

©2021. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International <http://creativecommons.org/about/downloads>



This is the accepted version of this paper. The version of record is available at:  
<https://doi.org/10.1016/j.chilyouth.2021.106188>

# Exploring the impact of Ofsted inspections on performance in children's social care

## Abstract

Children's social care services in England are inspected by a government agency, Ofsted. This paper reports on the impact of Ofsted inspections on rates of intervention and other performance measures, such as expenditure and workforce, drawing on a quantitative analysis of national datasets. Annual published returns from 150 English local authorities from 2009-19 were combined with results from Ofsted inspections under the Single Inspection Framework (2014-17), which covered all local authorities. Interrupted time series (ITS) analysis was undertaken to find out whether an inspection had a significant effect on performance in the context of trends pre- and post-inspection, and whether this effect varied according to the type of judgement. The results showed discontinuities in a range of indicators in the year of an inspection and the year afterwards. An increase in rates of child protection interventions occurred at the time of an inspection, which was most pronounced in local authorities receiving an inadequate judgement. An inadequate judgement led to a steep rise in spending and use of agency workers compared to other local authorities. The findings suggest that a short-term spike in intervention rates is an unintended consequence of inspections, which is sustained only in local authorities judged to be performing badly. Implications are discussed for performance in the sector and the purpose of inspection.

## Keywords:

Children's social care, child welfare, Ofsted, inspection, performance measures, quantitative analysis

This is the authors' accepted manuscript. Suggested citation:  
Hood, R. and Goldacre, A. (in press) Exploring the impact of Ofsted inspections on performance in children's social care. *Children and Youth Services Review*.

## Introduction

Many countries have seen an increased focus on state monitoring of child welfare provision, which may reflect the sector's political sensitivity in the context of gradual marketisation and privatisation in some areas (Jones, 2018). Regulation varies between jurisdictions, ranging from the licensing of residential care providers in Sweden (Pålsson and Shanks, 2020) to the wholesale inspection of social care services in the constituent countries of the United Kingdom (Lundy et al., 2019; Bach-Mortensen and Montgomery, 2020). In England, the Office for Standards in Education, Children's Services and Skills (Ofsted) is a non-ministerial government department that inspects services providing care for children and young people, and education and training for learners of all ages. Originally established to inspect schools, Ofsted's remit was expanded in 2007 to include children's services, as well as social care providers, for example children's homes and fostering agencies. While the latter are increasingly found in the private sector, the bulk of statutory children's services are provided by municipal governments, which in England are called local authorities (LAs). Such services include the assessment of need, care plans for children judged to be 'in need' of support, protective interventions and investigations, court work, and acting as corporate parent for children in public care. The current convention is to bracket the LA services together as 'children's social care' (CSC). A detailed description of the CSC system in England is provided by Own Author (2020).

Ofsted usually takes around three years to inspect all LAs in the country, although political and regulatory pressures have sometimes resulted in more frequent changes to inspection frameworks (see Table 1). The current framework (Ofsted, 2017b) is called the 'Inspection of Local Authority Children's Services' (ILACS) and has been in operation since January 2018. All inspections result in an overall judgement of effectiveness on a four-point scale: 1 – Outstanding, 2 – Good, 3 – Requires Improvement to be Good, and 4 – Inadequate. Judgements on the same scale are also given for sub-categories, which broadly distinguish between child protection services and provision for looked after children and care leavers. Inspection reports for all LAs and providers, including the judgements given, are publicly available on the Ofsted website (<https://reports.ofsted.gov.uk/>), along with guidance notes and annual reports on performance in the sector.

The political sensitivity of child protection (CP) has magnified the importance of Ofsted inspections. The government has taken a keen interest in the sector's performance, which over the years has been impacted by media scandals following deaths from child abuse (Jones, 2014) and the recommendations of public inquiries (Munro, 2004). Penalties for an inadequate judgement are correspondingly high. Alongside rigorous performance improvement measures, often accompanied

This is the authors' accepted manuscript. Suggested citation:

Hood, R. and Goldacre, A. (in press) Exploring the impact of Ofsted inspections on performance in children's social care. *Children and Youth Services Review*.

by managerial and organisational churn (Jones, 2015), some LAs have been effectively stripped of operational control over CSC after being judged inadequate. Nor has the inspectorate itself been immune to accountability pressures, notably during the crisis that engulfed the sector following the tragic death of Peter Connelly in Haringey (Jones, 2014), and a few years later with exposure of the disastrous handling of concerns about child sexual exploitation (CSE) in Rotherham (Jay, 2014). These scandals raised questions about the link between inspection judgements and performance, as discussed below. A parliamentary committee in 2015 questioned the procedural focus of Ofsted inspections and the lack of direct examination of social work practice (House of Commons Education and Skills Committee, 2007; see also Craven and Tooley, 2016). Such pressures add to an already difficult remit for the inspectorate and highlight the challenge of assessing performance in such a complex field of work.

## Inspections and performance

When it comes to gauging the standard of CSC services, the only nationally comparable alternative to Ofsted judgements are LAs' own administrative data, which are reported annually to the government (Own Author, 2019; Wilkins and Antonopoulou, 2019). These annual returns consist of process measures, such as assessment timescales, 'negative' quality measures, such as rates of re-referrals within 12 months, and workforce measures, such as rates of agency workers (Own Author et al., 2016). Ofsted also examines these statistics as part of the inspection process, although they are not the principal means through which judgements are formed and there is no linear relationship between them and inspection outcomes (La Valle *et al.*, 2016; Wilkins and Antonopoulou, 2019). Nonetheless, some performance measures seem to be important markers of quality, in the sense of being reinforced by inspection outcomes. For example, Own author et al. (2016) found that a combination of (less) timely completion of initial assessments, (higher) rates of re-referrals, and (higher) agency worker rates was enough to predict an inadequate judgement, although these measures are not formal criteria for such a judgement. Wilkins and Antonopolou (2019) found that an increase in LA average deprivation and a higher rate of 'overdue' assessments were associated with a lower likelihood of a 'good' or 'outstanding' judgement. Webb and Bywaters (2017) analysed the relationship between deprivation, expenditure and quality, finding that among high- or very high-deprivation LAs, those awarded a 'good' or 'outstanding' judgement spent on average 20% more per child than those awarded 'inadequate' or 'requires improvement to be good'. Meanwhile, Craven and Tooley (2016) found that LAs that improved their Ofsted judgement between one inspection and the next tended to have a smaller increase in children protection plans during the intervening period than LAs whose judgement got worse.

This is the authors' accepted manuscript. Suggested citation:

Hood, R. and Goldacre, A. (in press) Exploring the impact of Ofsted inspections on performance in children's social care. *Children and Youth Services Review*.

The political salience of risk perceptions in child protection co-exists somewhat uneasily with notions of quality and effectiveness. When CSC services in Rotherham were inspected in 2012, for example, they were judged as 'adequate' (judgement 3). Two years later, following a public scandal about the LA's failings in relation to CSE, those services were judged to be inadequate. From the regulator's perspective, there is an institutional risk (Rothstein, 2006) in appearing to give a positive judgement to services later found to be failing in some way, or even to be 'tolerating mediocrity' (Ofsted, 2013). In line with changes to school inspections, the Single Inspection Framework (SIF) for CSC, which started in 2013, saw the 'Adequate' judgement re-designated as 'Requires improvement' (later amended to 'Requires Improvement to be Good'). This meant that over the course of the whole SIF cycle, nearly two thirds of LAs were found to be either inadequate or as requiring improvement (to be good). They are quite distinct judgements, of course; nonetheless, senior leaders saw such results as presenting an unduly negative picture of provision in the sector (ADCS, 2015). The next framework, the ILACS, retained a similar methodology to the SIF within a 'proportionate, risk-based approach to inspection' (Ofsted, 2017: 6). In the first two years of the ILACS framework over 50% of LAs received either a 'good' or 'outstanding' judgement, a significant picture of improvement compared to earlier cycles. The comparison is shown in Table 1 below, which shows a breakdown of judgements under each inspection framework from 2009-19. During this period, about 43% of all LAs (66 out of 152) have been judged inadequate at least once in terms of their *overall effectiveness*, while only 17% (26 out of 152) have remained good or outstanding throughout.

*Table 1. Ofsted inspection frameworks and overall judgements 2009-19 (compiled from Ofsted, 2019b)*

Fluctuations in outcomes are to be expected in different inspection cycles, as is a robust response to emerging risks and the exposure of systemic failures. However, socio-economic and political realities inevitably intrude into technocratic notions of quality. Among the 'austerity' measures undertaken by central government to reduce public expenditure have been large-scale cuts to local government budgets, which have been borne disproportionately by more deprived LAs (Webb and Bywaters, 2018). This has made it harder for more deprived LAs to deliver the same standard of provision as less deprived areas, which is reflected in their Ofsted judgements (Wilkins and Antonopoulou, 2019). The context of demand has added further pressure, with a steep rise in the use of protective interventions relative to the level of referrals (Own Author et al., 2019). In its annual reports, Ofsted has acknowledged the impact of LA deprivation on the chances of a good inspection judgement (Ofsted, 2017a), although it has disputed the link with spending. The inspectorate has also

This is the authors' accepted manuscript. Suggested citation:

Hood, R. and Goldacre, A. (in press) Exploring the impact of Ofsted inspections on performance in children's social care. *Children and Youth Services Review*.

recognised the pressures caused by scandals about child abuse and (more recently) child sexual exploitation, which have fuelled 'public perception that LAs are failing children' (Ofsted, 2018: 89). Part of the challenge for Ofsted is that it must position itself as an objective arbiter of effectiveness and improvement at the same time as being itself an actor in a highly challenging regulatory environment. As an inspectorate, Ofsted operates in a public sphere increasingly dominated by risks and their management (Power, 2004). In this context, institutional anxiety about reputational damage, not least that inflicted by a poor inspection rating, has the potential to affect the performance of services.

### Institutional anxiety and the 'regulatory logic' of risk

Rothstein et al. (2006) argue that governance in the early 21st century is characterised by the regulation of risk. Efforts to address 'societal risks', i.e. potential harm to members of the public and the environment, have given rise to a multitude of regulatory bodies – including government agencies, consumer organisations and the media – that in turn seek out further risks that need to be addressed. When it comes to child safeguarding, for example, LA services are tasked with monitoring private family life to protect children from harm; these agencies monitor their own performance via audit and quality assurance processes but are also inspected by Ofsted, which as we have seen may have its own activities scrutinised by parliamentary committee (and by the media) in the event of a child abuse scandal. The proliferation of regulatory activity has magnified what Rothstein et al. (2006) call 'institutional risks', which they define as 'risks to organizations (state or non-state) regulating and managing societal risks, and/or risks to the legitimacy of their associated rules and methods' (Rothstein et al., 2006, p. 92). Institutional risks are particularly acute for regulators operating within constraints that significantly limit their control over societal risks. Most of these constraints are evident in CSC: high levels of unpredictability and uncertainty, rising demand combined with severe budgetary pressures, complex multi-agency processes, conflicting priorities and interests, and problem areas that do not always map onto existing frameworks for intervention. The combination of regulatory scrutiny, severe consequences for failure, and operational constraints on performance, are likely to provoke anxiety amongst professionals and within institutions (Menzies-Lyth, 1990; Woodhouse and Pengelly, 1991).

With agencies under growing pressure to justify their actions in the face of political, executive, and judicial scrutiny, more evidence is needed about the relationship between inspection and performance. The empirical work described below seeks to examine some of the dynamics of this relationship. Its aim was to identify the impact of Ofsted inspections on a range of performance

This is the authors' accepted manuscript. Suggested citation:  
Hood, R. and Goldacre, A. (in press) Exploring the impact of Ofsted inspections on performance in children's social care. *Children and Youth Services Review*.

measures for CSC, focusing on a completed cycle of inspections under the Single Inspection Framework (SIF). The research questions were:

1. What is the effect of an inspection on intervention rates and other performance measures, such as expenditure and workforce stability?
2. Do any identified effects vary according to the inspection outcome?

## Methods

To address the research questions, national returns for 152 English LAs were combined using statistical tables that are freely available and published online by the government (see Own Author et al., 2016a and 2019 for details). These tables consist of aggregated measures of demand and provision, such as referrals, assessments, children in need, child protection (CP) plans and care orders, along with other indicators such as timescales for completion, case closures, re-referrals, expenditure, caseloads, staff turnover and vacancies (see Own Author et al., 2020). Most comparable child welfare systems have similar measures to these (Gilbert, 2012). However, one category that is specific to the English system is a 'child in need' (CIN), defined by the Children Act 1989 as a child who is unlikely to reach or maintain a satisfactory level of health or development, or whose health or development will be significantly impaired without the provision of children's social care services, or a child who is disabled. Around 64% of referrals lead to a child being assessed as 'in need' and therefore eligible to receive statutory CSC services (Department for Education, 2020).

The administrative data described above were supplemented with a summary of inspection outcomes from 2009-19 (Ofsted, 2019). All years of expenditure were adjusted for inflation using the 2019 GDP deflator index published by the government (HM Treasury, 2019). A table of indicators for CSC and their sources can be found in the online supplementary material to this article. Following data cleaning, two LAs with very small numbers of children (Isles of Scilly and the City of London) were excluded from the analysis, leaving a population of 150 LAs. For this study it was important not only to be able to compare trends in performance measures in the period before and after an inspection, but also to have a reliable basis for comparing judgements between different LAs. For this reason, it was decided to focus on outcomes from the first inspection under the SIF, the only fully completed cycle of inspections for which it was possible to explore relevant trends data from the CIN Census and other indicators that were mostly introduced after 2010 (see Table 1 for a breakdown of judgements under the SIF). In line with previous work (Own Author et al., 2016; Wilkins and Antonopolou, 2019), Overall Effectiveness (OE) was chosen as the outcome variable. OE

This is the authors' accepted manuscript. Suggested citation:  
Hood, R. and Goldacre, A. (in press) Exploring the impact of Ofsted inspections on performance in children's social care. *Children and Youth Services Review*.

is an overall judgement based on the main areas of inspection: children who need help and protection; children looked after and achieving permanence; and leadership, management and governance. There is a very high correspondence between the OE judgement and these sub-categories (Wilkins and Antonopoulou, 2019), which means that results from statistical analysis based on either would be very similar.

Interrupted time series (ITS) analysis using multilevel segmented regression was chosen to answer the research questions. Firstly, time trends and annual percentage changes were modelled using multilevel Poisson regression, assuming a constant rate of change. Following this, the time series was segmented to identify step changes and slope changes (Bernal *et al.*, 2017). We chose the impact model based on i) the patterns found in the analysis of median rates, and ii) the fit statistics including the Bayesian Information Criterion and Akaike Information Criterion scores, which were lowest for the model. In order to carry out the model we categorised year into a variable which enabled the comparison of four different stages: annual change pre-inspection, step change at the year of inspection, step change the year after inspection and annual change post-inspection. The analysis was carried out within a multilevel framework in order to adjust the analysis for LA-membership; this limits the measurement of both within-LA variation (variation over time) and between-LA variation. Since time will have a different effect depending on which local authority it is controlled for by including it as a random effect. The results are presented graphically as % change over a seven-year period: 3 years prior to inspection to 3 years after inspection. The % changes were based on the model parameter estimates (i.e. the predicted multiplicative change in rates) for the four different stages (included in the appendix). The multilevel ITS analysis was firstly carried out for all inspections regardless of outcome (n=150), and then stratified by inspection outcome: 1 – Outstanding or Good (n=52), 2 – Requires Improvement to be Good (n=65), and 3 – Inadequate (n=33). Outstanding and Good were combined into a single group due to only three LAs being rated Outstanding. The analysis was carried out using Stata version 16.0.

## Findings

### Analysis of median rates

A longitudinal analysis of median rates of intervention is shown in Figure 1, centred on the year of SIF inspection (Y0) and covering a three-year period before (Y-3 to Y-1) and afterwards (Y1 – Y3).



This is the authors' accepted manuscript. Suggested citation:

Hood, R. and Goldacre, A. (in press) Exploring the impact of Ofsted inspections on performance in children's social care. *Children and Youth Services Review*.

Since LAs were inspected under the SIF between 2013 and 2017, these results include administrative data from 2010 – 2019. The shaded area on each graph highlights changes in the time series from Y-1 to Y1. The graphs show a general upward trend in rates of intervention over the whole period, particularly evident for child protection (CP) interventions. They also suggest a discontinuity around the time of an inspection, characterised by a steep rise during Y0 followed by a fall in Y1. This 'spike' in activity was observed for all intervention measures except for care orders and was particularly noticeable for Section 47 enquiries, CP conferences, and CP plans. The validity and significance of the observed pattern were tested through an interrupted time series (ITS) analysis, as reported below.

*Figure 1. Time series for median rates of intervention (centred on inspection year)*

### Interrupted time series (ITS) analysis

The ITS analysis was first carried out for all inspections regardless of outcome. The model can be described as a slope with two step changes, which was best suited to examining the discontinuities observed in Figure 1. The modelled % change along with 95% confidence intervals (CI), for 24 different performance measures, are shown in Table 2. The results are presented for four stages: annual change pre-inspection, step change at the year of inspection, step change the year after inspection and annual change post-inspection. Next, the predicted percentage changes were converted into changes year-on-year over a seven-year period: 3 years prior to inspection to 3 years after inspection, which made the results easier to interpret and visualise. Figure 2 is a visualisation of the ITS analysis for rates of service provision, centred on year of inspection. The shaded column in the middle highlights the two step changes sandwiched between pre- and post-inspection trends. The percentage changes show a rise or fall in from the baseline (100%) in each stage of the analysis. The results relating to service provision show evidence for the following discontinuities:

- a. A rise during the year of inspection, which is higher than the pre-inspection trend.
- b. A fall the year after inspection, which follows a rise the previous year.
- c. A post-inspection trend that is the same direction but less steep than the pre-inspection trend
- d. A post-inspection trend that changes direction from pre-inspection trend

A combination of a) and b) represents the best evidence for a 'spike' in values associated with an inspection. Evidence of a subsequent 'levelling off' in either an upwards or downwards trend is represented by c). Evidence of a reversal in the direction of a trend post-inspection is represented by d).

This is the authors' accepted manuscript. Suggested citation:  
Hood, R. and Goldacre, A. (in press) Exploring the impact of Ofsted inspections on performance in children's social care. *Children and Youth Services Review*.

*Table 2: ITS analysis results for CSC indicators*

The annotated results in Table 2 show evidence of inspection-related discontinuities for a range of indicators. These table show, for each measure, the model parameter estimates for each segment. The predicted rate ratios are converted into percentages for ease of interpretation. The 95% CIs tell us how precise the estimates are; a narrower CI indicates a more precise estimate, whilst a wider CI indicates a less precise estimate. The CIs are used to infer differences between the estimates in the time-series (for each performance indicator) by examining the lower and upper limits and checking for overlapping and non-overlapping segments. The ITS model fitted particularly well for rates of service provision, most of which were characterised by an upward trend with a spike around the year of inspection followed by a levelling off in the rate of increase. Rates of children in need (CIN) saw a spike followed by a downward trend post-inspection. Rates of children accommodated in care under Section 20 of the 1989 Children Act (a voluntary arrangement that does not require a court order) were an exception in this category, in that the rate of increase slowed during the year of inspection before dipping the year afterwards and persisting in a negative trend. It was noticeable that CP indicators (investigations, conferences and plans), which had the biggest annual increases, also saw the most significant spike around inspection. There was no evidence of a spike in care orders, although there seemed to be a levelling-off post-inspection. The findings on rates of provision are illustrated in Figure 2 below.

*Figure 2. ITS analysis for rates of service provision, centred on year of inspection.*

Quality indicators for service provision in Table 2 include the proportion of CIN and CP plans that were closed quickly (within three and six months respectively), timescales for completion of assessments and CP conferences, and the proportion of re-referrals and repeat CP plans. The effect of an inspection was less consistent for these indicators. A spike was evident for CIN plans ceasing within three months (the shortest reported time period), which otherwise trended downwards pre- and post-inspection, while CP plans ceasing within six months showed a dip the year after inspection before resuming an upward trend. Timescales for completion seemed to go from an upward to a downward trajectory following inspection. While there no significant inspection-related

This is the authors' accepted manuscript. Suggested citation:  
Hood, R. and Goldacre, A. (in press) Exploring the impact of Ofsted inspections on performance in children's social care. *Children and Youth Services Review*.

discontinuities in re-referrals, the proportion of repeat CP plans saw a two-year dip around inspection before going up again.

Spend on children's safety and spend on Early Help saw a spike. Spend on CLA saw a significant fall during the year of inspection and the year afterwards, despite the general upwards trend for this category pre- and post-inspection. The spike in spend on Early Help was short-lived, coming in the context of an otherwise steep decline over this period. Finally, workforce indicators suggested that inspections were associated with a spike in CIN per social worker and in vacancy rate, as well as a large, two-year increase in agency worker rates. Vacancy rates and agency worker rates in the post-inspection period then went into decline, contrasting with an upward trend pre-inspection. CIN per social worker, having been falling pre-inspection, stabilised following a sharp dip in the year after inspection.

### ITS analysis stratified by inspection outcome

Stratifying by inspection outcome revealed a more nuanced picture. Results for inadequate LAs often differed from those for LAs receiving other judgements. A supplementary table showing the coefficients and confidence intervals for the stratified analysis accompanies the online version of this paper. Key findings for each category of indicator are summarised here.

In the service provision category, the spike in intervention was characterised by a much steeper rise in the year of inspection for Inadequate LAs than in LAs with other judgements. Whereas the general trend was for rates to go down the year after inspection, the Inadequate LAs tended to diverge from other LAs; CIN rates increased the year after the inspection (before declining gradually in subsequent years), whilst CP Investigations and CP Plans were sustained at higher levels several years after inspection (rather than declining or levelling off), as illustrated in Figure 3. Levels of provision for LAs rated good or outstanding showed little discontinuity from pre-inspection trends during the year of inspection but then tended to dip temporarily the year after the inspection. A similar pattern emerged for LAs rated 'requires improvement to be good'. Care orders were an exception in that good or outstanding rated LAs diverged from the others, falling in both years around inspection before rising again in subsequent years.

*Figure 3. ITS analysis for provision of CIN and CP plans (by inspection outcome)*

This is the authors' accepted manuscript. Suggested citation:

Hood, R. and Goldacre, A. (in press) Exploring the impact of Ofsted inspections on performance in children's social care. *Children and Youth Services Review*.

Indicators in the 'period of service' category suggested that Inadequate LAs experienced a significant rise in CIN Cease at 3 months the year after an inspection, whereas other LAs experienced a rise during the year of inspection. For CP plans that ceased within 6 months, there was a steep rise in figures for Inadequate LAs both pre-inspection as well as in the year of inspection, followed by a dip the year afterwards, which again diverged from the pattern in other LAs. In the timescales category, Inadequate LAs tended to undertake less timely assessments and CP conferences during the year of inspection, but more timely the year after inspection, something that was particularly evident for CP conferences. Good or outstanding LAs were also less timely with CP conferences during the year of inspection, while LAs requiring improvement saw an increase. The indicator for re-referrals within 12 months went up for good and outstanding LAs during the year of inspection but down for other LAs, whereas Inadequate LAs had the opposite pattern with a decline in re-referrals at inspection followed by a steady rise in subsequent years. Results for repeat CP Plans were not as conclusive, with wide confidence intervals due to small numbers, but suggested that rates of repeat plans went down in good and outstanding LAs during the year of inspection, and went down in Inadequate LAs the year after inspection. This was in the context of a general upwards trend in repeat CP plans for all LAs over the period.

As illustrated in Figure 4, workforce and expenditure indicators also showed a divergence between Inadequate LAs and other LAs. Vacancy rates and agency worker rates in Inadequate LAs rose steeply during the year of inspection and the year afterwards before declining in subsequent years. In contrast, good and outstanding LAs saw a fall in the use of agency workers around the year of inspection. Rates of CIN per social worker saw little change, except for Inadequate LAs, who experienced a spike in this indicator during the year of inspection. Expenditure figures told a similar story, with total spend on children and young people's services going down in good or outstanding LAs but up in Inadequate LAs, while spending on children's safety also saw much bigger increases in Inadequate LAs than elsewhere (see Figure 4). The dip in CLA expenditure was also less pronounced in Inadequate LAs due to an increase the year after inspection. Inadequate LAs saw a two-year rise in spending on Early Help around inspection but this was then followed by a steeper decrease than in other LAs.

*Figure 4. ITS analysis for workforce and expenditure indicators (by inspection outcome)*

## **Summary of findings**

This is the authors' accepted manuscript. Suggested citation:  
Hood, R. and Goldacre, A. (in press) Exploring the impact of Ofsted inspections on performance in children's social care. *Children and Youth Services Review*.

- The analysis examined the impact of Ofsted inspections on the performance of CSC services.
- Analysis of median rates centred on the year of inspection suggested that rates of most types of intervention rose during the year of an inspection and then fell the year afterwards.
- An interrupted time series (ITS) analysis for all LAs confirmed evidence of a general spike in provision, particularly in CP interventions.
- When results were broken down by inspection outcome, it was found that increases in provision during the year of inspection were mostly experienced by inadequate rated LAs (Inadequate LAs). Rates of CP interventions for Inadequate LAs also continued to rise in the years following inspection.
- LAs that were rated good or outstanding generally did not have a rise in rates of provision during the year of inspection but did experience a dip the year afterwards. A similar pattern was seen for LAs rated as 'requires improvement to be good'.
- Workforce indicators also showed Inadequate LAs diverging from other LAs, with vacancies and use of agency workers in Inadequate LAs rising steeply during the year of inspection and the year afterwards before declining in subsequent years.
- In contrast, good and outstanding LAs saw a drop in agency worker rates around the time of inspection.
- Expenditure figures told a similar story, with total spend on children and young people's services going down in good or outstanding LAs but up in Inadequate LAs, while spending on children's safety also saw much bigger increases in Inadequate LAs than elsewhere

## Discussion

The findings show that Ofsted inspections do have an impact on performance measures in CSC services, which is generally experienced during the year of inspection and is significantly different from underlying trends. This impact is particularly pronounced for LAs receiving an inadequate judgement and seems to lead to a more sustained effect in those areas. However, it is necessary to bear in mind some limitations when interpreting these results. First, administrative data are not a particularly good measure of outcomes for children (see Own Author, 2019a). Accordingly, there is little consensus about what constitutes an optimal level of child welfare provision, particularly given the impact of inequalities (Bywaters *et al.*, 2020) and an overall shift to late intervention (Action for Children *et al.*, 2017). Second, the data considered for the study covers a nine-year period centred on the Single Inspection Framework; different effects may be observed when more data becomes

This is the authors' accepted manuscript. Suggested citation:

Hood, R. and Goldacre, A. (in press) Exploring the impact of Ofsted inspections on performance in children's social care. *Children and Youth Services Review*.

available, and when it becomes possible to replicate the analysis using outcomes under the current framework, the ILACS. Third, the relationship between Ofsted inspections and performance measures may be an indirect, or even unintended, consequence of implementing recommendations in Ofsted reports, and will be mediated by other factors, such as external influences on demand, policy changes, and the impact of judgements in the higher courts. With these limitations in mind, the discussion centres on three key questions: 1) why might inspections be associated with a spike in provision and other indicators for Inadequate LAs, 2) why might LAs with other judgements see a dip in many of the same indicators the year following inspection, and 3) are these effects desirable and consistent with the aims of inspection?

To address the first question, a spike in provision for Inadequate LAs could be explained by measures being taken to improve services in response to feedback from inspectors. For example, thresholds for intervention would be made more robust if there were concerns that children who required statutory services were not consistently receiving them. An increase in CP provision might also be deemed necessary to safeguard children while improvements are made to referral and assessment and other 'front door' processes. It could therefore be hypothesised that a spike reflects the sudden incorporation of demand that was previously – and inappropriately – being assessed as below the threshold for CSC. Relevant to this hypothesis is a recent analysis of all-England data for CSC services, which showed that LAs with high levels of demand tend to i) experience higher levels of deprivation ii) screen out more cases at the stages of referral and assessment, iii) spend less on children's safety (per CIN child) and on children in care (per CLA child) and iv) experience higher rates of re-referrals and repeat CP Plans (Own Author et al., 2020). High IMD scores were found by Wilkins and Antonopoulou (2019) to lower the likelihood of receiving a good or outstanding inspection judgement. In other words, more deprived LAs are likely to have fewer resources relative to the level of demand *and* to do less well in inspections, which makes them good candidates for tightening up thresholds in the wake of a critical report.

However, there are reasons to question this hypothesis. A recent thematic analysis of Ofsted reports (Own Author et al., 2019b), including from the SIF cycle, found that recommendations around casework focused mainly on timely assessment and intervention, consistent standards of recording and analysis, and outcomes-focused care plans. In relation to thresholds, LAs were generally recommended to apply these consistently and appropriately share information about risk between agencies. There is nothing to suggest that Inadequate LAs were generally being advised to lower their thresholds or increase intervention rates; indeed Craven and Tooley (2016) found LAs with a

This is the authors' accepted manuscript. Suggested citation:

Hood, R. and Goldacre, A. (in press) Exploring the impact of Ofsted inspections on performance in children's social care. *Children and Youth Services Review*.

more limited rise in CP interventions stood a better chance of improving their inspection outcome next time. Results from this study did show that the timeliness of assessments and CP conferences improved for Inadequate LAs in the year following inspection, and it is plausible to connect such improvements with inspection recommendations. However, timeliness itself has been shown to be unconnected to rates of intervention (Own Author et al., 2020). When it comes to an immediate spike in CIN and CP interventions for Inadequate LAs, it would therefore seem that factors other than (or additional to) Ofsted recommendations were involved.

One of the key contextual factors for Inadequate LAs is workforce instability. High rates of agency workers have previously been found to increase the chances of an inadequate judgement (Own Author et al., 2016), while in this study rates of agency workers (and associated expenditure) were found to climb steeply in Inadequate LAs both in the year of inspection and the year afterwards, from the already high levels beforehand. This resonates with qualitative evidence about the negative impact of an inadequate inspection on workforce morale, recruitment and turnover (Jones, 2015; Own Author, 2016b, 2020a) but again is not linked (at least directly) with recommendations made by Ofsted reports. The picture of Inadequate LAs that emerges is of organisations that prior to inspection seem to have been experiencing difficulties in terms of demand relative to resources, managerial grip on processes, and an unstable workforce. Such problems are likely to affect how the organisation responds to a period of intense scrutiny, contributing to heightened levels of anxiety among practitioners and managers (Ferguson *et al.*, 2019) both during the inspection itself and after the outcome is known.

From an organisational perspective, receiving (or being likely to receive) an inadequate judgement can be equated with a high level of institutional risk (Rothstein et al., 2016). This means that Inadequate LAs are likely to focus their attention on achieving and demonstrating a picture of performance that will meet the regulator's requirements and therefore lead to a better judgement. It is not an easy journey; for example, the ITS analysis results suggest that even stabilising the rate of agency workers at pre-inspection levels can take a number of years. In contrast, LAs that are rated good or outstanding seem to benefit from a reduction in institutional risk, stemming from an enhanced reputation and its knock-on effects on recruitment, workforce stability, morale and relationships with partner agencies (see Own Author et al., 2020). This may help to explain the dip in rates of service the year following inspection, which highly-rated LAs experienced even in the context of a steep upwards trend in CP interventions over the period in question. Another explanation for the dip is the converse of the screening effect noted earlier in relation to high-

This is the authors' accepted manuscript. Suggested citation:

Hood, R. and Goldacre, A. (in press) Exploring the impact of Ofsted inspections on performance in children's social care. *Children and Youth Services Review*.

deprivation high-demand LAs. In other words, low deprivation LAs are both more likely to receive a good or outstanding inspection outcome *and* have more resources relative to the level of demand; they tend to be more interventionist and so have more scope to adjust thresholds upwards.

The final question is whether the effects discussed above are desirable and consistent with the aims of inspection. It is hard to give a conclusive answer to this, given that the CSC performance measures are hard to align with outcomes for children (La Valle *et al.*, 2016), or even with Ofsted judgements themselves (Wilkins and Antonopoulou, 2019). The finding that indicators for Inadequate LAs tended to diverge quite strongly from other LAs, including an improved picture of timeliness the year after inspection, arguably reflects the inspectorate's success in identifying LAs whose institutional context was unusually problematic and where drastic action was needed to drive up standards. More troubling, perhaps, is the sense that other consequences were unintended, particularly the spike in CP interventions and the deteriorating stability of the workforce in Inadequate LAs post-inspection. While the findings are consistent with Ofsted inspectors correctly identifying poorly performing LAs, they provide some support for the view that such LAs are under-resourced to meet demand for services, particularly in very deprived areas that have been disproportionately affected by government cuts (Webb and Bywaters, 2017). Indeed, the last decade of Ofsted inspections has coincided not only with huge fiscal retrenchment, but also a steep rise in use of CP interventions. Arguably missing from this picture is the importance of leadership, which Ofsted considers to be the crucial factor enabling services in highly deprived LAs to perform well (Ofsted, 2015). Nonetheless, the findings suggest that the post-inspection trajectory for Inadequate LAs may replicate and even worsen some of the conditions that get them into trouble in the first place (Own author *et al.*, 2016b).

Ofsted's current inspection framework, the ILACS, replaced the SIF in 2018. The guidance (Ofsted, 2017) set out a change of direction away from a standard four-week inspection and towards a more flexible and proportionate system. At its heart is a 'tailored inspection menu' based on three 'inspection pathways' designed to adjust Ofsted's engagement with LAs to reflect available intelligence about their performance, including the last inspection report. The approach seems aligned with responsive and risk-based theories of regulation (Ayres and Braithwaite, 1992; Black and Baldwin, 2010), although this literature is not cited directly. The guidance also states the intention to focus less on process and more on direct social work practice with families. Compared with the SIF, the new framework seems to encourage less adversarial relationships with LAs, pay more attention on professional expertise rather than procedural compliance, and gear accountability



This is the authors' accepted manuscript. Suggested citation:

Hood, R. and Goldacre, A. (in press) Exploring the impact of Ofsted inspections on performance in children's social care. *Children and Youth Services Review*.

towards the shared goal of positive outcomes for children. Inspection outcomes under the first two years of the ILACS showed a marked improvement compared to the SIF, with three times as many outstanding judgements awarded than in the whole of the previous cycle (see Table 1).

Given these developments, could the way organisations experience and respond to inspections be changing under the ILACS? A recent review of ILCAS implementation (Ofsted, 2019a) included an independent report by Ferguson *et al.* (2019), which discussed feedback from agencies on changes such as short inspections, observations of direct practice, and the awarding of performance grades. It described agencies as being preoccupied with the outcome of an inspection: 'the grade Ofsted issues, the perceived meaning of this label, and its potential consequences, resulted in high levels of anxiety for staff at all levels in the LAs' (2019: 58). Short inspections, introduced with the ILACS, were particularly stressful 'when there is evidence that suggests a service has deteriorated and no longer appears to be Good'. Although these were preliminary findings from an early stage of implementation, they suggest that pressure to manage institutional risk is likely to persist under the ILACS, which could have consequences for performance as explored in this paper.

## Conclusion

This paper has presented evidence that rates of statutory provision undergo a short-term spike during the year of an inspection. The spike was found to be particularly pronounced for CP interventions in LAs that received an inadequate judgement. These LAs also experienced a rise in safeguarding expenditure and rates of agency workers. In contrast, LAs receiving a good or outstanding judgement benefited from a dip in intervention rates in the year after an inspection and a more stable workforce. The divergence in indicators between inadequate-rated LAs and other LAs may be attributed partly to pre-existing problems, partly to actions undertaken to improve performance, and partly to unintended consequences. The effect of institutional risk perceptions on workforce stability and confidence among partner agencies, combined with the unequal distribution of resources relative to demand, may hinder the ability of LAs to recover from an inadequate judgement. While the current ILACS framework has sought to develop a more proportionate approach to inspection, the institutional context remains one of rising demand, budgetary constraints, and anxiety about regulatory failure. With demand for statutory services set to rise in the wake of the Covid-19 pandemic, and financial constraints likely to tighten further, the sector continues to face a highly challenging regulatory context.

This is the authors' accepted manuscript. Suggested citation:  
Hood, R. and Goldacre, A. (in press) Exploring the impact of Ofsted inspections on performance in children's social care. *Children and Youth Services Review*.

This is the authors' accepted manuscript. Suggested citation:

Hood, R. and Goldacre, A. (in press) Exploring the impact of Ofsted inspections on performance in children's social care. *Children and Youth Services Review*.

## References

- Action for Children, National Children's Bureau and The Children's Society (2017) *Turning the Tide*, available online: <https://www.childrensociety.org.uk/what-we-do/resources-and-publications/turning-the-tide>, Last accessed Last accessed 26 July 2018
- Association of Directors of Children's Services (ADCS) (2015) *Comment on the Future of Children's Services Inspection*, available online: <http://adcs.org.uk/inspection/article/comment-on-the-future-of-childrens-services-inspection>, Last accessed
- Ayres, I. and Braithwaite, J. (1992) *Responsive regulation: Transcending the deregulation debate*, Oxford University Press on Demand.
- Bernal, J.L., Cummins, S. and Gasparrini, A. (2017) 'Interrupted time series regression for the evaluation of public health interventions: a tutorial', *International journal of epidemiology*, **46**(1), pp. 348-355.
- Black, J. and Baldwin, R. (2010) 'Really responsive risk-based regulation', *Law & policy*, **32**(2), pp. 181-213.
- Bywaters, P., Scourfield, J., Jones, C., Sparks, T., Elliott, M., Hooper, J., McCartan, C., Shapira, M., Bunting, L. and Daniel, B. (2020) 'Child welfare inequalities in the four nations of the UK', *Journal of Social Work*, **20**(2), pp. 193-215.
- Craven, B.M. and Tooley, J.N. (2016) 'Safeguarding Children: Ofsted and Regulatory Failure', *Economic Affairs*, **36**(1), pp. 64-79.
- Ferguson, H., Gibson, M. and Plumridge, G. (2019) 'Appendix 2: Independent Evaluation of the Implementation of Ofsted's Framework for Inspection of Local Authority Children's Services (ILACS)', in Ofsted (ed), *Inspection of local authority children's services framework implementation review*, available online: <http://www.gov.uk/government/organisations/ofsted>, pp.
- Gilbert, N. (2012). 'A comparative study of child welfare systems: Abstract orientations and concrete results', *Children and Youth Services Review*, **34**(3), 532-536.
- Jay, A. (2014) 'Independent Inquiry into Child Sexual Exploitation in Rotherham ', Rotherham, Rotherham Metropolitan Borough Council.
- Jones, R. (2014) *The story of Baby P: Setting the record straight*, Bristol, Policy Press.
- Jones, R. (2015) 'Ofsted doesn't recognise the practice chaos and professional carnage it leaves in its wake', *Community Care*.
- La Valle, I., Holmes, L., Gill, C., Brown, R., Hart, D. and Barnard, M. (2016) 'Improving Children's Social Care Services: Results of a feasibility study', London, CAMHS Press.
- Munro, E. (2004) 'The impact of child abuse inquiries since 1990', in Stanley, N. and Manthorpe, J. (eds), *The Age of the Inquiry: Learning and Blaming in Health and Social Care*, London, Routledge, pp.
- Ofsted (2015) *Joining the dots... Effective leadership of children's services*, available online: [www.gov.uk/government/organisations/ofsted](http://www.gov.uk/government/organisations/ofsted)
- Ofsted (2017a) *Annual report into children's social care 2016-17*, London, Ofsted, available online: [www.gov.uk/government/organisations/ofsted](http://www.gov.uk/government/organisations/ofsted)
- Ofsted (2017b) *Framework, evaluation criteria and inspector guidance for the inspections of local authority children's services*, London, Ofsted, available online: [www.gov.uk/government/organisations/ofsted](http://www.gov.uk/government/organisations/ofsted)
- Ofsted (2018) *Annual report into children's social care 2017-18*, London, Ofsted, available online: [www.gov.uk/government/organisations/ofsted](http://www.gov.uk/government/organisations/ofsted)
- Ofsted (2019a) 'Inspection of local authority children's services framework implementation review', available online: <http://www.gov.uk/government/organisations/ofsted>.
- Ofsted (2019b) *Local authority and children's homes in England inspections and outcomes autumn 2019 char*

This is the authors' accepted manuscript. Suggested citation:

Hood, R. and Goldacre, A. (in press) Exploring the impact of Ofsted inspections on performance in children's social care. *Children and Youth Services Review*.

*ts\_\_tables\_and\_underlying\_data*, London, Ofsted, available online:

<https://www.gov.uk/government/statistics/local-authority-and-childrens-homes-in-england-inspections-and-outcomes-autumn-2019>

Pålsson, D., & Shanks, E. (2020). Missed opportunities? State licencing on the Swedish residential care market. *European Journal of Social Work*, Advance access,

<https://doi.org/10.1080/13691457.2019.1709162>

Power, M. (2004) *The Risk Management of Everything*, London, Demos.

Rothstein, H. (2006) 'The institutional origins of risk: A new agenda for risk research', *Health, Risk and Society*, **8**(3), pp. 215-221.

Webb, C. and Bywaters, P. (2017) 'There is clear evidence that links deprivation, expenditure and quality in children's services', *Community Care*.

Webb, C.J.R. and Bywaters, P. (2018) 'Austerity, rationing and inequity: trends in children's and young peoples' services expenditure in England between 2010 and 2015', *Local Government Studies*, **44**(3), pp. 391-415.

Wilkins, D. and Antonopoulou, V. (2019) 'Ofsted and children's services: what performance indicators and other factors are associated with better inspection results?', *The British Journal of Social Work*, Advance Access, <https://doi.org/10.1093/bjsw/bcy134>