimpfel et al., 2012, Wallace et al.,
43
44 77 2009), loss of job satisfaction, adverse effects on health and wellbeing, absenteeism and intention to
45
46 78 leave (Dall' Ora et al., 2015, Dall'Ora et al., 2019).
47
48
49
50

51
52 79 There may also be adverse implications for the quality and safety of patient care associated with 12-
53
54 80 hour shifts. Longer shifts have been associated with increased risk of making an error, e.g. not
55
56 81 washing hands, not checking identity bracelets (Chudleigh et al., 2005, Ilhan et al., 2006) and
57
58 82 decreased quality of care (Todd et al., 1989, Vik and MacKay, 1982). However, other studies have
59
60
61
62

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

83 not found significant differences in terms of quality of care or number of incidents (Bloodworth et
84 al., 2001, Stone et al., 2006). An analysis of the European RN4CAST survey, a multi-country cross-
85 sectional nurse workforce study, shows that working 12-hour shifts is associated with higher odds of
86 poor quality of care, greater risk of necessary nurse care left undone and higher odds of reporting
87 being dissatisfied with their jobs (Griffiths et al., 2014). Nurses from the RN4CAST survey also
88 reported that longer shifts reduced the opportunities to discuss patient care and to participate in
89 continuing education activities (Dall’Ora et al., 2020).

90 There is a positive association between the amount of working hours and impairments to employees
91 physical and mental health (Raediker, 2006). Previous studies, using cross-sectional and subjective
92 nurse-reported data, have explored the effect of longer shifts on nurse health outcomes. Past
93 findings show that nurses working longer shifts reported increased fatigue (Chen et al., 2014, Geiger-
94 Brown et al., 2012, Lea and Bloodworth, 2002), greater stress burnout (Hoffman and Scott, 2003,
95 Stimpfel et al., 2012) and higher risk of suffering work-related musculoskeletal disorders (Lipscomb
96 et al., 2002, Trinkoff et al., 2006) all of which are likely to lead to increased sickness absence.
97 Sickness absence is a widely used outcome in health occupational research as it relates to workers’
98 social, economic and psychological processes (Merkus et al., 2012). Moreover, sickness absences are
99 associated with several negative outcomes such as salary loss to the employee, salary costs of
100 replacement staff, productivity losses and reduced quality of services.

101 This paper contributes to the literature by reporting an evaluation of the impact of a change from 8
102 to 12-hour shift patterns on sickness absences in a large National Health Service (NHS) mental health
103 care provider in England by means of two causal inference methods: Interrupted Time Series and
104 Difference-in-Difference. The aim is to identify whether there is a statistically significant association
105 between longer shifts and sickness absence by comparing staff short-term sickness absence rates
106 before and after the policy implementation. This is the second study, after Dall’Ora et al. (2019),
107 that explores the association between longer shifts and sickness absences. These authors analyse

108 longitudinal staffing data for 32 inpatient wards in England and find a significant association
109 between longer shifts and increased sickness absences, however their analysis does not cover
110 mental health staff. As nursing mental health patients can be physically and emotionally more
111 demanding than nursing patients with physical health problems, the extension to 12-hour shifts in
112 mental health wards might lead to worse staff outcomes than those observed for general inpatient
113 wards.

114 2. Methods

115 2.1. Data

116
117 The subjects of this study are nurses and health care assistants from six geographically dispersed
118 inpatient mental health wards from a large mental health NHS provider in England. The 12-hour
119 shifts were introduced on three different dates: 1) Adult Mental Health Wards (wards A, B and C) on
120 the 21st June 2017; 2) Older People Services Wards (wards D and E) on 18 September 2017 and 3) a
121 Learning Disability Ward (ward F) on 16 October 2017. The data are a combination of roster records
122 and routine hospital administrative data available before and after the introduction of the 12-hour
123 shifts. The ward-level data includes information on daily staff sickness records, i.e. there is a record
124 for each sickness absence registered, with each record containing information on the start date of
125 the absence, the total number of sickness hours, the ward, and staff category. General ward
126 information was available such as the total full time equivalent contracted staff per ward/month,
127 information on patients' casemix and staff demographics per ward/month. Daily data were
128 converted to weekly data. There were a total of 463 observations, one per week and ward, and this
129 data is distributed as follows: N = 69 for ward A (Observations from Oct-16 to Jan-18); N=73 for ward
130 B (Oct-16 to Jan-18); N =87 for ward C (from Jun-16 to Feb-18); N=56 for ward D (from Feb-17 to
131 Feb-18); N=91 for ward E (from Jun-16 to Feb-18) and N=87 for ward F (from Jun-16 to Jan-18).

2.2. Dependent variable: sickness absences

The duration of sickness absences ranges from 0 (where a sickness absence is reported but the staff is not on duty) to a maximum of 1912 hours (this is equivalent to a total of 239 8-hour shifts, and represents a long term sickness absence). As in Dall'Ora et al. (2019), a cut-off of 7 days is used to distinguish short and long term sickness absences. In this paper, given the limited time frame and number of observations after the implementation of the 12-hour shifts, the focus is on short-term sickness absences. This focus results in dropping 30% of individual observations for sickness absences before aggregating the sample to the ward / week level.

The outcome variable is defined as the ratio of the total sickness hours per ward/week and the total number of contracted hours per ward/week. The outcome measure is expressed as the percentage of total sickness hours per ward/week.

2.3. Control variables

Additional covariates are included to control for the effect that other factors can have on the outcome variable such as variations in patients' severity and complexity of mental ill-health, and staff demographic characteristics, all measured at the week/ward level. The following control variables are included: the monthly average Health of the Nation Outcome Scale (HoNOS) score, a clinician-rated outcome measure capturing patient's severity at ward level. The Health of the Nation Outcome Scale score is a widely used instrument measuring symptom severity and social functioning and it ranges from 0 (best) to 48 (worst) (Jacobs, 2009). Mental health providers are required to categorise all patients into groups designed to capture similar levels of need, called clusters, which in turn are categorised into three superclusters (Moscelli et al., 2018) namely: 1) non-psychotic, 2) psychotic and 3) organic. In our sample, the majority of organic patients are in wards with only other organic patients, so in order to capture these groups, but avoid multicollinearity, we only include a

157 dummy variable indicating the percentage of psychotic (category 2) patients per ward and per
158 month.

159 The inclusion of control variables improves the estimation in the presence of time-varying
160 confounders. Month dummies are included to capture variations in sickness absences that might be
161 associated with seasonal factors (e.g. flu season). In order to account for workload pressures, a
162 variable indicating the volume of patients is included, measured as the number of patients per
163 month per ward. The total number of staff (headcount) per month is included as workload has been
164 identified as a workplace stressor (Wallace et al., 2009). The percentage of registered nurses
165 controls for different capacities of staff available per week due to qualification level. We control for
166 various staff characteristics: average staff age, percentage of female staff and percentage of white
167 staff as previous literature indicates that younger mental health staff suffer more from burnout,
168 with mixed results by gender (Moore and Cooper, 1996), and that ethnic minority mental health
169 staff can be exposed to additional stressful situations linked to cultural misunderstandings with
170 patients, or to racist behaviours (Kavanagh, 1991).

171 2.4. Econometric strategy

172
173 Two different estimation approaches are followed: Interrupted Time Series and Difference-in-
174 Differences. Both methods have been used extensively in observational studies to address causality
175 (Angrist and Pischke, 2009, Cameron and Trivedi, 2005, Penfold and Zhang, 2013). For both
176 estimation methods three model specifications are estimated: (1) that only includes the policy effect
177 parameter and time dummies; (2) which includes all the variables in (1) plus patient casemix and (3)
178 which includes all the variables in (2) plus staff demographic variables.

179 Interrupted Time Series

180

181 The Interrupted Time Series analysis uses data gathered before the implementation of the 12-hour
182 shifts to forecast the trajectory of sickness absences, had the shift pattern change not come into
183 practice, and compares it to the observed sickness absences to estimate the effect associated with
184 the longer shifts. In the Interrupted Time Series analysis, the time of the implementation of the
185 policy is standardised to zero, as if the policy change was implemented in all wards at the same time,
186 and then all the data points observed before and after the policy are used to estimate an overall
187 effect of the policy in all six wards.

188 In particular, the interest is in estimating the effect of the switch from 8 to 12-hour shifts on the
189 percentage of sickness absences per week, so the equation of interest for wards $j = 1, \dots, J$ is estimated
190 by the following specification:

$$y_{jt} = \tau_{ITS}D_j + \delta D_j t + \alpha_0 t + \alpha_1 S + X'_{jt}\beta + \gamma_j + \epsilon_{jt} \quad (1)$$

193 where y_{jt} represents the percentage of sickness absences in ward j and week t ; D_j is a dummy
194 variable that captures the switch to 12-hour shifts and takes the value 1 for those observations after
195 the policy implementation and 0 for those observed before. $D_j t$ is the policy dummy interacted with
196 trend, t captures the weekly trend, and a set of dummy variables S , captures monthly seasonality.
197 X_{jt} is a vector that represents patients' characteristics and staff demographics per ward, γ_j is the
198 ward fixed effect and ϵ_{jt} the error term. The effect of the policy in this setting is measured by τ_{ITS}
199 (Cook et al., 2002).

201 The use of covariates in the interrupted time series requires that all covariates considered in X_{jt}
202 should be independent of the policy, in other words they should be stable and not affected by the
203 introduction of the policy (Lee and Lemieux, 2010). This assumption for each variable is tested in a
204 regression discontinuity framework (Calonico et al., 2019) and a failure of the assumption would
205 invalidate the use of these covariates in the regression.

206

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Differences in Differences

208
209 The Difference-in-Differences approach takes advantage of the different timings of introducing the
210 policy, in this case the period after Adult Mental Health Wards (wards A, B and C) had experienced
211 the introduction of 12-hour shifts and we use the data of the other three wards that were still operating
212 under the 8-hour shift policy (wards D, E and F) as controls. Between June and September 2017, there
213 were three wards affected by the policy and three unaffected which allows us to utilise data from the
214 latter group to estimate what would have happened to the sickness absences of affected wards if the
215 policy had not been implemented. The Difference-in-Differences approach is specified as:

$$y_{jt} = \theta Treated_j + \lambda time + \tau_{DID} Treated_j time + \alpha_1 S + X'_{jt} \beta + \gamma_j + \epsilon_{jt} \quad (2)$$

218 where *Treated* takes the value of 1 for the wards where the policy was implemented first, in this
219 case wards A, B and C and zero for wards D, E and F. The variable *time* corresponds to zero before
220 the introduction of the policy in the treated wards, in this case 21 June 2017, and one after this.

221 Following the standard Difference-in-Differences approach, in this equation τ_{DID} is the difference-
222 in-difference estimation which corresponds to the interaction term between treated and time
(Cameron and Trivedi, 2005). The sample was restricted to 18 September 2018 where the other
223 three wards (D, E and F) served as controls, because wards D and E were also affected by the policy
224 at that time.

226 The option of applying the Difference-in-Differences estimation to the period between the
227 introduction of the policy in wards D, E and ward F was explored. However, this was not possible due
228 to data limitations as there is only one month of difference between the implementation of the
229 policy on these wards.

Robustness Checks

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

232 To test the robustness of the estimates, a series of checks were carried out: (i) A placebo test is used
233 that assumes the introduction of an arbitrary extended shift policy on the same day and month, one
234 year before. (ii) The regressions are fitted using pooled Ordinary Least Squared instead of fixed
235 effect panel data and year and week dummies are added. (iii) The effect of the 12-hour shifts on
236 alternative outcome variables are tested: longer sickness absences up to 14 days, and shorter
237 absences up to 2 days. (iv) The policy implementation date is artificially moved up to six weeks prior
238 to the introduction of the policy and up to six weeks after to test for anticipatory and delayed
239 effects.

240 2.5 Ethical approval

241 This study received ethics approval from the Health Research Authority - 18-HRA-0454.

242 3. Results

243 3.1. Descriptive statistics

244
245 Figure 1 shows the change in the percentage of sickness absence before and after the policy
246 introduction, grouped by type of wards, while Figure 2 displays all wards grouped together, by
247 standardising the timing of the introduction of the policy to zero. It shows the percentage of sickness
248 absences per ward before and after the policy introduction. From a visual examination, it appears
249 that the introduction of 12-hour shifts increased the percentage of sickness absences per week in all
250 wards.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

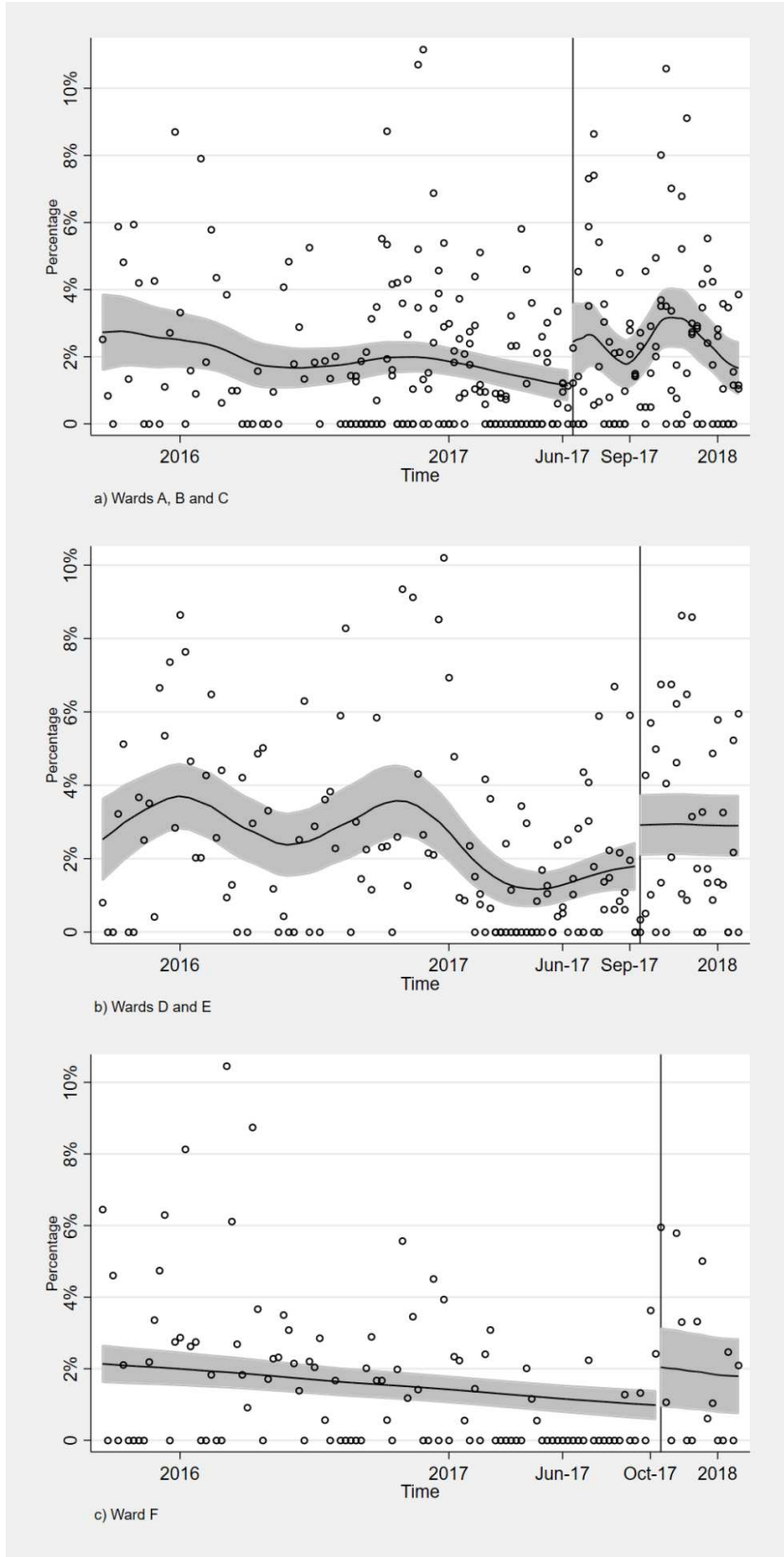
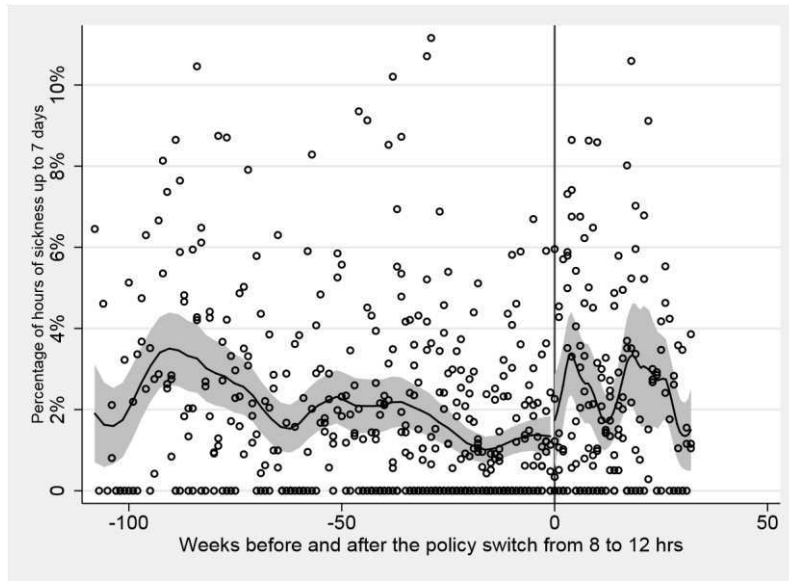


Figure 1: Change in policy timing per ward. The solid horizontal line is the local polynomial smooth of the dotted scatter values of percentage of hours of sickness absence up to 7 days using a triangle kernel function; the shaded area represents the 95% confidence interval around it and the vertical line indicates the timing of the introduction of the policy.

256



257

258 *Figure 2: Sickness absence in percentage up to 7 days, before and after policy implementation. Policy timing cut-off*
 259 *standardised at time zero. The solid horizontal line is the local polynomial smooth of the dotted scatter values using a*
 260 *triangle kernel function; the shaded area represents the 95% confidence interval around it and the vertical line indicates the*
 261 *timing of the introduction of the policy.*

262

263 Table 1 shows the number of observations and the descriptive statistics for all the variables included
 264 in the analysis (means and standard deviation) for the overall sample (1) and for the observations
 265 before (2) and after (3) the introduction of 12-hour shifts. The difference in means (t-test) (4)
 266 between (2) and (3) for all the covariates is also included.

267

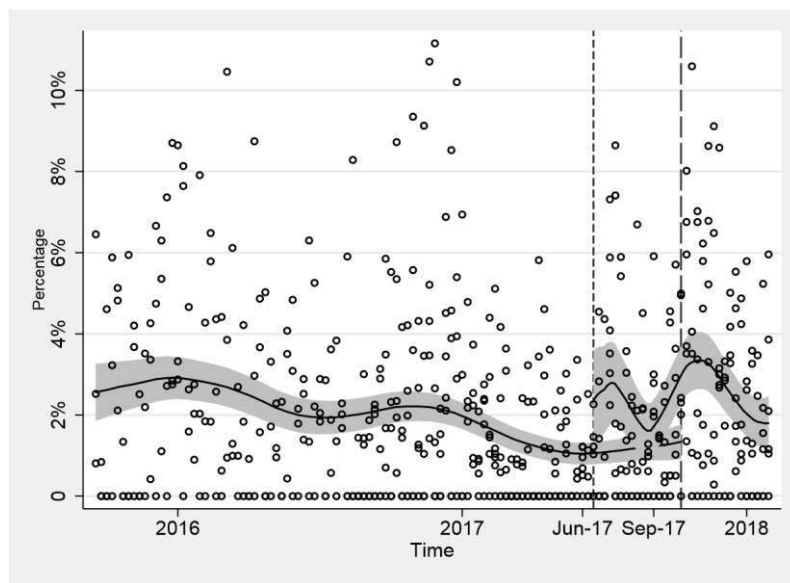
268 *Table 1: Descriptive statistics before and after the introduction of the 12-hour shift policy (means and standard deviation in*
 269 *parentheses) and difference in means test*

	(1)	(2)	(3)	(4)
	Overall	Before	After	
	Mean (SD)	Mean (SD)	Mean (SD)	Diff. Means
Outcome variables sickness up to 7 days				
Sickness hours per week (Perc. x 100)	1.96 (2.26)	1.68 (2.12)	2.45 (2.43)	0.76***
Control variables				
Casemix adjustment monthly average HoNOS score	18.72 (2.25)	18.65 (2.24)	18.85 (2.27)	0.20
Supercluster group 2 (Perc. x 100)	43.75 (33.73)	42.20 (34.06)	46.48 (33.06)	4.28
Number of patients per month	11.08 (2.96)	10.86 (2.91)	11.47 (3.01)	0.61*
Staff average age (in years)	45.01 (4.09)	45.16 (4.12)	44.74 (4.05)	0.42
Female staff (Perc. x 100)	78.90 (6.48)	79.40 (5.40)	78.03 (7.98)	1.37*
Staff ethnicity white (Perc. x 100)	94.32 (4.63)	94.75 (4.75)	93.57 (4.30)	1.17**
Registered Nurses (Perc. x 100)	35.88 (6.80)	36.93 (7.05)	34.02 (5.93)	2.91***
Number of staff per month	30.58 (8.96)	31.35 (10.49)	29.21 (4.98)	2.14*
Observations number of week wards	463	296	167	463

Note: * <0.1 , ** <0.05 , *** <0.01 . Pooled descriptive statistics, the sample is divided between before and after by the time of the implementation of the 12-hour shift policy for each ward.

For the outcome variable, the average percentage of sickness absences in hours per week (first row, Table 1), before the implementation of the policy, is 1.68% whilst after, it is 2.45% resulting in an increase of 0.76%, statistically significant at the 99% confidence level. There are no statistically significant differences associated with patient casemix before and after the introduction of the policy. However, there are some differences in the staff demographic composition (sex, ethnicity) of the wards before and after the introduction of the 12-hour shifts. These differences might be linked to changes in the total number of staff per ward as the percentage of registered nurses per ward has decreased by nearly 3% following the introduction of longer shifts.

283 Figure 3 shows the difference in timing in the introduction of the policy, that justifies the Difference-
 1
 284 in-Differences approach. In the intersection, the lower smoothed curve represents the wards (D,E
 3
 285 and F) with delayed timing in the application of the policy, between June and October 2017, i.e. the
 4
 286 control group, while the upper smoothed line represents the wards (A, B and C) with policy
 5
 287 implementation in June 2017, i.e. the treatment group. The figure suggests that the wards in the
 6
 288 control group maintained a more or less stable percentage of sickness absence, whereas in the
 7
 289 treatment group the sickness absence increased.



290
 291 *Figure 3: Difference-in-Differences identification strategy: The policy timing varies, three wards (A, B and C) introduced the*
 292 *12-hour shifts in June, two in September (wards D and E) and one in October 2017 (ward F). Between June 2017 to*
 293 *September 2017 there are three wards affected by the policy and three wards unaffected by the policy. The solid horizontal*
 294 *line is the local polynomial smooth of the dotted scatter values using a triangle kernel function; the shaded area represents*
 295 *the 95% confidence interval around it and the vertical lines indicate the timing of the introduction of the policy.*

3.2. Regression results

296
 297
 298 Table 2 displays the main estimation results. Columns (1)-(3) show the interrupted time series
 299 estimates whilst columns (4)-(6) present the difference-in-difference estimation results. The
 300 preferred specification includes all control variables given by column (3) for the Interrupted Time
 301 Series analysis and column (6) for the Difference-in-Differences analysis. According to the
 302 Interrupted Time Series results the switch to 12-hour shifts increased sickness absences by 0.7%,
 303 whilst Difference-in-Differences results estimates a slightly larger increase of approximately 1%.

304 However, there is no statistically significant difference between both estimations according to their
 305 confidence intervals. For full results, refer to Appendix A.

306 The test by means of a regression discontinuity framework shows no evidence of changes in the
 307 trend of covariates included in the analysis before and after the introduction of the 12-hour shifts,
 308 meaning that the policy has affected sickness absences, the outcome variable, but not the control
 309 variables. Estimation results of this test are reported in Appendix, Table A1.

310 Regarding the Interrupted Time Series specification, a test between fixed and random effects using
 311 the robust Hausman test ($\chi^2= 114.58$) finds in favour of the correlated random effect framework
 312 (Wooldridge, 2010). The Interrupted Time Series estimation reported in the results section
 313 corresponds to the fixed effect panel data estimation, which is the preferred estimation.

314 *Table 2: Interrupted time series (ITS) regression results and Difference-in-differences (DID) regression results*

	Interrupted Time Series			Difference-in-differences		
	(1)	(2)	(3)	(4)	(5)	(6)
Policy effect	0.736*** (0.177)	0.606** (0.250)	0.730** (0.307)	0.811** (0.230)	1.019** (0.305)	0.975** (0.347)
Control variables						
Monthly dummies	Yes	Yes	Yes	Yes	Yes	Yes
Patients casemix		Yes	Yes		Yes	Yes
Staff demographics			Yes			Yes
Observations	463	463	463	331	331	331

Note: * <0.1 , ** <0.05 , *** <0.01 . Standard errors (in parentheses) are clustered at ward level. Regressions include patient casemix indicated by the total HoNOS score, the psychotic supercluster and number of patients per month, and staff demographics are indicated by the average staff age, percentage of female staff, percentage of white staff, and the percentage of registered nurses.

3.3. Robustness checks

The series of checks performed suggest that our estimates are robust. A discussion of the results can be found in the Appendix.

4. Discussion

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

326 This study reports an evaluation of the impact of extending mental health nurses and health care
327 assistants' shifts from 8 to 12 hours. In particular, it explores the association between longer shifts
328 and sickness absence, a widely used outcome within occupational health research, by comparing
329 staff sickness absence rates before and after the introduction of a 12-hour shift pattern by means of
330 Interrupted Time Series and Differences-in-Differences estimation approaches. The change in shift
331 patterns from 8 to 12-hour shifts had a negative impact on 7-day sickness absence which increased
332 between 0.73% according to Interrupted Time Series analysis and 0.98% according to the Difference-
333 in-Differences estimation. All estimates are statistically significant at 95% confidence levels and
334 robust to various sensitivity checks. The results are far from negligible as in a typical ward where
335 there are 30 members of staff working full time (30 full-time equivalent), this would lead to
336 approximately 12 hours of sickness absence per week – i.e. sickness absences will be increased by a
337 complete shift per week, per ward as result of the new shift regime. These results are consistent
338 with findings from the only previous study from Dall'Ora et al. (2019) examining general healthcare
339 staff, who found that if a shift was scheduled to be 12-hours long, it would be 1.18 times more likely
340 that staff would miss a shift when compared to an 8-hour shift schedule. The odds increase to 1.24
341 when long sickness absences are also included in the sample.

342 This study makes three key contributions: (i) This is the first study to analyse the impact of 12-hour
343 shifts on sickness absences for mental health wards in England. The work from Dall'Ora et al. (2019)
344 is the only previous study on sickness absence, however their data came from inpatient general
345 adult wards. As previously stated, nursing mental health patients can be physically and emotionally
346 more demanding than nursing patients with physical health problems. The former often can be
347 impulsive, unpredictable, aggressive and can also often suffer physical health co-morbidities.
348 Therefore, the extension to 12-hour shifts in mental health wards might lead to worse staff
349 outcomes than those observed for general inpatient wards. (ii) The study uses longitudinal
350 administrative data of all staff members from the sample wards, moving away from self-reported
351 cross-sectional survey data often used to evaluate the impact of shift work on employee wellbeing.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

352 Jenkins and Elliott (2004) find that studies with self-reported questionnaires show inflation in
353 reported stress. The switch to 12-hour shifts in this sample was compulsory, thus overcoming
354 limitations of self-selection from previous studies where the policy has been optional. Furthermore,
355 in this study all the wards are observed before and after the policy, allowing an unbiased estimation
356 of the effect of the extended hours. This contrasts with Dall'Ora et al. (2019) who compare
357 outcomes of two groups of individuals (often working in different wards) who are exposed to two (or
358 more) different shifts patterns. In addition, the estimation strategy in this study accounts for a
359 comprehensive set of covariates that control for contextual factors. Previous studies fail to report in
360 sufficient detail contextual factors, such as staff skill mix or patient to nurse ratio, that are important
361 determinants of the impact of different shift patterns (Harris et al., 2015). (iii) Finally, the application
362 of Interrupted Time Series and Difference-in-Differences approaches is a methodological innovation
363 in this context since it allows the separation of the effect of the 12-hour shifts on sickness absences
364 from other confounders by using two different types of control groups. To our knowledge, this is the
365 first study to use causal inference in the estimation of the impact of 12-hour shifts on sickness
366 absences for mental health staff.

367 While this study makes important contributions, there are also limitations in terms of the
368 generalisability of these findings which are limited since only a small number of wards were
369 analysed, all of them belonging to the same hospital provider and therefore they might be different
370 in some respects to other wards in other hospitals. It is important to note that as Reid et al. (1991)
371 point out nurses' behaviours, in this case short term sickness absence, might not be independent of
372 their attitude towards the shift patterns they were working. With the data available, we are not able
373 to disentangle if some of the observed sickness absences might be the result of a protest against 12-
374 hour shifts rather than linked to health problems associated with the longer shifts. There is a small
375 chance that there might be some unobserved factors affecting sickness absence that have not been
376 accounted for. The latter is a common limitation of observational studies. Further, while our sickness
377 absence data was available at the individual level, we did not have individual identifiers to track

1
2 378 staff, and examine their characteristics, over time and therefore had to aggregate our analyses to
3 379 ward level.

4
5 380 Further research should look at the effect of 12-hour shifts over a longer period of time, at the
6
7 381 individual level, and evaluate the long run effects of working longer shifts. Not only the length of the
8
9 382 shifts can have negative effects on staff sickness absences, but also how those shifts are distributed
10
11 383 throughout the week. Ohayon et al. (2002) found that staff members from a Danish psychiatric
12
13 384 hospital working on rotating day-night shifts were more likely to take sick leave when compared
14
15 385 with those in the daytime group. Bourbonnais et al. (1992) found similar results for nurses from
16
17 386 seven hospitals in Quebec. Ball et al. (2015) highlight the difficulties in assessing the effect of 12-
18
19 387 hour shifts with little information about the practical ways the system is operated, i.e. how many
20
21 388 shifts are worked in a row, number and lengths of breaks, etc. The data in this study does not
22
23 389 include information on health care worker shift types or patterns and hence it is no possible to
24
25 390 identify their effect on sickness absences. Further information on the distribution of shifts is needed
26
27 391 to shed light on this issue.

28
29
30
31
32
33
34 392 In addition, most research evaluating the effect of longer shifts only focuses on staff acceptability
35
36 393 and work-life balance (Harris et al., 2015) and few studies evaluate the effects on patient safety and
37
38 394 experience (Griffiths et al., 2014). Future research should put more emphasis on patient measures as
39
40 395 well as including an analysis of the cost-effectiveness of longer shifts. Therefore, further evidence on
41
42 396 the costs and a wider set of outcomes on both patients and healthcare workers of a 12-hour shift
43
44 397 system is needed to allow managers and policymakers a complete assessment of the policy.

45
46
47
48
49 398 Moreover, the data used in this study is at an aggregate level, where sickness absences as well as
50
51 399 control variables are captured at week/ward level. Longer shifts entail different risks and benefits for
52
53 400 different staff at different times (Ball et al., 2015), e.g. there might be greater risks for the older
54
55 401 person working long shifts (Chudleigh et al., 2005). Further research could consider individual data
56
57 402 to disentangle which demographic groups are more likely to suffer the negative effects of 12-hour
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

403 shifts. According to Harris et al. (2015) the success of the 12-hour shift will ultimately be dependent
404 upon the support and cooperation of the staff involved. A qualitative study on the effects of 12-
405 hours shifts on the same mental health hospital from this study, (Suter et al., 2019) , shows that,
406 whilst a 12-hour shift was received positively by some staff as it offered them flexibility, when
407 implemented as a mandatory work pattern, the element of employee choice is eliminated leading to
408 dissatisfaction and disinterment amongst others. If staff have negative views and negative work
409 outcomes persist in the long-run then the policy is likely to fail.

410 5. Conclusion

411 Our findings suggest that the switch to 12-hour shifts has increased sickness absences and, whilst
412 wider aspects of the policy change are important, including alternative outcomes and overall cost-
413 effectiveness, if only sickness absences were taken into account, hospital managers should consider
414 reverting to 8-hour shifts. Our analysis is relevant to other hospitals within England, and
415 internationally, which are increasingly moving towards these shift patterns.

416 Acknowledgments

417 The research team acknowledges the support of the National Institute for Health Research Clinical
418 Research Network (NIHR CRN).

419 Ethical Approval

420 This study received ethics approval from the Health Research Authority - 18-HRA-0454

421 Funding sources

422 This work is part funded by the Wellcome Trust [ref 204829] through the Centre for Future (CFH) at
423 the University of York

424 Conflict of interest

425 None

429 References

- 430
- 431 Angrist, J.D., Pischke, J.S., 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*.
432 Princeton University Press, London, UK.
- 433 Ball, J., Maben, J., Murrells, T., Day, T., Griffiths, P., 2015. 12-hour shifts: Prevalence, views and
434 impact. King's College London, National Nursing Research Unit.
- 435 Bloodworth, C., Lea, A., Lane, S., Ginn, R., 2001. Challenging the myth of the 12-hour shift: a pilot
436 evaluation. *Nurs Stand* 15 (29), 33-36.
- 437 Bonner, G., Lowe, T., Rawcliffe, D., Wellman, N., 2002. Trauma for all: a pilot study of the
438 subjective experience of physical restraint for mental health inpatients and staff in the UK.
439 *Journal of Psychiatric Mental Health Nursing* 9 (4), 465-473.
- 440 Bourbonnais, R., Vinet, A., Vezina, M., Gingras, S., 1992. Certified sick leave as a non-specific
441 morbidity indicator: a case-referent study among nurses. *Occupational Environmental
442 Medicine* 49 (10), 673-678.
- 443 British Medical Association, 2018. *Lost in transit? Funding for mental health services in England
444 2018*. British Medical Association.
- 445 Calonico, S., Cattaneo, M.D., Farrell, M.H., Titiunik, R., 2019. Regression Discontinuity Designs
446 Using Covariates. *Review of Economics and Statistics* 101 (3), 442-451.
- 447 Cameron, A.C., Trivedi, P.K., 2005. *Microeconometrics: Methods and Applications*. Cambridge
448 University Press.
- 449 Chen, J., Davis, K.G., Daraiseh, N.M., Pan, W., Davis, L.S., 2014. Fatigue and recovery in 12-hour
450 dayshift hospital nurses. *Journal of Nursing Management* 22 (5), 593-603.
- 451 Chudleigh, J., Fletcher, M., Gould, D., 2005. Infection control in neonatal intensive care units. *Journal
452 of Hospital Infection* 61 (2), 123-129.
- 453 Cook, T.D., Campbell, D.T., Shadish, W., 2002. *Experimental and quasi-experimental designs for
454 generalized causal inference*. Houghton Mifflin Boston.
- 455 Dall' Ora, Chiara., Griffiths, P., Ball, J., Simon, M., Aiken, L.H., 2015. Association of 12 h shifts and
456 nurses' job satisfaction, burnout and intention to leave: findings from a cross-sectional study
457 of 12 European countries. *BMJ open* 5 (9), e008331.
- 458 Dall' Ora, C., Griffiths, P., Emmanuel, T., Rafferty, A.M., Ewings, S., RN4CAST Consortium,
459 Sermeus, W., Van den Heede, K., Bruyneel, L., Lesaffre, E. and Aiken, L., 2020. 12- hr
460 shifts in nursing: Do they remove unproductive time and information loss or do they reduce
461 education and discussion opportunities for nurses? A cross- sectional study in 12 European
462 countries. *Journal of clinical nursing*, 29 (1-2), 53-59.
- 463 Dall'Ora, C., Ball, J., Redfern, O., Recio Saucedo, A., Maruotti, A., Meredith, P., Griffiths, P., 2019.
464 Are long nursing shifts on hospital wards associated with sickness absence? A longitudinal
465 retrospective observational study. *Journal of Nursing Management* 27 (1), 19-26.
- 466 Edwards, D., Burnard, P., 2003. A systematic review of stress and stress management interventions
467 for mental health nurses. *J Adv Nurs* 42 (2), 169-200.
- 468 Fothergill, A., Edwards, D., Burnard, P., 2004. Stress, burnout, coping and stress management in
469 psychiatrists: findings from a systematic review. *Int J Soc Psychiatry* 50 (1), 54-65.
- 470 Geiger-Brown, J., Rogers, V.E., Trinkoff, A.M., Kane, R.L., Bausell, R.B., Scharf, S.M., 2012. Sleep,
471 sleepiness, fatigue, and performance of 12-hour-shift nurses. *Chronobiol Int* 29 (2), 211-219.
- 472 Griffiths, P., Dall'Ora, C., Simon, M., Ball, J., Lindqvist, R., Rafferty, A.-M., Schoonhoven, L.,
473 Tishelman, C., Aiken, L.H., 2014. Nurses' shift length and overtime working in 12 European
474 countries: the association with perceived quality of care and patient safety. *Medical Care* 52
475 (11), 975.
- 476 Hagen, J., Knizek, B.L., Hjelmeland, H., 2017. Mental Health Nurses' Experiences of Caring for
477 Suicidal Patients in Psychiatric Wards: An Emotional Endeavor. *Archives of psychiatric
478 nursing* 31 (1), 31-37.
- 479 Harris, R., Sims, S., Parr, J., Davies, N., 2015. Impact of 12 h shift patterns in nursing: A scoping
480 review. *International Journal of Nursing Studies* 52 (2), 605-634.
- 481 Hoffman, A.J., Scott, L.D., 2003. Role stress and career satisfaction among registered nurses by work
482 shift patterns. *Journal of Nursing Administration* 33 (6), 337-342.

483 Holloway, F., Szmukler, G., Carson, J., 2000. Support system. 1.Introduction. *Advances in*
 1 484 *Psychiatric Treatment* 6 (3), 226-235

2 485 Ilhan, M.N., Durukan, E., Aras, E., Turkcuoglu, S., Aygun, R., 2006. Long working hours increase
 3 486 the risk of sharp and needlestick injury in nurses: the need for new policy implication. *Journal*
 4 487 *of Advance Nursing* 56 (5), 563-568.

5 488 Jacobs, R., 2009. Investigating patient outcome measures in mental health. In: CHE Research Paper.
 6 489 Centre for Health Economics.

7 490 Jenkins, R., Elliott, P., 2004. Stressors, burnout and social support: nurses in acute mental health
 8 491 settings. *Journal of Advance Nursing* 48 (6), 622-631.

9 492 Kavanagh, K.H., 1991. Invisibility and selective avoidance: gender and ethnicity in psychiatry and
 10 493 psychiatric nursing staff interaction. *Cult Med Psychiatry* 15 (2), 245-274.

11 494 Kings Fund, 2015. Mental health services under pressure. Kings Fund.

12 495 Knauth, P., 2007. Extended work periods. *Industrial Health* 45 (1), 125-136.

13 496 Lea, A., Bloodworth, C., 2002. Longer shift, better care? *Nurs Older People* 14 (1), 32-33.

14 497 Lee, D.S., Lemieux, T., 2010. Regression Discontinuity Designs in Economics. *Journal of Economic*
 15 498 *Literature* 48 (2), 281-355.

16 499 Lipscomb, J.A., Trinkoff, A.M., Geiger-Brown, J., Brady, B., 2002. Work-schedule characteristics
 17 500 and reported musculoskeletal disorders of registered nurses. *Scand J Work Environ Health* 28
 18 501 (6), 394-401.

19 502 Merkus, S.L., Van Drongelen, A., Holte, K.A., Labriola, M., Lund, T., Van Mechelen, W., Van der
 20 503 Beek, A.J., 2012. The association between shift work and sick leave: a systematic review.
 21 504 *Occup Environ Med* 69 (10), 701-712.

22 505 Moore, K.A., Cooper, C.L., 1996. Stress in mental health professionals: A theoretical overview.
 23 506 *International Journal of Social Psychiatry* 42 (2), 82-89.

24 507 Moscelli, G., Jacobs, R., Gutacker, N., Aragon, A., Monserratt, M.J., Chalkley, M.J., Mason, A.,
 25 508 Boehnke, J.R.J.H.E., 2018. Prospective payment systems and discretionary coding-Evidence
 26 509 from English mental health providers.

27 510 NHS Evidence, 2010. Moving to 12-hour shift patterns: to increase continuity and reduce costs.
 28 511 Basingstoke and North Hampshire NHS Foundation Trust.

29 512 O'Connor, K., Muller Neff, D., Pitman, S., 2018. Burnout in mental health professionals: A systematic
 30 513 review and meta-analysis of prevalence and determinants. *Eur Psychiatry* 53, 74-99.

31 514 Ohayon, M.M., Lemoine, P., Arnaud-Briant, V., Dreyfus, M., 2002. Prevalence and consequences of
 32 515 sleep disorders in a shift worker population. *Journal of psychosomatic research* 53 (1), 577-
 33 516 583.

34 517 Penfold, R.B., Zhang, F., 2013. Use of interrupted time series analysis in evaluating health care
 35 518 quality improvements. *Academic Pediatrics* 13 (6), S38-S44.

36 519 Raediker, B., Janßen, D., Schomann, C. and Nachreiner, F., 2006. Extended working hours and
 37 520 health. *Chronobiology International*, 23(6),1305-1316.

38 521 Reid, N., Todd, C., Robinson, G., 1991. Educational activities on wards under 12 hour shifts. *Int J*
 39 522 *Nurs Stud* 28 (1), 47-54.

40 523 Renwick, L., Lavelle, M., Brennan, G., Stewart, D., James, K., Richardson, M., Williams, H., Price,
 41 524 O. and Bowers, L., 2016. Physical injury and workplace assault in UK mental health trusts:
 42 525 An analysis of formal reports. *International journal of mental health nursing*,25(4), 355-366.

43 526 Stimpfel, A.W., Sloane, D.M., Aiken, L.H., 2012. The longer the shifts for hospital nurses, the higher
 44 527 the levels of burnout and patient dissatisfaction. *Health Affairs* 31 (11), 2501-2509.

45 528 Stone, P.W., Du, Y., Cowell, R., Amsterdam, N., Helfrich, T.A., Linn, R.W., Gladstein, A., Walsh,
 46 529 M., Mojica, L.A., 2006. Comparison of nurse, system and quality patient care outcomes in 8-
 47 530 hour and 12-hour shifts. *Med Care* 44 (12), 1099-1106.

48 531 Suter, J., Kowalski, T., Rodriguez Santana, I., Jacobs, R., Chalkley, M., Anaya Montes, M., In Press,
 49 532 2019. The impact of extended shifts on work-life conflict: a qualitative study of experiences
 50 533 of mental healthcare staff. University of York In Press.

51 534 Todd, C., Reid, N., Robinson, G., 1989. The quality of nursing care on wards working eight and
 52 535 twelve hour shifts: a repeated measures study using the MONITOR index of quality of care.
 53 536 *International Journal of Nursing Studies* 26 (4), 359-368.

537 Trinkoff, A.M., Le, R., Geiger-Brown, J., Lipscomb, J., Lang, G., 2006. Longitudinal relationship of
1 538 work hours, mandatory overtime, and on-call to musculoskeletal problems in nurses.
2 539 American Journal of Industrial Medicine 49 (11), 964-971.
3 540 Vik, A.G., MacKay, R.C., 1982. How does the 12-hour shift affect patient care? Journal of Nursing
4 541 Administration 12 (1), 11-14.
5 542 Wallace, J.E., Lemaire, J.B., Ghali, W.A., 2009. Physician wellness: a missing quality indicator. The
6 543 Lancet 374 (9702), 1714-1721.
7 544 Wooldridge, J.M., 2010. Correlated random effects models with unbalanced panels. Michigan State
8 545 University, Department of Economics.
9 546 World Health Organization, 2007. Atlas: nurses in mental health 2007. World Health Organization.

11 547

12 548

13 549

14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

550 Appendix

551 *Table A1: Discontinuity of covariates at the time of introduction of the policy*

	RD Effect	Robust P-value
Casemix adjustment monthly average HoNOS score	-0.47	0.366
Supercluster group 2 (Perc. x 100)	0.919	0.944
Number of patients per month	-0.28	0.914
Number of staff per month	-0.147	0.987
Female staff (Perc. x 100)	0.403	0.983
Staff ethnicity white (Perc. x 100)	-0.16	0.913
Staff average age (in years)	-0.086	0.968
Registered Nurses (Perc. x 100)	0.468	0.757

552 Note: Test of discontinuity of covariates at the time of introduction of the policy by means of the 'rdrobust'
 553 Stata command. The first column, RD effect, reports the regression discontinuity coefficient for each covariate.
 554 The second column reports the Robust P-value. None of the RD effects are statistically significant.

555 Robustness Checks

556 Regarding the robustness check number one (i.e the placebo test), Figure A1 shows that this
 557 arbitrary policy does not seem to change the percentage of sickness absences per ward/week.
 558 Moreover, Table A2 presents the regression results for the Interrupted Time Series and Difference-
 559 in-Differences for the placebo test and no effect on the rate of sickness absence is found, suggesting
 560 that the findings are robust.

561 Regarding the robustness checks two, using pooled Ordinary Least Squares instead of fixed effects
 562 panel data, and three, using sickness absences of 14 days and 2 days, all estimations with the
 563 exception of sickness absences up to 2 days provide estimates with confidence intervals that contain
 564 the main results presented (details in Table A3), suggesting results are driven by sickness absences
 565 from 3 to 7 days. Robustness check number four, where the policy implementation date is moved
 566 artificially six weeks prior and six weeks post the actual date, shows no evidence for anticipatory
 567 effects up to two weeks before the policy; however a consistent and increasing effect was found from
 568 the introduction of the policy and up to three weeks after, with the greatest effect found in the third
 569 week. The latter suggests the existence of some delayed effects which are expected, since the effects
 570 on sickness absence due to the introduction of the policy are more likely to increase in the
 571 medium/long run when tiredness, stress and burnout accumulates due to the ongoing experience of
 572 longer shifts. Results from robustness check (iv) are available upon request.

573

1

574

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

575

23

24

576

25

577

26

578

27

579

28

29

580

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

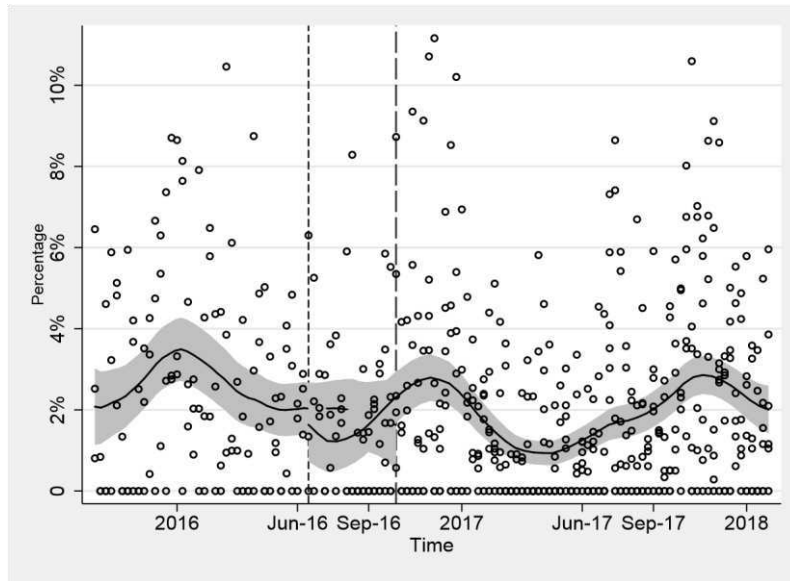


Figure A1 Placebo test, introduction of the policy one year before. The solid horizontal line is the local polynomial smooth of the dotted scatter values using a triangle kernel function; the shaded area represents the 95% confidence interval around it and the vertical lines indicate the timing of the introduction of the policy

581

582 *Table A2: Main estimation results for the Interrupted Time Series (ITS) (1) and Difference-in-Differences (DID) (3) models*
 583 *and estimation results for the ITS (2) and DID placebo tests (4)*

	ITS		DID	
	(1)	(2)	(3)	(4)
Policy effect (ITS)				
Policy effect (T)	0.729** (0.307)	-0.684* (0.390)		
Trend	0.016*** (0.004)	0.024 (0.041)		
Iteration trend x T	-0.009 (0.016)	-0.03 (0.041)		
Policy effect (diff-diff)				
Policy effect (iteration time of policy x treated wards)			0.975** (0.347)	0.350 (1.319)
Time of the policy			-0.496 (0.336)	0.612 (1.319)
Treated wards			0.595 (0.457)	-17.979 (9.299)
Patient control variables				
Casemix adjustment monthly average HoNOS score	0.046** (0.022)	0.064*** (0.018)	0.066 (0.035)	-0.442 (0.355)
Supercluster group 2 (Perc. x 100)	-0.031*** (0.008)	-0.030*** (0.008)	-2.720*** (0.629)	15.263 (8.252)
Number of patients per month	0.01 (0.029)	0.01 (0.037)	0.042 (0.021)	-0.518*** (0.007)
Staff control variables				
Staff average age (in years)	0.136 (0.181)	0.2 (0.174)	-0.126** (0.048)	-0.471 (0.342)
Female staff (Perc. x 100)	0.001 (0.030)	0.004 (0.032)	3.897** (1.319)	-32.202** (3.673)
Staff ethnicity white (Perc. x 100)	-0.001 (0.043)	-0.024 (0.045)	-1.746 (2.778)	-230.413* (74.593)
Registered Nurses (Perc. x 100)	0 (0.038)	-0.005 (0.043)	0.859 (2.191)	34.375** (4.567)
Number of staff	0.038 (0.024)	0.04 (0.026)	0.025* (0.010)	-0.003 (0.025)
Time variables				
Monthly dummies	Yes	Yes	Yes	Yes
Observations	463	463	331	45

584 Note: * <0.1 , ** <0.05 , *** <0.01 . Standard errors (in parentheses) are clustered at ward level. Columns (1) and
 585 (3) correspond to main results in the fixed effect model and difference-in-difference, respectively. Columns (2)
 586 and (4) show the results for the placebo test, that assumes the policy was introduced one year before on the
 587 same month and day, for the ITS and the DID models respective

Table A3: Robustness check Interrupted Time Series (ITS) specification

	(1)	(2)	(3)	(4)	(5)
Policy effect					
Policy effect (T)	0.813** (0.324)	0.730** (0.298)	0.706** (0.302)	0.042 (0.241)	0.900** (0.446)
Trend	0.013** (0.006)	0.069 (0.058)	0.188 (1.123)	0.006 (0.007)	-0.022** (0.008)
Iteration trend x T	-0.006 (0.018)	-0.013 (0.024)	-0.015 (0.024)	-0.012 (0.011)	0.004 (0.010)
Patient control variables					
Casemix adjustment monthly average HoNOS score	0.047 (0.030)	0.042 (0.027)	0.042 (0.027)	-0.031 (0.025)	0.101*** (0.039)
Supercluster group 2 (Perc. x 100)	-0.020*** (0.002)	-0.031*** (0.009)	-0.031*** (0.009)	-0.016** (0.004)	-0.037*** (0.011)
Number of patients per month	0.036* (0.019)	0.006 (0.038)	0.001 (0.042)	-0.016 (0.019)	0.051 (0.045)
Staff control variables					
Staff average age (in years)	-0.177*** (0.024)	0.145 (0.177)	0.161 (0.176)	0.184 (0.117)	0.031 (0.223)
Female staff (Perc. x 100)	0.014* (0.008)	0 (0.027)	0.001 (0.027)	0.001 (0.025)	0.023 (0.050)
Staff ethnicity white (Perc. x 100)	-0.025 (0.029)	-0.005 (0.038)	-0.006 (0.039)	-0.042* (0.021)	0.035 (0.073)
Registered Nurses (Perc. x 100)	-0.025 (0.018)	-0.003 (0.034)	-0.002 (0.038)	-0.028* (0.014)	0.007 (0.057)
Number of staff	0.01 (0.016)	0.038* (0.023)	0.040* (0.024)	0.029 (0.022)	0.036 (0.041)
Time variables					
Monthly dummies	Yes	Yes	Yes	Yes	Yes
Yearly dummies		Yes	Yes		
Weekly dummies			Yes		
Observations	463	463	463	463	463

Note: * <0.1 , ** <0.05 , *** <0.01 . Standard errors (in parentheses) are clustered at ward level. Column (1) shows the estimation results for the pooled OLS; Column(2) shows the estimates for fixed effect with yearly dummies and (3) with the addition of weekly dummies Column (4) shows results for sickness absence up to two days and (5) up to 14 days

Reviewer #1:

Thanks for inviting me to review this interesting and very timely paper on 12-h shifts and sickness absence in MH nurses.

Comment	Response	Action
<p>1. Authors used a very robust analytic approach, which increases our ability to estimate causality although using observational data. My main concern is the level of aggregation - by focussing on sick hours per week as opposed to looking at the individual level, I feel the authors are losing on granularity.</p>	<p>We agree. Ideally, we would have liked to carry out an individual level analysis; however, the nature of the data didn't allow it. Our data on sickness absences was at the individual level (i.e. we observed that a person was sick for X hours between two specific dates), however we didn't have individual identifiers to be able to track nurses over time. In addition, we didn't have information about the staff sociodemographic characteristics at the individual level. We were however given information about sociodemographic composition of the wards at aggregate level (e.g. % female nurses per ward per month). The approach we used was the most granular analysis we could undertake given the available data. We now discuss this limitation in more detail.</p>	<p>Line 378:</p> <p>Further, while our sickness absence data was available at the individual level, we did not have individual identifiers to track staff, and examine their characteristics, over time and therefore had to aggregate our analyses to ward level.</p> <p>Further research should look at the effect of 12-hour shifts over a longer period of time, at the individual level, and evaluate the long run effects of working longer shifts.</p>
<p>2. I have some specific comments I would like authors to address:</p> <p>The introduction, while quite informative could perhaps be shortened to make sure the focus is on 12-h shifts in mental health and sickness absence. As it stands now, the first 2 paragraphs of the paper are about MH and this detracts attention from the focus of the paper.</p>	<p>We acknowledge this background section could have been too detailed. We have shortened these two paragraphs significantly and combined them into one.</p>	<p>Line 44:</p> <p>The mental health workforce face all sorts of stressors; many of those are common to all health care staff, such as limited resources or overcrowded wards, whilst others are specific to the mental health care setting. Examples of the latter are: dealing with patients' physical and/or verbal violent behaviour (Holloway et al., 2000; Renwick et al., 2016), the use of coercive measures such as restraint and detention of patients (Bonner et</p>

		al., 2002) and continuous monitoring of patients at risk of self-harm (Hagen et al., 2017), in the most extreme cases with patients' suicide (Fothergill et al., 2004) and with the public inquiries associated with these deaths, that often allocate blame to staff members (Holloway, Szmukler et al. 2000). According to O'Connor et al. (2018), the mental health workforce reports higher emotional exhaustion than emergency nurses and equal burnout to cancer professionals. The described stressors can explain, at least partially, the difficulties in recruiting and retaining mental health workforce. According to the World Health Organization (2007), there is a growing concern worldwide around mental health workforce capacity, characterized by staff shortages and high attrition rates. With excessive stress stemming from intense patient interactions over extended periods of time exposure to this working environment over a longer shift may have a greater impact on employee wellbeing (Edwards & Burnard, 2003).
3. The Griffiths et al 2014 paper did not find that handovers and overlaps are regarded as unproductive and associated with a larger number of errors due to discontinuity of care, can you please find a more appropriate reference?	We have corrected this with an alternative reference.	Line 73: ...handovers and overlaps are regarded as unproductive and associated with a larger number of errors due to discontinuity of care (NHS Evidence, 2010).
4. Also, the Griffiths et al paper considered all EU countries, not only England as you wrote at line 94.	We have modified the sentence accordingly.	Line 83: An analysis of the European RN4CAST survey, a multi-country cross-sectional nurse

		<p>workforce study, shows that working 12-hour shifts is associated with higher odds of poor quality of care, greater risk of necessary nurse care left undone and higher odds of reporting being dissatisfied with their jobs (Griffiths et al., 2014).</p> <p>We also added on line 86:</p> <p>Nurses from the RN4CAST survey also reported that longer shifts reduced the opportunities to discuss patient care and to participate in continuing education activities (Dall’Ora 2020)</p>
<p>5. Please remove "quantitative" from the sentence at line 98. Similarly, please remove "quantitative" from the sentence at line 105.</p>	<p>This is now done.</p>	<p>Line 102:</p> <p>This paper contributes to the literature by reporting an evaluation of the impact of a change from 8 to 12-hour shift patterns on sickness absences in a large National Health Service (NHS) mental health care provider in England by means of two causal inference methods: Interrupted Time Series and Difference-in-Difference.</p> <p>Line 105:</p> <p>This is the second study, after Dall’Ora et al. (2019), that explores the association between longer shifts and sickness absences.</p>
<p>6. Lines 115 to 130 belong to the methods.</p>	<p>This content has now been placed in the method section and repetition removed.</p>	<p>Line 116:</p> <p>The subjects of this study are nurses and health care assistants from six</p>

		<p>geographically dispersed inpatient mental health wards from a large mental health NHS provider in England.</p> <p>Line 120:</p> <p>The data are a combination of roster records and routine hospital administrative data available before and after the introduction of the 12-hour shifts.</p>
<p>7. Can you please incorporate Lines 131-148 to the discussion session - that's where the contributions of the paper should be stated.</p>	<p>This has been now placed in the discussion section.</p>	<p>Lines 333-356:</p> <p>This study makes three key contributions: (i) This is the first study to analyse the impact of 12-hour shifts on sickness absences for mental health wards in the UK. The work from Dall'Ora et al. (2019) is the only previous study on sickness absence, however their data came from inpatient general adult wards. As previously stated, nursing mental health patients can be physically and emotionally more demanding than nursing patients with physical health problems. The former often can be impulsive, unpredictable, aggressive and can also often suffer physical health co-morbidities. Therefore, the extension to 12-hour shifts in mental health wards might lead to worse staff outcomes than those observed for general inpatient wards. (ii) The study uses longitudinal administrative data of all staff members from the sample wards, moving away from self-reported cross-sectional survey data often used to evaluate the impact of shift work on employee wellbeing.</p>

		<p>.....</p> <p>(iii) Finally, the application of ITS and Diff-in-Diff approaches is a methodological innovation in this context since it allows the separation of the effect of the 12-hour shifts on sickness absences from other confounders by using two different types of control groups. To our knowledge, this is the first study to use causal inference in the estimation of the impact of 12-hour shifts on sickness absences for mental health staff.</p>
<p>8. The background should be about what the problem is, why it is important, what has been studied before and what do we still need to know. Please restructure your background to include these elements.</p>	<p>We have significantly shortened (see response to point 2 above) and restructured the Introduction. We moved up the section on previous literature and the importance of sickness absence as an outcome. We then state the aim of the paper and the key contribution of our paper compared to what has gone before.</p>	<p>Lines 89-99:</p> <p>There is a positive association between the amount of working hours and impairments to employees physical and mental health (Raediker, 2006). Previous studies, using cross-sectional and subjective nurse-reported data, have explored the effect of longer shifts on nurse health outcomes. Past findings show that nurses working longer shifts reported increased fatigue (Chen et al., 2014, Geiger-Brown et al., 2012, Lea and Bloodworth, 2002), greater stress burnout (Hoffman and Scott, 2003, Stimpfel et al., 2012) and higher risk of suffering work-related musculoskeletal disorders (Lipscomb et al., 2002, Trinkoff et al., 2006) all of which are likely to lead to increased sickness absence.</p> <p>Sickness absence is a widely used outcome in health occupational research as it relates to workers' social, economic and</p>

		<p>psychological processes (Merkus et al., 2012). Moreover, sickness absences are associated with several negative outcomes such as salary loss to the employee, salary costs of replacement staff, productivity losses and reduced quality of services.</p> <p>Line 105:</p> <p>The aim is to identify whether there is a statistically significant association between longer shifts and sickness absence by comparing staff short-term sickness absence rates before and after the policy implementation.</p> <p>Line 111:</p> <p>As nursing mental health patients can be physically and emotionally more demanding than nursing patients with physical health problems, the extension to 12-hour shifts in mental health wards might lead to worse staff outcomes than those observed for general inpatient wards.</p>
<p>9. At line 177 you start mentioning control variables but you have not specified the main model you are deploying to measure the association between 12-h shifts and sickness absence yet. Please do report this before moving on to control variables.</p>	<p>We no longer mention control variables in the Introduction. We do however state what our main estimations methods are now in the Introduction.</p> <p>We also specify the three main model specifications in the section about econometric strategy.</p>	<p>Line 102:</p> <p>This paper contributes to the literature by reporting an evaluation of the impact of a change from 8 to 12-hour shift patterns on sickness absences in a large National Health Service (NHS) mental health care provider in England by means of two causal inference methods: Interrupted Time Series and Difference-in-Difference.</p>

		<p>Line 174: For both the ITS and DID three model specifications are estimated: (1) that only includes the policy effect parameter and time dummies; (2) which includes all the variables in (1) plus patient casemix. and (3) which includes all the variables in (2) plus staff demographic variables.</p>
<p>10. Also, all these information come under the section "data" but this is not correct as it is already describing variables, outcome and analysis aspects.</p>	<p>We have divided the Methods section into three Sub-sections, including 2.3 Control variables.</p>	<p>Line 131 2.2 Dependent variable: sickness absences</p> <p>Line 143: 2.3 Control variables</p> <p>Line 169: 2.4 Econometric strategy</p>
<p>11. I suggest placing the whole section on robustness check as an appendix - it is very specialist and would probably be of interest to a minority of readers</p>	<p>The results of robustness checks have been placed in the Appendix now.</p>	<p>Line 314: The series of checks performed suggest that our estimates are robust. A discussion of the results can be found in the Appendix.</p>
<p>12. I applaud the authors for the robustness of their analysis approach, however, by looking at "sick hours per week", they lose granularity - they do not know what shifts were worked by individuals before they went off sick, and they cannot control for personal factors. I think authors should expand discussion of this limitation in the discussion.</p>	<p>Please see response to point 1 above.</p>	<p>Line 366: Further, while our sickness absence data was available at the individual level, we did not have individual identifiers to track staff, and examine their characteristics, over time and therefore had to aggregate our analyses to ward level.</p>

		Further research should look at the effect of 12-hour shifts over a longer period of time, at the individual level, and evaluate the long run effects of working longer shifts.
--	--	---

Reviewer #2:

I think it is an interesting paper but my main reservations are the following:

Comment	Response	Action
<p>1. The introduction should include a theory for why long work hours may be related to increased risk for ill-health, i.e. describe potential mechanisms.</p>	<p>We have significantly amended the Introduction (see response to Reviewer #1 points 2,8,9). Whilst we have shortened sections about the pressures that mental health staff are under, we have kept some text on why these stressors could affect their health and wellbeing and we have added text on why specifically mental health nurses may experience more sickness absence.</p>	<p>Line 56: With excessive stress stemming from intense patient interactions over extended periods of time exposure to this working environment over a longer shift may have a greater impact on employee wellbeing (Edwards & Burnard, 2003). Line 89: There is a positive association between the amount of working hours and impairments to employees physical and mental health (Raediker, 2006). Line 109: As nursing mental health patients can be physically and emotionally more demanding than nursing patients with physical health problems, the extension to 12-hour shifts in mental health wards might lead to worse staff outcomes than those observed for general inpatient wards.</p>
<p>2. In the methods: make it more clear how many persons actually are included in the analysis. It's the sentence "this focus results in dropping long term sickness absences which constitutes approximately 30% of the sample" that makes me a bit confused. Are those to be removed from the N reported in the</p>	<p>N= 463 as data is aggregated and reported by week and ward. The 30% was a reference to the individual observations (those that captured long term absences) before being aggregated into ward/week level data. We acknowledge this may lead to confusion and we have now clarified that we are dropping individual observations.</p>	<p>Line 138: This focus results in dropping 30% of individual observations for sickness absences before aggregating the sample to the ward / week level.</p>

paragraph above?		
3. If possible, I would like to have a description of what kind of shift patterns they have (e.g. rotating, nights etc)	We recognise the importance of different types of shift patterns on understanding the effect on sickness absences. Whilst we know that these healthcare workers work day and night shifts, we don't have information about individual healthcare workers' patterns and hence cannot include these as control variables.	Line 378: The data in this study does not include information of health care worker shift types or patterns and hence it is no possible to identify their effect on sickness absences. Further information on the distribution of shifts is needed to shed light on this issue.
4. they are followed for a limited time, are these fluctuations an effect of 12 shifts or some other factors that happens at that time?	We control for a set of observable characteristics and apply two different econometric techniques with the aim of capturing the causal effect of 12-hour shifts on sickness absences. We exploit all the available data we have following the introduction of the policy changes and have a great level of confidence that the reported estimates show the effect of 12-hour shifts on sickness absences. Our methods control for seasonality effects, ward and hospital effects in addition to all the observable characteristics. We also run several robustness checks. Nonetheless, as is common in this type of observational study, there is small chance that there might be some unobserved factors which could affect sickness absence that we have not been able to account for.	Line 362: With the data available, we are not able to disentangle if some of the observed sickness absences might be the result of a protest against 12-hour shifts rather than linked to health problems associated with the longer shifts. There is a small chance that there might be some unobserved factors affecting sickness absence that have not been accounted for. The latter is a common limitation of observational studies.
5. I would like to know how big N is needed for sufficient power.	Our data captures the universe of nurses and healthcare assistants of the six studied wards, so we are using all the available data for the six wards where the longer shifts were introduced. The study is therefore fully powered.	No changes made.

Conflict of Interest: NONE

Credit Author Statement

IRS, MAM, RJ, and MC: contributed to the design and implementation of the research and to the analysis of the results. MAM: carried out the data analysis. IRS: wrote the initial draft with help from MAM. RJ: Supervised the study. JS, TK, RJ, IRS and MC: worked on the funding acquisition. All authors provided critical feedback and worked in the writing, reviewing and editing process .