CESifo WORKING PAPERS

8205
2020

April 2020

Children's Social Care and Early Intervention Policy in England

Dan Anderberg, Christina Olympiou

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 8205*

Children's Social Care and Early Intervention Policy in England

Abstract

We study the relationship between a key early intervention policy designed to support families with children up to the age of four and the rate at which children are taken into social care. The gradual build-up of over 3,600 Sure Start Children's Centres (SSCC), operated by Local Authorities across England, created large spatial and cohort variation in the provision of a range of services that include childcare, early education, health and parenting support. Local Authorities are also responsible for the safeguarding of children and about 25 children per 10,000 are taken into social care annually, in the majority of cases to protect them from abuse and neglect. We find that SSCC provision is associated with a higher rate of entry into care for children aged 0-4, but a lower rate of entry for children aged 5-9. The findings are consistent with the policy improving longer-term outcomes while identifying cases in urgent need of care.

JEL-Codes: I380, J120, J130.

Keywords: social care for children, early intervention, abuse and neglect.

Dan Anderberg
Royal Holloway University of London &
Institute for Fiscal Studies / London / UK
dan.anderberg@rhul.ac.uk

Christina Olympiou
Royal Holloway University of
London / UK
christina.olympiou.2016@live.rhul.ac.uk

March 26, 2020
We are grateful to comments from Monica Costa Dias, Malte Sandner, Arnaud Chevalier and seminar participants at Royal Holloway University of London and at the March 2020 London workshop on "Care for Children and Children's Outcomes".

1 Introduction

Over the past decade, more than 30,000 children have been taken into care in England each year, in the majority of cases in order to protect them from abuse and neglect. Looking at a specific cohort, for instance all children born in England in 2001 (who just turned 18), about 3 percent have spent some time in care and many of them have spent a significant part of their childhood in a care placement. While the safeguarding of children in England is the responsibility of the Department for Education who set out policy, legislation and statutory guidance, local safeguarding arrangements are led by Local Authorities (LAs) along with the police. Working together with other relevant agencies, including the National Health Service, they have the responsibility for coordinating and ensuring the effectiveness of work to protect and promote the welfare of children, including identifying and supporting children at risk of harm.

A recent literature in economics has started investigating potential causes and consequences of child maltreatment showing how the incidence varies over time and space (Doyle & Aizer, 2018). For instance, Brown & De Cao (2018) recently studied the effect of unemployment on child maltreatment at county level in the United States, finding that a one percentage point increase in the unemployment rate causes a 20 percent increase in neglect.[1] However, relatively little is known about the role of public policies with respect to parents and children in the process of reducing abuse and neglect of children. Recent work by Sandner & Thomsen (2018) indicates that policies can make a difference: exploiting the rapid expansion of universal child-care provision in Germany between 2002 and 2014 demonstrates how public childcare provision can reduce the incidence of abuse and neglect. Other key policies are designed specifically to support parents and children, promote child and family health, provide advice on parenting, offer financial support etc. in order to reduce tensions and the development of dysfunctional family dynamics.

One such key policy in England is Sure Start, aimed at families with children aged 0 to

[1] For earlier contributions to this literature, see Lindo et al. (2013), Paxson & Waldfogel (1999), Paxson & Waldfogel (2002), Paxson & Waldfogel (2003)

4. Sure Start was initially set up in the late 1990s as relatively small scale Local Programmes in the most deprived areas. However, in 2003 the government announced a long-term plan to transfer Sure Start into the control of local government and to create a Sure Start Children's Centre in every community. During the period 2004-2010 around 3,600 centres were established in England providing a broad range of services including family health services, early years care, early years education, well-being programmes for children and parents and parenting skills support. Since 2010, however, funding has been reduced and about 1 in 7 of the centres that existed in 2010 have closed, leading to a discussion in the media and among practitioners about the impact that this may have had on the care system (see e.g. The Economist, 2018).

However, the relationship between policies and local organizational structures that engage with and support parents and children, on the one hand, and the care system, on the other hand, is not unambiguous. While a primary aim of early intervention is to offer support and thereby reduce the need for statutory care, LAs also face the challenge of identifying children at risk of harm. Studies consistently indicate that only a minority of children who are maltreated are identified. In the United States, the National Incidence Study has consistently found that child protective agencies investigate maltreatment of only a minority of the children who have experienced it (Sedlak et al., 2010). In the UK, research conducted by the Children's Commissioner for England (2015) indicated that it is likely that only 1 in 8 victims of sexual abuse come to the attention of the police and children's services, while the ONS has identified that around one in seven adults who called the National Association for People Abused in Childhood's helpline had not told anyone about their abuse before.[2] In this context, schools have had long-standing legal safeguarding responsibilities, including taking appropriate action and working with services as needed. Hence a second possible impact of early intervention policy is to help identify children in need of statutory care.

The aim of the current paper is to study how the provision of family-based services through Sure Start Children's Centres (SSCC), as a key early intervention policy, has impacted the rate of entry of children into care in England between 2007 and 2017. Over this period, the rate of

[2] See ONS' analysis "Child abuse in England and Wales" here: https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/bulletins/childabuseinenglandandwales/march2020

entry into care has increased steadily, raising questions about the relative roles played by the recession after 2008, the austerity that followed, and by potential behavioural changes of social workers in light of high media profile cases of child abuse.

The gradual build-up of Sure Start centres between 2004 and 2010, along with the subsequent closures created rich geographical and temporal variation in SSCC provision at LA level. We will relate this variation in early intervention provision to the rate of entry by age for children aged 0 to 9. In doing so, we will pay particular attention to the threshold between age 4 and 5 as this delineates the Sure Start target age group (0-4) from schooling age.[3]

We hypothesize two potential effects of SSCC provision, reducing entry into care both in the present and in the future through a family support investment effect, but potentially increasing entry through an identification effect. If both effects are present, SSCC provision should have differential effects on the rate of entry of children within and beyond the target age group. For children who benefited from SSCC provision when in the target age group but who are now aged 5 or above, the effect of SSCC provision should be to reduce entry into care. For children who are currently within the target age group, an identification effect would operate in the opposite direction, thus potentially increasing the rate of entry. Our findings indeed indicate sharply distinct effect on current and past eligible age groups, where we find that increased SSCC provision – as measured by the number of SSCC in an LA per 1,000 children aged 0 to 4 – increases the rate of entry into care for children in that target age group. In contrast, for children aged 5-9 we find that increased SSCC provision when they were of eligible age reduces their current rate of entry into care.

Our analysis thus contribute to the literature by offering insights into the role of early intervention policy for the safeguarding and care for children. The finding that the SSCC provision reduces the rate of entry into care for children who are beyond the target age group is indicative of the policy having a positive investment effect. This finding is in line with Cattan et al. (2019) who recently explored the effect of SSCC provision on children's health outcomes and found that hospital admissions for respiratory illnesses and infections rise in the early years

[3]While there is some flexibility with regards to school starting age, most children start school full-time in the academic year in which they turn five.

of Sure Start access as children are exposed to more pathogens, but subsequently fall as the child develops a stronger immune system. Our finding that SSCC provision increases the rate of entry into care for children within the target age group suggests that early intervention policy may play a role in helping identifying children at risk of harm.

Our paper more broadly contributes to following literatures. First, we contribute to the literature on child abuse and neglect by investigating how a key early intervention policy targeted at young children may affect incidence and identification (e.g. Sandner & Thomsen, 2018; Howard & Brooks-Gunn, 2009; Avellar & Supplee, 2013). Our results also provides further evidence on the potential link to local labour market conditions as we include controls for, for instance, local unemployment rates. Second, we contribute to literatures that explore the causal effects of similar early intervention policies targeting preschool-age children in deprived areas, for instance Head Start in the US (Garces et al., 2002; Carneiro & Ginja, 2014). Third, by looking for longer term benefits of an early intervention policy, we contribute to a rapidly growing literature on child development by exploring whether such a policy can have longer term benefits on child welfare, health and well-being (Currie & Almond, 2011; Conti et al., 2016).

The paper proceeds as follows. Section 2 provides institutional background and some descriptive statistics on the population of looked-after children in England. Section 3 describes further details of the looked-after children data that we will use on the estimation and of additional data on LA characteristics. Section 4 describes the history and evolution of Sure Start and the data on children's centres. Section 5 discusses the potential channels through which SSCC may affect care entry rates by age and also our empirical strategy. Section 6 presents our results, and Section 7 concludes.

2 Looked-After Children in England

In this section we will use publicly available data on population of looked-after children in England going back to 2002, looking at trends and demographic composition in the aggregate and at some variation across LAs.[4]

[4]No data on category of need and ethnicity of children entering care were found for 2002. Analysis for these characteristics starts from 2003.

According to Children Act 1989, a child is defined as "looked-after" by an LA if she or he is provided with accommodation for more than 24 hours, is subject to a care order, or is subject to a placement order.[5] Children's route through care involves decision making on various stages. These decisions are made or managed by social workers in each LA. Social workers judge when to take children into care, assess their needs and the type of placements required, and recommend when the child should leave care. The Department for Education sets out what councils must do but not how they should do it. Rather than take a national lead, the Department supports sector-led improvement, and relies on LAs to develop practice (National Audit Office, 2014).

2.1 Entry-into-care rates

In recent years, more than 30,000 children have annually been taken into care in England. At the same time, the number of children in care has been around 70,000, reflecting that the average duration of care episodes is around 2.5 years. The annual rate of entry into care, defined as the number of children entering care in a year per 10,000 children in the relevant population, has not always being as prevalent as it is now. Figure 1 shows the annual rate of entry into care in each year over the period 2002 - 2017. The figure shows that the rate of entry has increased by about 30 percent since 2008, with the largest increase taking place between 2008 and 2010, when the entry-to-care rate jumped from around 21 children per 10,000 to 25 children. Reflecting that the time in care has not substantially changed, the figure also shows that the increase in the entry rate has been associated with corresponding increase in the proportion of children in care.

The increase in entry-to-care rates has been the topic of ongoing policy debate between social workers, policy analysts and the media. The 2018 Care Crisis Review identified the key factors contributing to national increases in numbers of looked-after children and applications for care orders, based on analysis of administrative data, review of research, surveys of families and professionals and sector consultation (Thomas, 2018). Some of the key factors identified are discussed next.

The period of the largest jump in entry rates (2008-2010) overlaps with the global economic

[5] A care order is an order which places a child under the care of the LA. A placement order is a court order authorising the LA to place a child for adoption with any prospective adopters it chooses.

recession and the following austerity policy in the UK. Recession and austerity are expected to have an impact on the demand for children's social care through putting more families under financial pressure and stress, but also due to the reduced resources available to LAs for family support and early intervention. The funding cuts have made it harder for the poorest families to cope, limited the number of LAs' programmes aimed to teach parenting skills and offer respite care and led to closure of hundreds of SSCC for children and young people since 2010.

In 2007, the London Borough of Haringey was severely criticised for failing to prevent the death of Peter Connelly, a toddler who suffered more than 50 injuries over an eight-month period. The year after the "Baby P" case care applications across England rose by 36 percent (The Economist, 2018), potentially reflecting increased risk aversion among social workers. Similar cases may have also affected the sector during the last 10 years. For example, in 2011, Blake Fowler died of a head injury at the age of seven in Southampton, after concerns had been raised about him since he was a toddler. Finally, changes in professionals' knowledge related to the impact of neglect on children's well-being and outcomes, as well as the fact that there is not enough guidance for social workers towards working in partnership with parents, may have contributed to the increasing care entry rates.

2.2 Characteristics of children entering care

Table 1 shows the demographic composition of children entering care in the aggregate data. The first column shows the characteristics of children (i.e. aged 0-17) in the general population. The second column shows the prevalence of each demographic characteristic within the population of all children taken into care during the period 2002-2017. As above, the entry rate is the annual number of children entering care per 10,000 children in the relevant population. For example, the first row of the table states that on average 51.2 percent of all children in England were boys during the period studied and, correspondingly, 48.8 percent were girls. However, the proportion of boys among all children entering care was 54.4 percent, reflecting that the entry into care rate is higher for boys (25.60) than for girls (22.52).

The table shows that the entry-to-care rate has been substantially higher for infants than for any other age group: children aged < 1 only make 5.7 percent of the child population, but have accounted for 18.4 percent of all children entering care. The rate of entry is generally U-shaped in age, with the lowest entry rates being for children aged 5-9. The table further shows that children with Black, Mixed and Other ethnicity backgrounds are overrepresented in the population of children taken into care, whilst Asians have a relatively low rate of entry.

Table 1: Rate of entry into care by demographic subgroup

		Population Share	**Share of Entry**	**Entry Rate**
Gender	Male	0.512	0.544	25.60
	Female	0.488	0.456	22.52
Age Group	< 1	0.057	0.184	78.15
	$1 - 4$	0.224	0.189	20.32
	$5 - 9$	0.274	0.171	15.00
	$10 - 17$	0.445	0.456	24.69
Ethnicity	White	0.794	0.670	20.33
	Asian	0.096	0.055	13.94
	Black	0.047	0.093	47.68
	Mixed/Other	0.062	0.116	45.42

Notes: Annual data on all children aged 0-17 in England, pooled over the years 2002 - 2017. The "entry rate" is the average number of children entering care in a year per 10,000 children.

Figure 2 shows the trends in entry rates by subgroup using 2006 as base year. The top panel shows that while the entry rates for all age and gender subgroups are increasing since 2008, the entry rates of male children and children younger than 1 year old and older than 9 years old are increasing faster. The bottom figure shows the variation in trends across groups of children with different ethnic backgrounds. While Table 1 showed that children with Black, Mixed and Other ethnic backgrounds are overrepresented in the children taken into care population compared to the general population, Figure 2 shows that the increase in entry rates is mainly driven by children with White ethnic background.

2.3 Category of need

The reason for a given child being taken into care is recorded by social workers under one of the following categories of need: abuse or neglect, family in acute stress, family dysfunction, child's socially unacceptable behaviour, child's disability, parent's illness or disability, low income and absent parenting. Figure 3 shows the variation of entry rates over time across 4 categories. As already mentioned, abuse and neglect is the most common reason for entering care and its prevalence has been continuously increasing since 2008. In 2016/17, around 16 children in 10,000 children entered care due to abuse and neglect, while the rest of entry rates varied from 0 to 9. Family circumstances, which includes family stress and dysfunction, is the second most common category and it has been roughly consistent over time, with only a very small increase since 2008. The increasing rates of abuse and neglect could potentially reflect improved social workers' knowledge of the impact of neglect on a child's development, as discussed in the literature (Thomas, 2018).

2.4 Local variation

One important feature of the sector is the variation in entry rates across LAs. Different factors including local demographic composition and labour market characteristics contribute to this variation. However, the decision to take a child into care is a local one and it is generally recognised that practice varies locally (National Audit Office, 2014), contributing to this spatial variation. Figure 4 shows the variation in entry rates across "cells" defined by LA and year, by age group. Entry rates of all age groups show wide variation. The figure shows that the entry rate for infants (age < 1) is not only substantially higher than for other age groups, but also showing a lot of variation. Panel A of Table 2 shows the variation in entry rates, within each age group, between LAs and within LAs (over time). For all age groups, there is large variation both across and within areas, with the between- and within-standard deviations being of similar size and quite large compared to the average entry rate. Panel B shows two versions of correlations between the entry rates of different age groups: the overall correlation across all cells below the diagonal and the (average) within-LA correlation above diagonal. Unsurprisingly,

Table 2: Variation in entry rates across local authorities and year by age group (Panel A) and correlation in entry-into-care rates across age groups (Panel B)

Panel A: Between- and within variation				
Age Group	**Entry Rate**	**St. Dev.**	**St. Dev. Between**	**St. Dev. Within**
< 1	78.15	45.07	31.34	32.48
$1 - 4$	20.32	11.92	8.73	8.14
$5 - 9$	15.00	9.20	5.89	7.09
$10 - 17$	24.69	14.76	12.17	8.40
Panel B: Overall and within-correlations				
Age Group	< 1	$1 - 4$	$5 - 9$	$10 - 17$
< 1	1	0.319	0.256	0.252
$1 - 4$	0.619	1	0.435	0.289
$5 - 9$	0.456	0.625	1	0.413
$10 - 17$	0.193	0.257	0.480	1

Notes: Annual data on all children aged 0-17 in England for the years 2002 - 2017. The "entry rate" is the average number of children entering care in a year per 10,000 children. Standard deviations and correlations are calculated over LA-year cells. Panel B presents overall and within-LA correlations below and above the diagonal respectively.

the overall correlation is strong and positive, reflecting permanent differences across LAs that are driving permanent variation in entry rates across all age groups. Perhaps more surprising is the substantial within-LA correlation. These within-LA correlations partly reflect the generally increasing entry rates.

3 Data

The care entry data that we will use for our main analysis uses finer age groups than is available in the published data. Through a request to the Department for Education, we have data, for the years 2007-2017, on the number of children entering care in each LA in each of five age groups $g \in G = \{g_1, ..., g_5\}$. The five age groups are defined as follows: g_1 is defined as age < 1, g_2 as age $1 - 2$, g_3 as age $3 - 4$, g_4 as age $5 - 6$ and g_5 as age $7 - 9$. This finer set of age groups has the benefit of allowing us to analyse more closely the care entry rates for children around the key Sure Start age threshold between 4 and 5.[6]

We use entry rates from 2007 onwards as this allows us to relate entry rates by age group to the level of SSCC provision at the time of eligibility. In particular, for age group g in LA j in year t, we characterize the SSCC provision experienced as that available in the specific LA when the members of this age group were two years old (or as close as possible to this age). This means that for age groups g_1 and g_2 we use the contemporaneous SSCC provision, whilst for age group g_3 we use the 1-year earlier provision, for age group g_4 we use the 3-years earlier provision, and for age group g_5 we use the 5-years earlier provision.[7] This data thus allows us to relate care entry rates to SSCC provision in the post-2003 policy period when SSCC were placed under direct LA control and were being rapidly expanded.

Apart from the main variables of interest, i.e. care entry rates and SSCC provision (described below), our analysis makes use of data on LA background characteristics. Two sets of variables were used: a set of time-varying LA characteristics and a set of "baseline" LA characteristics fixed in year 2004, the beginning of the SSCC expansion period.

The time-varying variables are used to control for LA characteristics that vary significantly over time and are expected, based on the literature on abuse and neglect noted above, to be correlated with entry-to-care rates. For example, as discussed in Section 1 labour market characteristics, e.g. unemployment, can be strongly correlated with child maltreatment. We use LA-level time-varying data for the period 2007-2017 with measures of ethnic composition, economic activity, unemployment claimants proportion, educational qualifications, median income, and local political leadership. A data description is provided in Appendix A.1.

The baseline LA characteristics are used to control for factors that are expected to have contributed to the timing and magnitude of the SSCC expansion in each LA. The selection of those characteristics was based on sector literature on Sure Start Local Programmes (SSLP) and SSCC expansion criteria (e.g. Bate & Foster, 2017) and SSCC impact studies (e.g. Cattan et al., 2019). The baseline characteristics used are: Income Deprivation Affecting Children Index

[6] These age groups constituted a natural refinement of the publicly available data, whilst avoiding going down to single age groups which would have resulted in several suppressed cells due to small numbers. For the same reason, the entry rates cannot be broken down by further demographic characteristics.

[7] For age group g_5 we use the 2004 level of provision if the 5-year lag would have gone further back.

(IDACI), teenage pregnancy, income inequality, school exclusions, low-birth-weight rates, and fertility. Detailed information on each baseline characteristic is provided in Appendix A.2, while more information on the selection of those variables and their relationship with SSCC provision is located in Appendix B.1.

4 Sure Start Children's Centres

4.1 The history of Sure Start

Sure Start Local Programmes

Sure Start is a policy of early intervention providing support to pre-school children and their parents in England. Sure Start was launched in 1998 with the first SSLP being announced in 1999 based on high levels of local deprivation, existing good practice in early years provision, high level of teenage pregnancy and low birth weights. By November 2003, 521 SSLP were operating in England offering: (i) outreach services and home visiting (including a visit to a new mother within 3 months of giving birth), (ii) support for families and parents, (iii) good quality play, (iv) childcare, (v) primary and community healthcare, and (vi) advice about child health and development (Bate & Foster, 2017).

The SSCC expansion 2004-2010

In 2003, Sure Start was universalised and Local Programmes were transitioned to SSCC, signalling a move from targeted intervention to a universal delivery model. In 2004, the "Choice for Parents" 10-year strategy announced the target of one children's centre in every community by 2010. This strategy was developed over three phases.

- 2004-06: Phase One - targeting 20 percent most disadvantaged areas, as defined by the Income Deprivation Affecting Children Index (IDACI).
- 2006-08: Phase Two - targeting 30 percent most disadvantaged areas, as defined by IDACI.
- 2008-10: Phase Three - one centre in every community.

By 31 July 2010, 3,633 SSCC were open in England. Children's centres opened in Phase 1 and Phase 2 offered a wide range of services similar to what was offered by SSLP and acted as a gateway to more specialised provision for young children and families. Each centre offered on average 28 services from a list of possible 50, ranging from a minimum of 13 to a maximum of 42. On the other hand, Phase 3 centres were more likely to operate on a part-time basis and offer a smaller range of services, while they were predominantly not located in disadvantaged areas (Smith et al., 2018).

The SSCC reduction since 2010

In April 2011, the ring-fence on Sure Start funding was removed and an Early Intervention Grant was introduced. Consequently, Sure Start was now competing for funding with other not child-focused early intervention services including careers services for young adults, substance misuse services, and young offender and crime prevention services.

In 2013, the new statutory guidance defined children's centres' core purpose as improving outcomes for young people and their families and reducing inequalities between families in *greatest need*. The new guidance was a signal for a targeted focus of SSCC on families facing difficulties and at a high risk instead of the previous aim of providing open access services to all local families.

More than 500 children's centres were closed since 2010. Apart from centres' closures, the limited funding and the changes in guidance led to other changes in provision. According to the Education Select Committee's report in December 2013, many LAs were redesigning their centres so that they operated in clusters, an increasing number of centres were targeting services, partly because of reductions in funding and partly because of the new core purpose, and in some cases, centres had reduced their services rather than closed (Bate & Foster, 2017).

Although the criteria for the roll-out of SSCC were specified by the government, less is known about any systematic characteristics of areas that have reduced their SSCC provision. According to the 2018 Sutton Trust's report's findings from a survey of 124 out of 152 LAs in England, the principal reasons for major changes in children's centre provision were funding pressures, change of focus, high cost of stand-alone centres and changed national priorities (Smith et al., 2018).

4.2 SSCC variation over time and across Local Authorities

Figure 5 shows the number of SSCC in operation in England from 2004 to 2017. The figure also shows the national average "SSCC coverage", defined as the number of SSCC per 1,000 children aged 0-4. As discussed above, the number of SSCC increased sharply after 2004, reaching more than 3,600 centres in 2010. Since 2010 the number is decreasing slowly.

Figure 6 shows the overall evolution of SSCC across England. The first map shows the average SSCC coverage in each LA in England between 2004-8, as expected the average coverage over this period was low with the majority of LAs having 0-0.35 SSCC per 1,000 eligible children. The second map shows the period with the highest coverage, i.e. 2009-2011. The vast majority of LAs have a coverage rate close to 1 centre per 1,000 eligible children, with a few exceptions having more than 1.4 centres. The third map represents the contraction period, and it shows that the closures were concentrated in a few specific LAs and there was not a nationally widespread decrease in the number of SSCC.

The strong expansion between 2004 and 2010, however, was not random but reflected that the policy was initially aimed at more deprived areas. As a result, the pace and timing of expansion varied across areas of different characteristics. This feature of the expansion build-up period is highlighted in Figure 7. The left panel of the figure classifies the LAs by their IDACI scores in the baseline year of 2004, splitting them into quartiles and then plotting the average SSCC coverage by IDACI quartile and year. This shows how the SSCC expansion was initially faster in the more deprived LAs. For example, in 2006, the average coverage was 0.3 higher in the most deprived 25 percent of LAs than in the 25 least deprived ones. However, by 2010 the systematic gap in coverage was largely eliminated. The middle and right panels do the same exercise but instead splitting the LAs into quartiles based on their 2004 rates of teenage pregnancy and low birth weights respectively. Both panels show similar patterns as for the IDACI scores, though slightly less pronounced. The non-random expansion of SSCC provision will be an empirical challenge discussed in some detail below. Section B.1 reports the results of a regression analysis of SSCC expansion determinants.

Table 3: Patterns of SSCC closures over the period 2011 to 2017

Proportion		**Concentration**		**Phase**		**Sub-period**	
Proportion closed	Local Authorities	Local Authorities	Share of closures	Phase created	Proportion closed	Year of closure	Share of closures
0	75	20	3.4%	1	9.8%	2011-13	10.6%
$0 - 0.5$	65	51	61.7%	2	17.0%	2014-15	27.1%
$0.5 - 1$	11	5	34.9%	3	24.2%	2016-17	62.3%
	151	76	100%				100%

Notes: The table describes SSCC closure patterns between 2011 and 2017. Over this period there were 519 SSCC closures out of the 3,615 that existed in 2010.

While the build-up of SSCC provision had partly known determinants, less is known about any potential systematic patterns to the closures that have taken place since 2010. In total, 519 SSCC were closed between 2011 and 2017. Table 3 provides some basic information about the pattern of closures. It first notes that, effectively, half of the LAs had not closed any SSCC by 2017. Next, it shows that, among the 76 LAs that close at least one children's centre, 20 LAs accounted for only 3 percent of closures (each closing exactly one children's centre). In contrast, five LAs accounted for more than a third of all closures. The next column shows later established SSCC were more frequently closed, while the final column shows that the majority of closures occurred only after 2015. The fact that later established SSCC were more likely to close also implies an indirect link to the local level of deprivation. This is highlighted in Appendix B.2 which provides a simple "survival analysis", linking the likelihood of a given children's centre to close to local area characteristics. In general, the link between area characteristics and SSCC closures is substantially weaker than is the link to timing of SSCC build-up during the expansion period.

5 Conceptual Framework and Empirical Strategy

The aim of Sure Start was "giving children the best possible start in life" through family support, improved childcare, health and early education (Department for Children, Schools and Families, 2008). As such, SSCC provision can be viewed as a facility for investment in young children's development, heath and well-being, as well as a vehicle for supporting parents' employment, parenting behaviour and mental health. Thus, benefits from SSCC provision can be expected to have longer run returns not only in terms of children's health and educational attainment, but also in terms of encouraging improved family outcomes.

These potential long term benefits have indeed so far been the focus of evaluations of Sure Start. While findings have been positive, as a general characterization, the effects have also been found to be relatively modest. The National Evaluation of Sure Start (NESS) team reported that families in SSLP areas showed less negative parenting – inducing less chaotic home environments and creating a more stimulating home learning environment with less harsh disciplining – along with some improved child outcomes – lowering BMIs, improved social behaviour, and higher rates of immunization and fewer accidental injuries (Melhuish et al., 2008; Melhuish et al., 2010; Melhuish et al., 2012). Additionally, the Department for Education (Sammons et al., 2015) evaluated the impact of children's centres in improving family and child outcomes for a broader sample of user families than earlier NESS evaluations. The findings were in line with previous studies, predicting improved home-learning environment, reduced parental distress and improved social child behaviour. More recent evidence by Cattan et al. (2019), exploiting the full expansion of SSLP and SSCC found that the availability of children's centres had led to improved BMIs and lower rates of hospitalization for respiratory illness and infections among children.

5.1 Hypothesized effects

The first channel through which SSCC provision may have affected the incidence of statutory care would thus be through a longer term positive investment effect, reducing dysfunctional family behaviour including abuse and neglect. Such improvements would not only occur contemporaneously but would reduce entry to care over time, including beyond the target age of SSCC. However, while the investment effect of SSCC provision would reduce the need for entry into care, such provision could also lead to the identification of children in need of statutory care. Such an identification effect would increase the rate of entry within the target age group.

Table 4: Predicted effects of SSCC provision on care entry rates by age group

Channel	**Currently Eligible (0-4)**	**Past Eligible (5+)**
Investment	-	-
Identification	+	-/0
Total	+/-	-

One version of this effect would be towards earlier identification. Such a timing of identification effect could contribute to reducing entry among the children beyond the SSCC target age.

Table 4 summarizes the hypothesized effects. Two key features should be noted. First, the threshold age between 4 and 5 provides key information on the underlying mechanisms: if there was no identification effect – only an investment effect – then we would expect SSCC provision to reduce entry into care among all age groups and, furthermore, there should be no discontinuity in the effect of SSCC provision at the end of the SSCC target age, that is the effect of SSCC provision should be similar for children aged closely up to 4 as for children just beyond that age. If a positive investment effect is present, then the impact of SSCC provision should be associated with lower entry into care among children aged 5 and above.

5.2 Empirical Model and Identification

To identify the effects of SSCC provision on care entry rates we exploit the variation introduced by the expansion of SSCC provision over time, which generated differences in coverage across LAs and, critically, across cohorts within LAs. The outcomes of interest that we observe in our data are the entry rates of children of different age groups. In particular, we will model the entry-to-care rate in age group g, area j and year t, denoted y_{gjt} as follows

$$\log\left(y_{gjt}\right) = \beta^{g}\widehat{cc}_{gjt} + \eta_g I_g + \delta_j I_j + \nu_t I_t + \omega_g I_g t + \mu_j I_j t + \rho Z_{jt} + \varepsilon_{gjt}. \qquad (1)$$

In this specification, $\widehat{cc}_{gjt}$, defined for children in age group g in year t and in area j, measures the level of SSCC coverage that was available to children in that area at the time when this cohort were of central SSCC target age. As noted above, for age groups g_1 (age < 1) and g_2 (age $1-2$) $\widehat{cc}_{gjt}$ is measured with the contemporaneous level of SSCC coverage while for age groups g_3 through to g_5 coverage $\widehat{cc}_{gjt}$ is measured one, three and five years earlier respectively for area j. Note that the model allows the effect of SSCC coverage to have age-group-specific effects on care entry rates, β^g. These age-group-specific effects are the key parameters of interest.

The estimating equation (1) further includes age-group fixed-effects η_g. The age-group specific effects model the fact that entry rates differ substantially with age, for instance being markedly higher among infants (age group g_1) than among the older age groups. The use of the outcome entry rates in log form also reflects the fact that the average entry rates are different across age groups as it allows us to interpret the various estimated effects as proportional effects.

The model further includes LA fixed-effects δ_j, and year fixed-effects ν_t. The area fixed-effects δ_j control for permanent unobserved differences across areas while the year fixed-effects ν_t control for a common national trend. As noted above, the national trends were fairly similar across age groups. Nevertheless, in order to allow for age-group-specific national trends, we include interactions between the age group dummies and time. The associated estimated parameters ω_g thus measure any additional age-group-specific (linear) national trends. The area fixed effects, δ_j, while controlling for any permanent unobserved differences across areas, will not account for any local time-varying characteristics that could influence both the local development of SSCC provision and the care entry rates. To tackle this threat to identification we adopt two complementary strategies. First, we include additional LA-specific linear time trends μ_j. Second, we include a set of time-varying LA characteristics Z_{jt}. These include e.g. labour market conditions, ethnicity and educational demographic composition and political leadership. In specification tests, we also make use of a set of baseline characteristics, denoted Z_j^b, that were found above to be related to the timing of SSCC expansion. We replace first the area-specific linear trends with Z_j^b interacted with time, and second we replace the LA fixed effects directly with Z_j^b.

Table 5: Illustration of variation in current/post SSCC exposure generated by variation in the timing of SSCC build-up across local authorities

	Local Authority A			**Local Authority B**	
Sub-period	SSCC	Cohorts		SSCC	Cohorts
$Period = 0$ ("02-05")	No	$CE(c_0) = 0$		No	$CE(c_0) = 0$
$Period = 1$ ("05-07")	Yes	$PE(c_0) = 0$ $CE(c_1) = 1$	⇔ (row 2)	No	$PE(c_0) = 0$ $CE(c_1) = 0$
$Period = 2$ ("07-09")	Yes	$PE(c_1) = 1$ $CE(c_2) = 1$	⇔ (row 1)	Yes	$PE(c_1) = 0$ $CE(c_2) = 1$
$Period = 3$ ("09-11")	Yes	$PE(c_2) = 1$ $CE(c_3) = 1$		Yes	$PE(c_2) = 1$ $CE(c_3) = 1$

A stylized illustration of the identification strategy is provided in Table 5. The figure contrasts two hypothetical LAs, A and B. Children born "early" (say 2002-2005) in either area had no current SSCC provision ($CE\,(c_0) = 0$) available whilst of eligible age. Correspondingly, when we observe the entry rates from these children a few years later (around 2005-2007) when they are beyond SSCC target age we know that they had no access to SSCC services when they were of eligible age ($PE\,(c_0) = 0$ in both areas). Suppose then that area A creates SSCC capacity at this time, but area B does not yet do so. The current young children in area A thus enjoyed SSCC availability while the corresponding children in area B do not ($CE\,(c_1) = 1$ in area A but $CE\,(c_1) = 0$ in area B). A further few years later we observe this cohort being of post-target age, with only the children in area A having earlier enjoyed SSCC availability. As area B now expands SSCC provision, we get a first cohort in this area benefiting from SSCC provision when young.

The figure thus illustrates how the build-up of SSCC coverage created cohort variation in SSCC availability both for children of current eligible age and for children beyond eligible age. Moreover, as the timing of SSCC expansion was different across LAs, the cohort variation was not synchronized across areas, facilitating the identification of SSCC effects from common trends. Finally, note that at key phases, children of the same age from different areas would have enjoyed different levels of current or past SSCC provision. However, such variation would confound differences in SSCC provision with area differences. In the stylized example, area A could for instance be expected to be characterized by a higher level of deprivation than area B. Hence it is key that specification (1) identifies the effect of SSCC provision from within-LA cohort variation and not from contemporaneous differences in SSCC provision across LAs.

6 Results

6.1 Main regression results

Table 6 presents our basic estimates of equation (1). The estimations model the care entry rates for the five age groups in G, for 147 LAs, and for 11 years. Hence the maximum number of observations is 8,085. As our main interest is in the estimated effects of SSCC provision, we only present the estimates of the β^g terms: further estimated coefficients on time-varying and baseline LA characteristics from selected specifications are presented in Appendix B.3.

Specification (1) in Table 6 starts by constraining the estimated effect of SSCC provision on care entry rates to be uniform across all age-groups. The estimated common effect is positive but fairly modest in size and statistically significant only at the five percent level. Specification (2) is the first estimated specification with age-group specific effects. This specification shows that the initially estimated common effect masks effects that vary substantially across age groups. For each of the youngest three age groups g_1 - g_3, the estimated effects of SSCC are strongly positive and statistically significant. In contrast, for the two higher age groups, g_4 and g_5, the estimated effects are negative but smaller in absolute terms and not statistically significant above the ten percent level.

These finding of a differential effects of SSCC provision on children who are still within the target age (0-4) and on children beyond this age (5-9) is thus consistent with the predictions from section 5. In particular, the estimated positive effect for age groups g_1 - g_3 indicates the presence of an identification effect. In contrast, the estimated entry-reducing effects of SSCC provision for age-groups g_4 and g_5 is consistent with a lasting investment effect of past eligibility and access to SSCC services leading to a lower current need for statutory care. Specification

Table 6: Effect of SSCC provision on care entry rates: main results

	(1)	(2)	(3)
Any age group	0.0584*		
	(0.0283)		
Aged < 1		0.252***	0.259***
		(0.0560)	(0.0575)
Aged $1 - 2$		0.198***	0.204***
		(0.0574)	(0.0592)
Aged $3 - 4$		0.148**	0.157**
		(0.0471)	(0.0484)
Aged $5 - 6$		-0.0618	-0.0578
		(0.0457)	(0.0466)
Aged $7 - 9$		-0.111+	-0.105+
		(0.0561)	(0.0541)
Observations	8,085	8,085	7,930
Year effects	Y	Y	Y
Age group effects	Y	Y	Y
Age group trends	Y	Y	Y
LA effects	Y	Y	Y
LA trends	Y	Y	Y
Time-Varying LA	N	N	Y

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The table presents estimated β^g coefficients from estimating equation (1) on data on log entry rates into care among five age groups across 147 Local Authorities in England between 2007 and 2017. The time-varying LA characteristics are described in Appendix A. Standard errors are clustered at the LA level.

(2) did not include any time-varying LA characteristics, Z_{jt}. Such characteristics are added in specification (3) and the estimated coefficients for these variables are presented in column (1) of Table B3 in Appendix B.3. Inspection of those estimated coefficients reveals that no single variable in Z_{jt} is statistically significant, suggesting that the already included area-level trends control well for local time-varying factors. Correspondingly, the estimated age-group-specific effects of SSCC provision on care entry rates are barely affected by the inclusion of the time-varying LA characteristics.

6.2 Alternative specifications

Table 7 presents a sequence of alternative specifications to test for the robustness of the results. The first column in Table 7 reiterates our preferred specification (3) from Table 6 for purposes of comparison. Specification (2) in Table 7 removes the area-specific linear time-trends and replaces them with baseline characteristics interacted with time $Z_j^b t$, while leaving in the time-varying LA characteristics. Specification (3) further replaces the area-fixed effects with the time-invariant baseline LA characteristics Z_j^b. In either case, the estimated β^g-effects are strikingly robust. The estimated coefficients on $Z_j^b t$ and Z_j^b from these two specifications are presented in Table B3 in Appendix B.3. These show that, once we no longer include the area-specific linear time trends, the measured economic activity rates in particular is negatively associated with the care entry rates, as would be consistent with the literature indicating that higher rates of economic activity are associated with lower rates of child maltreatment. When we further replace the LA fixed effects with the baseline characteristics Z_j^b we find that high levels of deprivation and teenage pregnancies are associated with higher entry rates.

Specification (4) in contrast confirm that controlling for permanent differences across LAs – either via fixed-effects or via baseline characteristics – is critical. This specification removes both the LA fixed effects and the use of baseline characteristics. Doing so implies that the β^g coefficients are now estimated using cross-LA variation in SSCC provision. In this case we should expect the estimated coefficients to be strongly biased upwards as the SSCC provision would typically have been higher in more deprived areas where the care entry cares would also have been substantially higher. Indeed, when the "effect" of SSCC provision is estimated using such cross-area variation, the estimated coefficients increase substantially and all turn (implausible) positive, thus indicating a strong upward bias.

As we will show below, there is some evidence of age-entry profiles varying with area-baseline characteristics. In order to account for this, specification (5) includes interactions between LA baseline characteristics and age group. However, doing so does not substantially affect the estimated parameters of interest.

Specification (6) reduces the sample period to 2007 - 2012, thus using measures of coverage

Table 7: Effect of SSCC provision on care entry rates: sensitivity analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Aged < 1	0.259***	0.217***	0.229***	0.549***	0.204***	0.358***	0.282***		
	(0.0575)	(0.0592)	(0.0607)	(0.0919)	(0.0534)	(0.0843)	(0.0625)		
Aged $1-2$	0.204***	0.162**	0.174**	0.491***	0.164**	0.232**	0.223***		0.210**
	(0.0592)	(0.0617)	(0.0638)	(0.0882)	(0.0557)	(0.0813)	(0.0628)		(0.0676)
Aged $3-4$	0.157**	0.120*	0.133*	0.480***	0.137**	0.213**	0.163**	0.135	
	(0.0484)	(0.0495)	(0.0564)	(0.0821)	(0.0475)	(0.0725)	(0.0529)	(0.0826)	
Aged $5-6$	-0.0578	-0.0848+	-0.0533	0.366***	-0.0693	-0.144+	-0.115+	-0.0764	-0.0775
	(0.0466)	(0.0498)	(0.0612)	(0.0844)	(0.0469)	(0.0814)	(0.0613)	(0.0792)	(0.0647)
Aged $7-9$	-0.105+	-0.105*	-0.0477	0.403***	-0.116*	0.00856	-0.155*		
	(0.0541)	(0.0526)	(0.0632)	(0.0899)	(0.0500)	(0.0911)	(0.0774)		
Observations	7,930	7,930	7,930	8,085	7,930	4,330	7,930	3,172	3172
Year effects	Y	Y	Y	Y	Y	Y	N	Y	Y
Age gr. effects	Y	Y	Y	Y	Y	Y	Y	Y	Y
Age gr. trends	Y	Y	Y	Y	Y	Y	Y	Y	Y
LA effects	Y	Y	N	N	Y	Y	Y	Y	Y
LA trends	Y	N	N	N	Y	Y	Y	Y	Y
Time-varying	Y	Y	Y	N	Y	Y	Y	Y	Y
Baseline LA	N	N	Y	N	N	N	N	N	N
Baseline×time	N	Y	Y	N	Y	N	N	N	N
Baseline×age	N	N	N	N	Y	N	N	N	N
Cohort effects	N	N	N	N	N	N	Y	N	N

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The table presents estimated β^g coefficients from estimating equation (1) on data on log entry rates into care among five age groups across 147 Local Authorities in England between 2007 and 2017. The time-varying and baseline LA characteristics are described in Appendix A. Standard errors are clustered at the LA level.

mainly from the SSCC expansion period. In particular, the specification avoids using the sample years when some SSCC were closed and, more generally, funding was being reduced. Doing so generally increases the estimated SSCC effects in absolute terms, but of course also makes the estimates less precise.

Specification (7) control for cohort effects rather than for year fixed effects.[8] In this case we find that the estimated values of β^{g_1} - β^{g_3} are largely unaffected, but the estimated entry-reducing effects on children aged five and above increases.

Specification (8) takes a closer look at the entry rates for children in the age groups just below and above the threshold delineating those currently in the Sure Start target age group and those beyond it. In other words, the specification only includes age groups g_3 and g_4, and hence only estimates β^{g_3} and β^{g_4}. These two groups are naturally close in age and can be expected

[8]Cohorts are defined as year minus age. But, of course, our age groups are banded. Hence we define the cohort of age group g_1 in year t as t, the cohort of age group g_2 as $t-1$, the cohort of age group g_3 as $t-3$, and the cohort of age group g_4 as $t-5$, and finally the cohort of age group g_5 as $t-7$. The specification further controls for general trends using year and year-squared.

to be similar in other respects, except age group g_3 are within SSCC target age while g_4 are instead just of schooling age. Estimating using only these two groups generates does change the estimated effects in any noticeable way. For comparison, specification (9) does a corresponding estimation using only age groups g_2 and g_4. Again, the results are largely unaffected.

6.3 Age-group pairwise comparisons

The final two columns of Table 7 introduced the idea of comparing age groups pairwise. The approach of making pairwise age-group comparisons can be taken further. Differencing the estimating equation (1) between any two age groups g and g', where $g < g'$, and decomposing the difference $\beta^g \widehat{cc}_{gjt} - \beta^{g'} \widehat{cc}_{g'jt}$, into a difference in coverage and a difference in impact of coverage, we obtain that

$$\begin{aligned} \Delta \log y_{jt}^{g,g'} &= \beta^g \left(\widehat{cc}_{gjt} - \widehat{cc}_{g'jt}\right) + \left(\beta^g - \beta^{g'}\right) \widehat{cc}_{g'jt} \\ &\quad + \left(\eta_g - \eta_{g'}\right) + \left(\omega_g - \omega_{g'}\right) t + \varepsilon_{gjt} - \varepsilon_{g'jt} \end{aligned} \tag{2}$$

This shows that the parameters of main interest can be recovered from estimations that make pairwise use of the various age groups. In particular, regressing the log difference in the entry rates on the *difference* in SSCC coverage experienced by the two groups as well as on the *level* experienced by group g' provides estimates of β^g and of $\Delta\beta^{g,g'} \equiv \beta^g - \beta^{g'}$. It consequently also provides an estimate of $\beta^{g'}$ as $\beta^{g'} \equiv \beta^g - \Delta\beta^{g,g'}$. The constant in the regression (2) estimates the difference in group averages as represented by the difference in the age-group fixed effects $\eta_g - \eta_{g'}$ in the original estimating equation. Furthermore, the estimated coefficient on time t captures the (linear) difference in the age-group trends.

An advantage of estimating in age-group differenced form is that it does not rely on the original estimating equation (1) being correct with respect to the linear specification of LA-specific trends and time-varying characteristics. For instance, these could take non-linear forms; as long as they are additive and can plausibly be assumed to be the same for the two age groups that are being compared, these effects are removed by the differencing.

Note that (2) can be estimated for any pair of age groups. Since we have five age groups in

Table 8: Effect of SSCC provision on care entry rates: pairwise regressions

	(g_1, g_2)	(g_1, g_3)	(g_1, g_4)	(g_1, g_5)	(g_2, g_3)	(g_2, g_4)	(g_2, g_5)	(g_3, g_4)	(g_3, g_5)	(g_4, g_5)
β^{g_1}	(.)	0.0842	0.230***	0.246***						
		(0.0763)	(0.0643)	(0.0475)						
β^{g_2}	(.)				0.0738	0.206**	0.207***			
					(0.0739)	(0.0656)	(0.0491)			
β^{g_3}		0.00509			0.0533			0.155*	0.154***	
		(0.0711)			(0.0689)			(0.0753)	(0.0462)	
β^{g_4}			-0.123			-0.0740		-0.0392		-0.0218
			(0.0638)			(0.0651)		(0.0800)		(0.0589)
β^{g_5}				-0.274***			-0.152***		-0.0922	-0.0436
				(0.0570)			(0.0587)		(0.0589)	(0.0701)
$\Delta\beta$	0.0547	0.0791+	0.353***	0.520***	0.0204	0.280***	0.359***	0.194***	0.246***	0.0218
	(0.0359)	(0.0429)	(0.0596)	(0.0702)	(0.0415)	(0.0608)	(0.0722)	(0.0567)	(0.0699)	(0.0668)
Obs.	1,617									

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Each column corresponds to a regression of the form (2) for a pair of age groups. g_1 = age < 1, g_2 = age $1-2$, g_3 = age $3-4$, g_4 = age $5-6$, g_5 = age $7-9$. For the pair (g,g') with $g < g'$, the directly estimated coefficients are β^g and $\Delta\beta = \beta^g - \beta^{g'}$. $\beta^{g'}$ is recovered as the difference between the estimated coefficients, $\beta^{g'} = \beta^g - \Delta\beta$.

G, we can form ten age-group pairs. The upshot of this is that we obtain (up to) four estimates of each β^g-term. The one exception is the comparison between g_1 and g_2 as, for both these age groups we measure SSCC coverage with the contemporaneous coverage. This implies that $\widehat{cc}_{g_1jt} - \widehat{cc}_{g_2jt}$ has no variation and we can only estimate the difference $\Delta\beta^{g_1,g_2}$ – we cannot separate out estimates of β^{g_1} and β^{g_2} from this age-group pair.

Table 8 shows the results from all pairwise regressions. The general findings is that every point estimate of the effect of SSCC provision on the entry rates of children *within* the Sure Start target age, that is every estimate of β^{g_1}, β^{g_2} and β^{g_3}, is positive. In contrast, every point estimate of the effect of SSCC provision on the entry rates of children *beyond* the target age, that is every estimate of β^{g_4} and β^{g_5}, is negative. Comparing the estimates for β^{g_1}, β^{g_2} and β^{g_3} the general pattern is that the estimates tend to decrease with age, both in terms of range and average.

The estimated positive coefficient on β^{g_3} in the (g_3, g_4)-pairwise regression suggests that recent expansions of SSCC provision would be associated with increases in the *relative* rate of entry into care for age group g_3 who are still within the target age group compared to the "just-beyond-target-age" group g_4. This is illustrated in Figure 8. Defining $\Delta \log y_{jt}^{g,g'} \equiv \log\left(y_{gjt}\right) - \log\left(y_{g'jt}\right)$ as the log difference in entry rates between age groups g and g', the top left panel plots the distribution of $\Delta \log y_{jt}^{g_3,g_4}$. The distribution is centered around 0.15, indicating that the entry rate is, on average, about 15 percent higher among children aged 3 - 4 compared to children aged 5 - 6.

Next we define $\Delta\widehat{cc}_{jt}^{g,g'} \equiv \widehat{cc}_{gjt} - \widehat{cc}_{g'jt}$. Note in particular that $\Delta\widehat{cc}_{jt}^{g_3,g_4}$ will be positive if the SSCC coverage for children now aged 3 - 4 was higher than that for children now aged 5 - 6, that is, if there was a recent expansion of SSCC provision. The right panel of figure 8 plots the average of $\Delta \log y_{jt}^{g_3,g_4}$ by quartile of $\Delta\widehat{cc}_{jt}^{g_3,g_4}$. The figure confirms that recent increases in SSCC provision are associated with higher relative entry into care for children aged $3-4$ compared to children aged $5-6$.

The lower two panels correspondingly compare age group g_2 (aged 1 - 2) to age group 4. In this case, the left panel shows that the distribution of $\Delta \log y_{jt}^{g_2,g_4}$ is somewhat further to the right, reflecting that the average rate of entry into care of children aged 1 - 2 is on average about 45 percent higher than the corresponding rate for those aged 5 - 6. Hence age groups g_2 and g_4 are, naturally, not as similar as are g_3 and g_4. Nevertheless, the same pattern emerges. When we characterize the change in coverage $\Delta\widehat{cc}_{jt}^{g_2,g_4} \equiv \widehat{cc}_{g_2jt} - \widehat{cc}_{g_4jt}$ for each (j,t)-cell we find that cells associated with recent increases in SSCC provision are also associated with an increase in the relative entry of children aged 1 - 2 compared with children aged 5 - 6.

The main concern with the above analysis is the possibility that some local time-varying factors could influence both the local SSCC provision and care entry rates. We tackled this issue by including a set of time-varying LA characteristics and by allowing for LA-specific linear time trends. Moreover, as our results have shown, for any unobserved time-varying factors to threaten the validity of our results, they would have to have differentially affected the entry rates of children within and beyond Sure Start target age.

Nevertheless, a further robustness test can be done by relating the relative age-group entry rates directly to baseline LA-characteristics. From Figure 7 we know that, in line with the stated Sure Start rollout principles, between 2004 and 2010, LAs that were "disadvantaged" – as measured by higher baseline values of the IDACI deprivation score, the baseline rate of teenage pregnancy, and the baseline rate of low birth weight – had higher levels of SSCC coverage.

This allows us to test the robustness of our finding that SSCC provision was associated with higher relative rates of entry into care for children within target age. Specifically, the finding would imply that disadvantaged areas should be observed to have had higher relative rates of entry *for age groups within the SSCC target age* in *among cohorts who were within eligible age during the Sure Start expansion period.* This prediction can be tested using a simple difference-in-difference approach. The attractiveness of such a test is that it relies only on observable *baseline* characteristics of LAs, making it unaffected by temporary local shocks that could have affected both SSCC provision and entry rates.

In order to implement such a test we proceed as follows. We start by creating a standardized index of disadvantage, denoted, ζ_j^b, which is a linear combination of the baseline (2004) IDACI, teenage pregnancy and low birth weight rate. The index is created – using a simple regression – as the linear combination of these three factors that best predicts SSCC coverage over the period 2004 to 2010.[9] Next we split our sample period for the entry analysis into (i) an "early" subperiod 2007-2012 that includes most entry observations associated with SSCC eligibility up to 2010, and (ii) a "late" subperiod 2013-2017 that includes most entry observations associated with post-2010 SSCC eligibility.[10] We then estimate the following regression, relating the log difference in entry rates between two age groups g and g', with $g < g'$, to the disadvantage index ζ_j^b, the index interacted with a dummy for "early" subperiod, and a set of year fixed effects where the latter allow us to control for any general national trend in relative entry rates.

$$\Delta \log y_{jt}^{g,g'} = \gamma_0 \zeta_j^b + \gamma_1 \zeta_j^b I_{2007 \leq t \leq 2012} + \eta_t I_t + \epsilon_{jt}^{g,g'}. \tag{3}$$

As we have five age groups in G we can, as before, form ten age group pairs and estimate (3) for each pair separately. The results from doing so are presented in Table 9. The ten estimated γ_0 values are presented below the diagonal while the corresponding ten γ_1 values are presented

[9] Specifically, we regress local SSCC coverage, cc_{jt}, over the period 2004-2010 on the standardized baseline (2004) IDACI, teenage pregnancy and low birth weight rate and generate predicted values – which per construction are time-invariant. Our index ζ_j^b, which loads positively particularly on the IDACI score and the teenage pregnancy rate, is obtained by standardizing the predicted values.

[10] Note that, as we will be using pairs of age groups, the age groups would have had different timing of SSCC eligibility. This leads to some inevitable misclassification, e.g. for infants, g_1, observed and contemporaneously eligible in 2011-12, and for oldest children, g_5, in 2013-14 who were mainly SSCC eligible pre-2010.

above the diagonal. Note that a positive value on γ_0 would indicate that relatively disadvantaged areas have a higher relative care entry rate for the younger age group g compared to older age group g'. In fact, most estimated γ_0 values are negative – and often statistically significantly – hence indicating that the more disadvantaged areas tend to have a slightly different entry age-profiles tilted more towards relatively older children than the less disadvantaged areas.[11]

Our main parameter of interest, however, is γ_1, and in particular for (g, g')-pairs that involve one age group within Sure Start target age ($g_1 - g_3$) and one beyond target age ($g_4 - g_5$). There are six such age group pairs and the associated estimated values of γ_1 are highlighted by the box in the top right corner of the table. Consistent with our main results, the coefficients on each of these six interaction terms are positive and in a number of cases statistically significant (in contrast to the remaining four estimated γ_1 values). Consider for instance age groups g_3 and g_4. The negative estimated value for γ_0 for this pair indicates that, in the "later" period 2013 to 2017, areas that were disadvantaged according to their 2004 baseline characteristics ($\zeta_j^b > 0$) had a higher entry rate for children aged 5-6 relative to children aged 3-4 than the corresponding relative entry rate among "'non-disadvantaged" areas ($\zeta_j^b \leq 0$). The positive and numerically larger value for γ_1 for the same age-group pair indicate that, the opposite was true during the early period 2007-2012: during this subperiod the disadvantaged areas had a higher entry rate for children aged 3-4 relative to children aged 5-6 than the corresponding relative entry rate among "'non-disadvantaged" areas.

Similarly, when comparing age groups g_2 and g_4, the table shows that in the later subperiod, the areas that were disadvantaged at baseline had a substantially higher rate of entry of children aged 5-6 relative to children aged 1-2 than the non-disadvantaged areas. In contrast, in the early subperiod the disadvantaged and non-disadvantaged areas had similar relative entry rates for the two age groups.

The results are thus consistent with the age-composition of children entering care changing over time differently in the disadvantaged and non-disadvantaged areas, and in particular with entry being relatively high for age groups within the SSCC target age in disadvantaged areas in

[11] This relationship between the age-profiles of care entry rates and local disadvantage motivated specification (5) in Table 7 that allowed for an interaction between LA baseline characteristics and age group

Table 9: The relationship between relative age group entry rates and baseline disadvantage over time

	g_1 (< 1)	g_2 $(1-2)$	g_3 $(3-4)$	g_4 $(5-6)$	g_5 $(7-9)$
g_1 (< 1)	-	-0.0251 (0.0188)	-0.0281 (0.0223)	0.0328 (0.0284)	0.0174 (0.0258)
g_2 $(1-2)$	0.0279* (0.0138)	-	-0.00317 (0.0215)	0.0574* (0.0289)	0.0420 (0.0263)
g_3 $(3-4)$	-0.0186 (0.0162)	-0.0460** (0.0156)	-	0.0619* (0.0293)	0.0457+ (0.0266)
g_4 $(5-6)$	-0.0543** (0.0205)	-0.0815*** (0.0208)	-0.0357+ (0.0210)	-	-0.0171 (0.0282)
g_5 $(7-9)$	-0.0793*** (0.0187)	-0.107*** (0.0190)	-0.0608** (0.0190)	-0.0251 (0.0202)	-
Obs.			1,617		

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: A "disadvantage index", denoted ζ_j^b, was created by regressing the level of SSCC coverage, cc_{jt}, in area j in year t for years 2004-2010 on the baseline (2004) IDACI score, teenage pregnancy rate, and rate of low birth weight. The (time-invariant) index, ζ_j^b, was generated as the standardized predicted values from this regression. The sample period 2007-2017 for the observed care entry rates was split into two subperiods, 2007-2012 and 2013-2017. For each age-group pair (g, g'), with $g < g'$, the log difference in entry rates, $\Delta \log y_{jt}^{g,g'}$, was regressed on the disadvantage index, ζ_j^b, the index interacted with a sub-period 1 dummy $\zeta_j^b \times I_{2007 \leq t \leq 2012}$, along with a full set of year dummies. The terms *below* the diagonal are the estimated coefficients on ζ_j^b in each pairwise regression. The terms *above* the diagonal are the estimated coefficients on the interaction term $\zeta_j^b \times I_{2007 \leq t \leq 2012}$.

the early subperiod. This is thus consistent with earlier entry rates being high in SSCC target age groups in LAs that, due to their baseline characteristics, developed SSCC capacity earlier.

6.4 Counterfactual care entry rates

The above findings strongly suggest that SSCC provision has led to higher rates of entry into care among children within the target age range $0-4$ but also to a reduction in the rate of entry for children age $5-9$. To gauge the potential impact of the SSCC on the rates of entry of the various age groups over time, we can use a counterfactual simulation. Specifically, we can compare the model-predicted care entry rate in age group g and year t under the actual level of SSCC provision to that predicted under the counterfactual scenario in which no SSCC services were provided. The result of this exercise is shown in Figure 9.

Each panel plots first the actual empirical care entry rate, followed by the model-predicted rate, and finally the care entry rate predicted for the counterfactual scenario without SSCC provision. Looking first at the older age groups, g_4 and g_5, the model predicts that provision of SSCC services reduced the entry into care among these children, aged $5-9$. The gap in the prediction is essentially zero for the first few sample years – this reflects that the level of SSCC provision experienced by children in these age groups at the time when they were of eligible age was very low. From about 2010 onwards, the predictions diverge as these children in these age groups would have benefited from more substantial levels of SSCC provision when younger.

In contrast, the model predicts that SSCC provision raised the entry rates for children in age groups g_1 through to g_3: that is, under the counterfactual the predicted entry rates are lower. The gaps in the predictions that appear already at the beginning of the sample period reflect that children aged $0-4$ in 2007 were already benefiting from some positive level of SSCC provision. Over time the gap between the factual and counterfactual predictions increase, as the level of SSCC provision increased.

7 Conclusions

As the number of children taken into care is continuously increasing in England, a growing literature in economics has started exploring the potential causes and consequences of child maltreatment. A key gap in the literature is the impact of public policies targeting children, parents and families on the prevalence of child abuse and neglect, as well as on the resulting need for safeguarding, protection and care.

Sure Start has been a flagship early intervention policy aimed at pre-school children and their parents in England. During the period 2004-2010, over 3,600 children's centres were established providing a wide range of services including child healthcare, childcare, good quality play, parenting support and advice on child health and development. Although SSCC were welcomed by the local communities, since 2011, more than 500 centres were closed. Ongoing policy debate on potential effects of SSCC closures, including the possibility of affecting the number of children entering care, has been taking place.

The theoretical impact of SSCC on entry rates is ambiguous. Early intervention, through supporting families, reduce current and future need for care (investment effect). At the same time SSCC provision may lead to identifying cases where care is needed, leading to additional or earlier care (identification effect). This paper studies the impact of SSCC provision, as a key early intervention programme, on children's rate of entry into care by exploiting the variation in SSCC provision across LAs and, more importantly, across cohorts. Our identification strategy allows for separate impacts on currently eligible children (aged 4 years old and younger), and children eligible in the past (children older than 4 years old). We control for year effects, age group effects and trends, permanent unobserved LA characteristics, LA-specific trends, time-varying LA characteristics and baseline LA characteristics interacting with time.[12]

Our findings suggest that SSCC impact varies substantially across age groups. The estimated effects of SSCC are strongly positive and statistically significant for children within the policy target age. In contrast, for past eligibility groups, the estimated effects are negative but smaller in absolute terms and not always statistically significant. Differencing the estimating equation between any two age groups produced consistent results and showed that the estimates tend to decrease with age, both in terms of range and average.

Our results indicate the presence a positive investment effect on children who have passed eligibility, as well as a significant role of SSCC provision in helping identifying children at risk of harm. Our paper presents the first empirical evidence related to the ongoing discussion on the relationship between Sure Start Children's Centres and the number of children entering care in England.[13] At the same time it contributes to the broader literatures on child abuse and neglect, early intervention policies targeting young children from disadvantaged backgrounds, and child development.

---

[12] Baseline LA characteristics are variables measured in 2004, i.e. at the beginning of the SSCC expansion period.

[13] Carpenter et al. (2007) also studied the relationship between the Sure Start intervention and the demand for children's safeguarding but focused on the impact of SSLP on child protection registrations. Their analysis of data concerning 10,000 referrals of children under four yeas old and 1,600 child protection registrations in four LAs revealed no short-term impact of SSLP on child protection registrations.

References

Avellar, S. A., & Supplee, L. H. (2013). Effectiveness of home visiting in improving child health and reducing child maltreatment. *PEDIATRICS*, *132*(Supplement), S90–S99. Retrieved 2020-03-20, from http://pediatrics.aappublications.org/cgi/doi/10.1542/peds.2013-1021G

Bate, A., & Foster, D. (2017). Sure Start (England). Retrieved from https://commonslibrary.parliament.uk/research-briefings/cbp-7257/

Brown, D., & De Cao, E. (2018). The impact of unemployment on child maltreatment in the United States. *ISER Working Paper Series*(2018-04). Retrieved from http://hdl.handle.net/10419/200368

Carneiro, P., & Ginja, R. (2014). Long-term impacts of compensatory preschool on health and behavior: Evidence from Head Start. *American Economic Journal: Economic Policy*, *6*(4), 135-73. Retrieved from http://www.aeaweb.org/articles?id=10.1257/pol.6.4.135

Carpenter, J., Brown, S., & Griffin, M. (2007). Prevention in integrated children's services: the impact of Sure Start on referrals to social services and child protection registrations. *Child Abuse Review*, *16*(1), 17–31. Retrieved from http://doi.wiley.com/10.1002/car.969

Cattan, S., Conti, G., Farquharson, C., & Ginja, R. (2019). The health effects of Sure Start. Retrieved from https://www.ifs.org.uk/uploads/The-health-effects-of-Sure-Start-Exec-Summary.pdf

Children's Commissioner for England. (2015). Protecting children from harm. Retrieved from https://www.childrenscommissioner.gov.uk/wp-content/uploads/2017/06/Protecting-children-from-harm-full-report.pdf

Conti, G., Heckman, J. J., & Pinto, R. (2016). The effects of two influential early childhood interventions on health and healthy behaviour. *The Economic Journal*, *126*(596), F28–F65. Retrieved from https://academic.oup.com/ej/article/126/596/F28-F65/5077840

Currie, J., & Almond, D. (2011). Human capital development before age five. In *Handbook of Labor Economics* (Vol. 4, pp. 1315–1486). Elsevier. Retrieved from https://linkinghub.elsevier.com/retrieve/pii/S0169721811024130

Department for Children, Schools and Families. (2008). The Sure Start journey. Retrieved from https://dera.ioe.ac.uk/8110/7/FINAL%20The%20Sure%20Start%20Journey_Redacted.pdf

Doyle, J. J., & Aizer, A. (2018). Economics of child protection: Maltreatment, foster care, and intimate partner violence. *Annual Review of Economics*, *10*(1), 87–108. Retrieved from https://www.annualreviews.org/doi/10.1146/annurev-economics-080217-053237

Garces, E., Thomas, D., & Currie, J. (2002). Longer-term effects of Head Start. *American Economic Review*, *92*(4), 999-1012. Retrieved from http://www.aeaweb.org/articles?id=10.1257/00028280260344560

Howard, K. S., & Brooks-Gunn, J. (2009). The role of home - visiting programs in preventing child abuse and neglect. *The Future of Children*, *19*(2), 119–146. Retrieved from http://muse.jhu.edu/content/crossref/journals/future_of_children/v019/19.2.howard.html

Lindo, J. M., Schaller, J., & Hansen, B. (2013). Economic conditions and child abuse. Retrieved from http://ftp.iza.org/dp7355.pdf

Melhuish, E., Belksy, J., & Leyland, A. (2010). The impact of Sure Start Local Programmes on five years olds and their families. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/182026/DFE-RR067.pdf

Melhuish, E., Belksy, J., & Leyland, A. (2012). The impact of Sure Start Local Programmes on seven year olds and their families. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/495329/The_impact_of_Sure_Start_local_programmes_on_7-year-olds_and_their_families.pdf

Melhuish, E., Belsky, J., Leyland, A. H., Barnes, J., & the National Evaluation of Sure Start Research Team. (2008). Effects of fully-established Sure Start Local Programmes on 3-year-old children and their families living in England: a quasi-experimental observational study. , *372*. Retrieved from https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(08)61687-6/fulltext

National Audit Office. (2014). Department for Education - Children in care. Retrieved from https://www.nao.org.uk/report/children-in-care/

Paxson, C., & Waldfogel, J. (1999). Parental resources and child abuse and neglect. *The American Economic Review*, *89*(2), 239–244. Retrieved from http://www.jstor.org/stable/117113

Paxson, C., & Waldfogel, J. (2002). Work, welfare, and child maltreatment. *Journal of Labor Economics*, *20*(3), 435-474. Retrieved from https://doi.org/10.1086/339609

Paxson, C., & Waldfogel, J. (2003). Welfare reforms, family resources, and child maltreatment. *Journal of Policy Analysis and Management*, *22*(1), 85–113. Retrieved from http://www.jstor.org/stable/3325847

Sammons, P., Hall, J., Smees, R., & Goff, J. (2015). The impact of children's centres: studying the effects of children's centres in promoting better outcomes for young children and their families. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/485346/DFE-RR495_Evaluation_of_children_s_centres_in_England__the_impact_of_children_s_centres.pdf

Sandner, M., & Thomsen, S. L. (2018). The effects of universal public childcare provision on cases of child neglect and abuse. *IZA Discussion Paper Series*(11687). Retrieved from http://ftp.iza.org/dp11687.pdf

Sedlak, A. J., Mettenburg, J., Basena, M., Petta, I., McPherson, K., Greene, A., & Li, S. (2010). Fourth National Incidence Study of Child Abuse and Neglect (NIS-

4): Report to Congress. Retrieved from https://www.acf.hhs.gov/opre/resource/fourth-national-incidence-study-of-child-abuse-and-neglect-nis-4-report-to

Smith, G., Sylva, K., Smith, T., Sammons, P., & Omonigho, A. (2018). Stop Start - Survival, decline or closure? Children's Centres in England. Retrieved from https://www.suttontrust.com/our-research/sure-start-childrens-centres-england/

The Economist. (2018). The troubling surge in English children being taken from their parents. Retrieved from https://www.economist.com/britain/2018/03/22/the-troubling-surge-in-english-children-being-taken-from-their-parents

Thomas, C. (2018). The Care Crisis Review: Factors contributing to national increases in numbers of looked after children and applications for care orders. Retrieved from https://www.frg.org.uk/involving-families/reforming-law-and-practice/care-crisis-review

Figures

Figure 1: Number of children entering care and in care by year

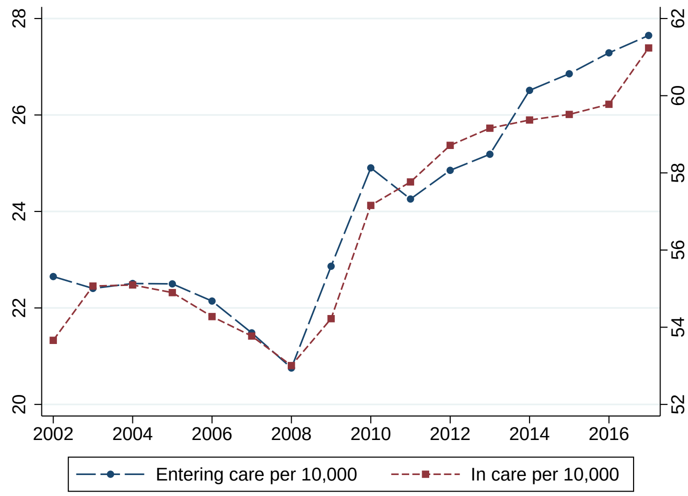


Figure 2: Trends in care entry rate by demographic subgroup (index 2006 = 100)

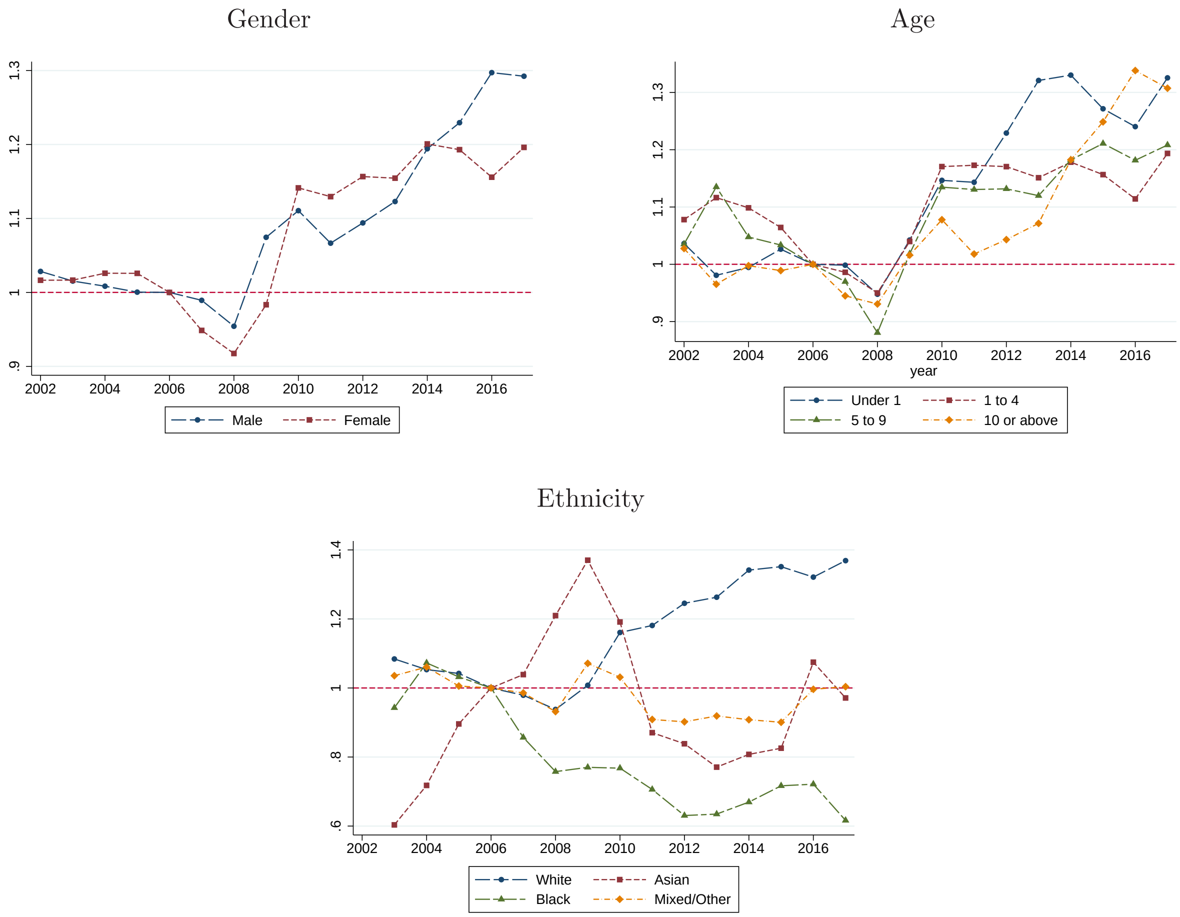


Figure 3: Care entry rate by reason for care

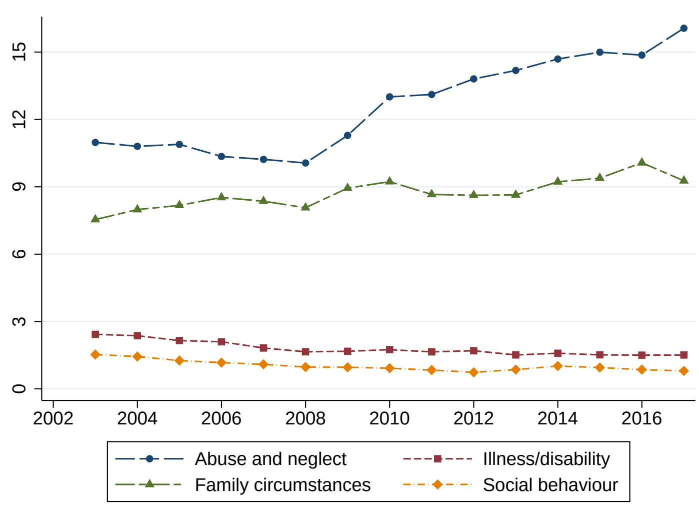


Figure 4: Variation in care entry rates across cells defined by Local Authority and year

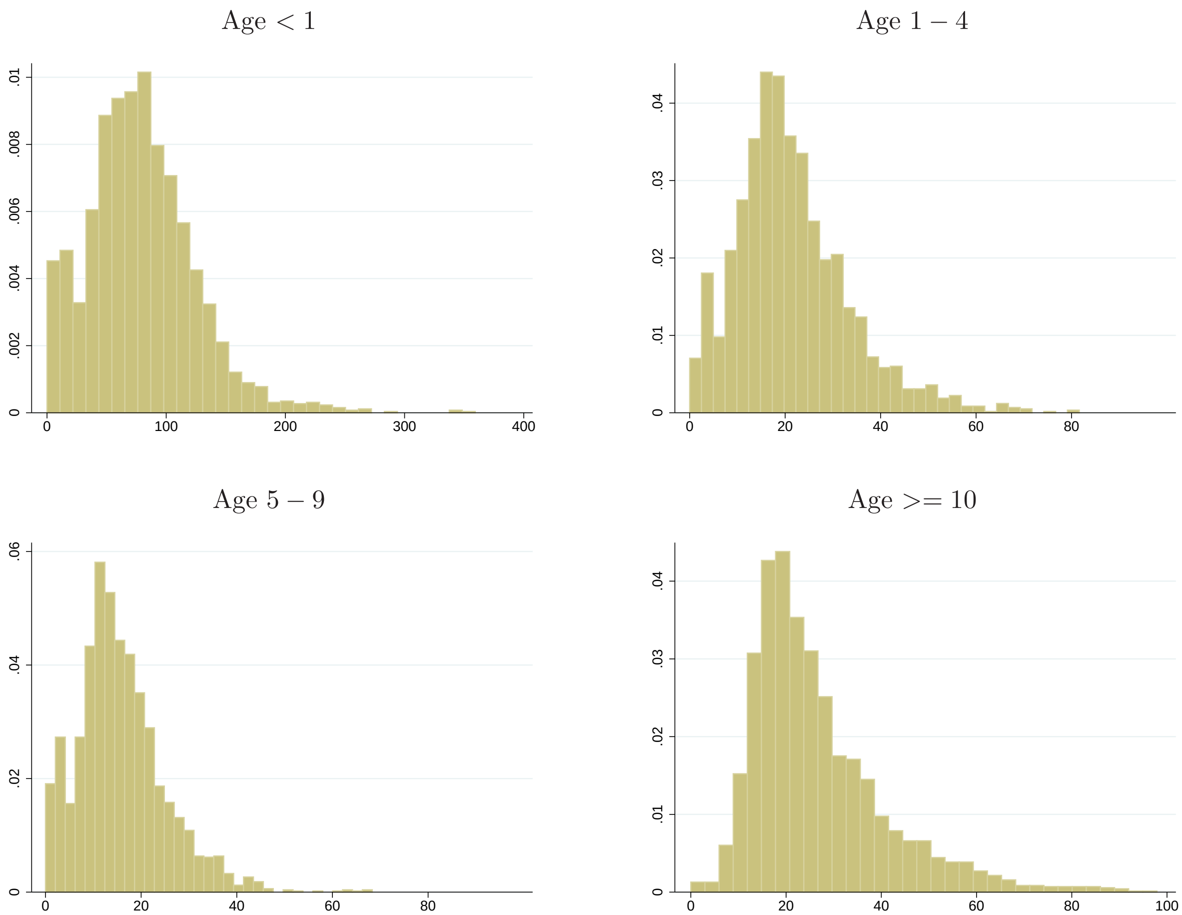


Figure 5: Number of SSCC in operation in England by year, and number of SSCC per 1,000 children aged 0-4

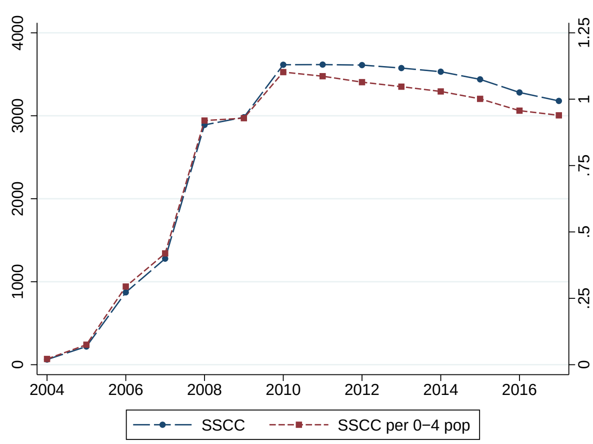


Figure 6: Evolution of SSCC coverage variation

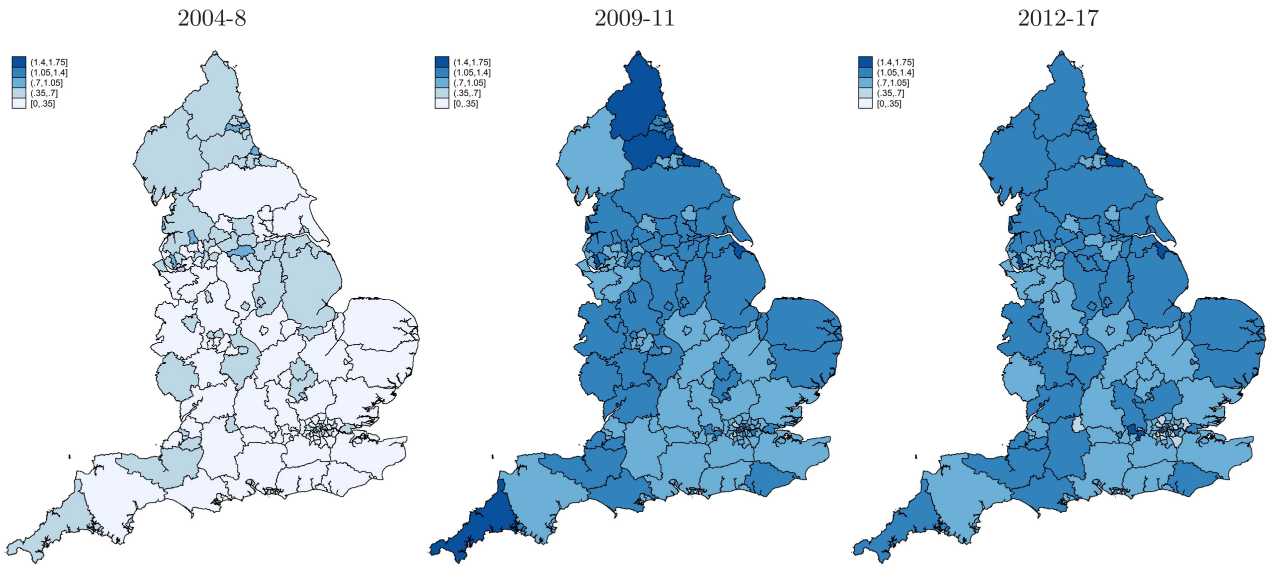


Figure 7: SSCC coverage: different expansion paths by local characteristics

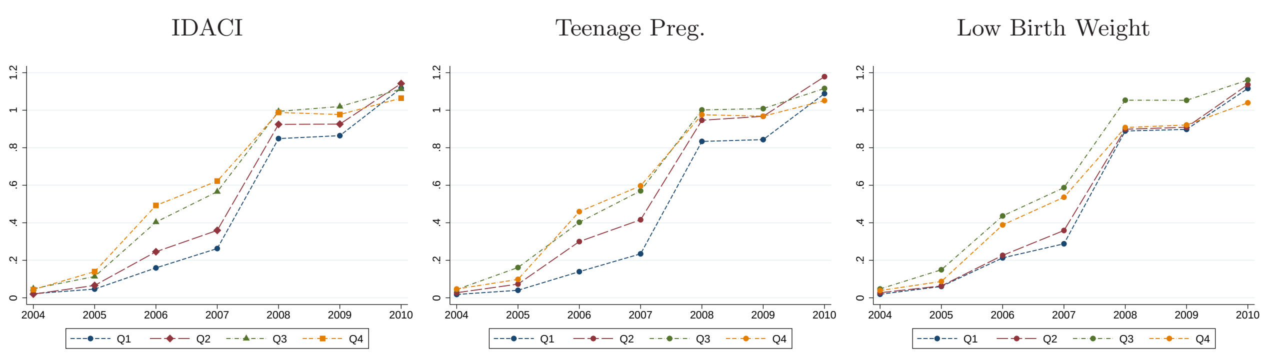


Figure 8: Relating relative entry rates to SSCC expansion

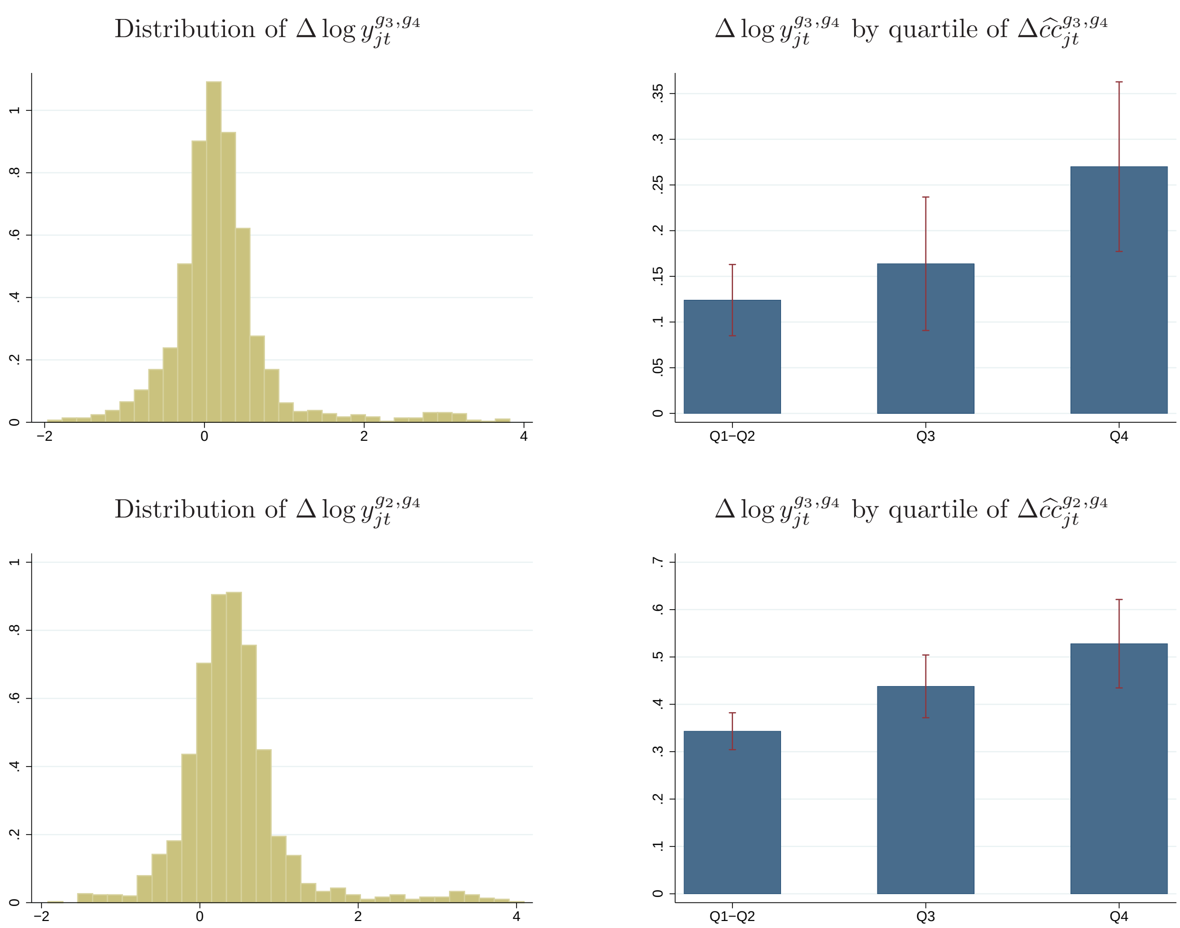
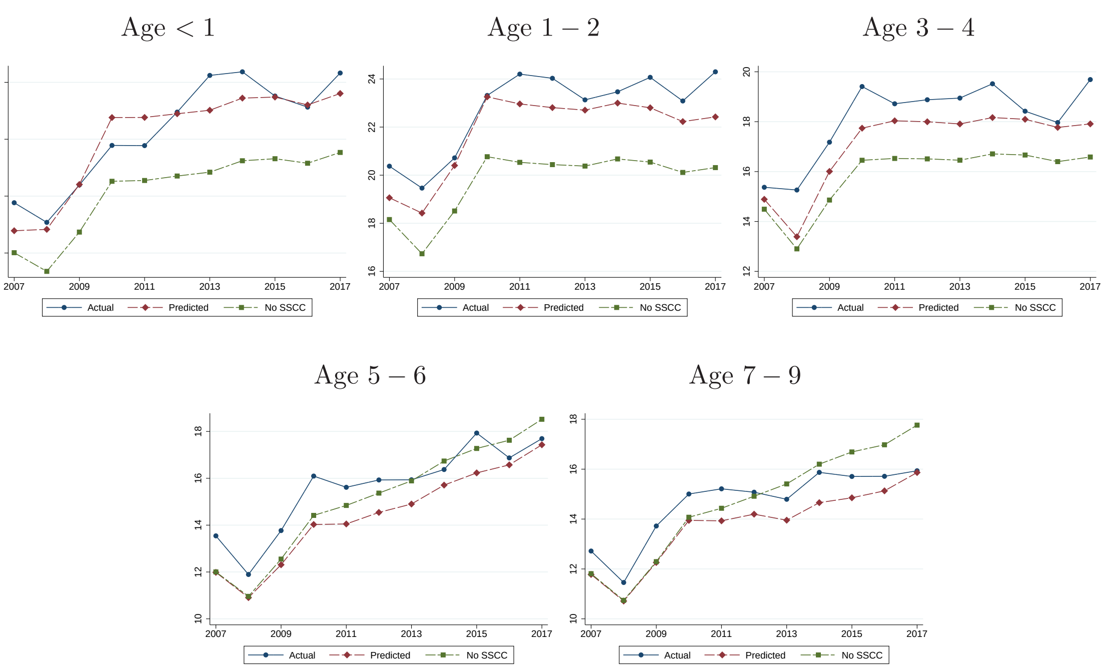


Figure 9: Factual and counterfactual entry rates



A Appendix A: Local Authorities' characteristics

This section presents detailed information on all control variables used in the analysis. The sample consists of 147 LAs in England over the period 2007-2017, while baseline characteristics refer to 2004 unless otherwise specified.[14]

A.1 Time-varying characteristics

- **Ethnicity**: Ethnicity is measured as the proportion of the adult population with each of the following ethnic backgrounds: White, Black, Asian and Mixed and Other. The data is provided by the Office for National Statistics (ONS) through the Nomis website and their original source is the Annual Population Survey (APS).

- **Economic activity rate**: Economic activity rate shows the percentage of people aged at least 16 years old who are economically active. As above, the data was provided by ONS through the Nomis platform, while their original source is APS.

- **Claimants proportion**: Claimants proportions is defined as the proportion of the population claiming benefits principally for the reason of being unemployed out of all people aged 16-64 years old. The data was provided by ONS through the Nomis platform and it was collected for the last month of each financial year (March).

- **Education qualifications**: Our indicators show the proportion of the population aged 16-64 owning a qualification within the following broad categories: university/college degree or above, GCE/GCSE or equivalent, any qualification of a lower level than the ones stated above. The data was collected through the Nomis platform (ONS), and their original source is the APS.

- **Income**: Our analysis uses the median annual gross income of full-time workers. The original data source is the Annual Survey of Hours and Earnings, and the data was accessed through the Nomis platform. The data was adjusted for inflation.

[14] City of London, Rutland and Isles of Scilly are not included in our sample due to very small numbers and missing data in many cases. Bedford Borough and Central Bedfordshire are merged together, and Cheshire East and Cheshire West and Chester are also merged together.

- **Political control**: The political party controlling each LA at each point in time is included in our analysis with the main categories being: the Conservative Party, the Labour Party, Liberal Democrats, and no overall control. The Elections Centre provides data on annual political composition of all local councils.

A.2 Baseline characteristics

- **Income Deprivation Affecting Children Index (IDACI)**: IDACI measures the proportion of all children aged 0 to 15 living in low income families. Our variable represents the percentage of LSOAs in each LA that are in the top percentile of the IDACI distribution. IDACI is provided by the Ministry of Housing, Communities and Local Government (MHCLG) as part of the Index of Multiple Deprivation.

- **Low-birth-weight rate**: The Low-birth-weight rate is defined as the proportion of low-birth-weight babies ($< 2,500$g) of 37 weeks or more gestation ("full-term") out of all full-term live births. The data is is published by the NHS England and the ONS.[15]

- **Teenage pregnancy**: Teenage pregnancy is measured as the number of pregnancies of girls younger than 18 years old during each financial year per 1,000 population of girls aged 15 to 17. Teenage conception data is provided by the ONS through the quarterly publication "Quarterly conceptions to women aged under 18".

- **Gross fertility rate** : The gross fertility rate is the number of live births per 1,000 women aged 15-44. Annual data on fertility are included in the ONS Birth Summary Tables.

- **Income Inequality**: We define income inequality as the ratio of the 80th percentile over the 20th percentile of the weekly income distribution in each LA. Income data was collected through the Nomis platform, while the original source is the Annual Survey of Hours and Earnings.

- **Exclusions**: Our analysis controls for school exclusions by using the ratio of the total number of fixed period exclusions and the number of all pupils in each LA. Data on

[15] No data has been found for 2004, so data from the 2005 calendar year were used instead.

schools exclusions are provided annually by the Department for Education through their publication "Permanent and fixed period exclusions in England ".

The data on care entry rates are provided on an annual financial year basis (i.e. April to March next year). Consequently, our priority was to collect data on the above variables based on financial years. Where this was not possible we used the period overlapping the most with the financial year.

Table A1 presents the descriptive statistics of LA characteristics described in this section over LA and year cells (baseline characteristics' statistics are analysed only over LAs and not over time).

Table A1: Descriptive statistics of Local Authority characteristics.

Panel A: Baseline Characteristics		
Variable	**Mean**	**St. Dev.**
IDACI (2004)	12.66	14.10
Low birth weight (2004)	3.12	1.14
Teenage pregnancies (2004)	45.10	14.31
Gross fertility rate (2004)	57.61	6.12
Inequality (2004)	2.28	0.24
Exclusions (2004)	4.59	1.73
Observations	147	

Panel B: Time Variant Characteristics		
Variable	**Mean**	**St. Dev.**
White	85.78	15.14
Black	3.48	5.02
Asian	5.98	8.06
Mixed-Other	4.68	4.70
Ec. Activity Rate	63.51	3.84
Claimant Prop.	3.00	1.51
Prop. High Qual.	24.66	9.86
Prop. Medium Qual.	52.72	8.97
Prop. Low Qual.	22.71	6.82
Median Income	29,999	4,481
Cons. control	0.436	0.496
Lib. Dem. control	0.032	0.177
Labour control	0.320	0.467
No overall control	0.210	0.407
Observations	1,617	

Notes: See text for variable definitions.

B Appendix B: Additional analysis

B.1 SSCC coverage: modelling the expansion period

This section presents a background econometric analysis on the SSCC coverage during the period of SSCC expansion, i.e. from 2004 to 2010, aiming to understand systematic differences in SSCC coverage across areas with different characteristics.

We developed a model predicting the SSCC coverage in each LA at the end of each year based on LA baseline characteristics in 2004 and time-varying political composition variables. Those characteristics were selected based on the SSLP and SSCC focus and aims. For example, the 2017 House of Commons Briefing Paper on Sure Start stated that the first 60 Sure Start districts were chosen based on high levels of deprivation and existing good practice in early years provision (Bate & Foster, 2017). Cattan et al. (2019) mention that SSLP targeted disadvantaged areas with high levels of teenage pregnancy and low birth weight. Phase 1 SSCC targeted 20 percent most disadvantaged areas and Phase 2 SSCC focused on 30 percent most disadvantaged areas. Phase 3 SSCC were focused on covering areas with lower provision, and thus they are expected to have been built on reverse criteria. The final list of SSCC predictors used is presented in Appendix A.2, apart from the additional time-varying political composition variables described in Appendix A.1.

Our model is presented in Equations (B1) and (B2).

$$\log\left(cc_{jt}\right) = \beta Z_j^b + \gamma Z_j^b t + \rho Z_{jt} + \eta_t + \varepsilon_{jt} \tag{B1}$$

whereby

$$\frac{\partial \log\left(cc_{jt}\right)}{\partial Z_j^b} = \frac{\partial cc_{jt}/\partial Z_j^b}{cc_{jt}} = \beta + \gamma t \tag{B2}$$

The dependent variable cc_{jt} denotes the number of children's centres per 1,000 children aged 0-4. Z_j^b includes all the area characteristics at baseline. Time is defined as $t = 0$ for 2004, $t = 1$ for 2005, etc., and thus $Z_j^b t$ includes interaction terms of baseline area characteristics with time. Finally, the model includes time-varying political control variables defined as Z_{jt} .

Our analysis uses baseline variables instead of time-varying ones to capture how the decision of expansion was formed based on the information available in the beginning of the intervention. Our methodology is consistent with the methodology used by Cattan et al. (2019). Additionally, those LA characteristics show very little variation over time within LAs. Interaction terms between baseline characteristics and time are used to explore the hypothesis that SSCC predictors defined mainly the provision in the first two phases of the expansion period, with Phase 3 targeting areas with lower provision.

Table B1 shows the results of the above analysis. The results are consistent with what is described in the literature. Areas with high deprivation, teenage pregnancy rates, and low-birth-weight rates in 2004 had higher SSCC coverage, but, as expected, the impact was decreasing over time. The results also show that LAs controlled by the Conservative and Liberal Democrats parties had significantly lower SSCC coverage during the period 2004-2010.

Cattan et al. (2019) in their study of the impact of SSCC on children's health outcomes present results in line with our findings. Their dataset is at a Lower Super Output Area level and thus the magnitude of the coefficients is not directly comparable but the direction of impact has the same interpretation. The paper shows that being in the 20-30 percent most deprived LSOAs is correlated with higher SSCC coverage. Low birth weight and teenage conception are also positively correlated with SSCC coverage, but their impact is much smaller and not always significant. Finally, LSOAs in Local Authorities aligned with the national government (Labour during the whole expansion phase) tended to open centres earlier and to have higher coverage.

Table B1: SSCC prediction analysis

	(1)	(2)	(3)
IDACI Index (2004)	0.0164***	0.0219***	0.0209***
	(0.00407)	(0.00427)	(0.00429)
Teenage Preg. Rate (2004)	0.0216***	0.0184***	0.0156***
	(0.00397)	(0.00440)	(0.00443)
LBW rate (2004)	-0.0136	0.0926*	0.0732
	(0.0427)	(0.0459)	(0.0459)
IDACI-04 × time	-0.00275*	-0.00353**	-0.00373**
	(0.00110)	(0.00116)	(0.00115)
TPR-04 × time	-0.00309**	-0.00262*	-0.00264*
	(0.00108)	(0.00119)	(0.00119)
LBW-04 × time	-0.00108	-0.0161	-0.0145
	(0.0115)	(0.0123)	(0.0122)
Gr. Fertility Rate (2004)		-0.0475***	-0.0442***
		(0.00756)	(0.00766)
Inequality (2004)		-0.299	-0.139
		(0.227)	(0.231)
School Exclusion Rate (2004)		0.0153	0.0141
		(0.0261)	(0.0262)
GFR-04 × time		0.00695***	0.00621**
		(0.00205)	(0.00204)
Inequality-04 × time		0.0687	0.0492
		(0.0622)	(0.0623)
SER-04. × time		0.00232	0.00199
		(0.00706)	(0.00707)
Cons. Control			-0.265***
			(0.0679)
Lib Dem control			-0.220*
			(0.0986)
No Overall Control			-0.0626
			(0.0610)
Constant	-5.082***	-1.974*	-2.226**
	(0.170)	(0.780)	(0.786)
Year Fixed Effects	Y	Y	Y
Observations	1,029	1,029	1,026
Adjusted R^2	0.830	0.838	0.840

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The unit of observation is an LA-year cell. The dependent variable is the SSCC coverage rate, defined as the number of operating SSCC per 1,000 children aged 0-4. The years included are 2004-2010. For variable definitions, see Appendix A.

B.2 SSCC closures: survival analysis

As discussed in Section 4.2, SSCC closures were concentrated in specific areas (five LAs accounted for more than a third of all closures that took place during the period 2011-2017). Consequently, the link between area characteristics and SSCC closures is expected to be substantially weaker than the link with the timing of SSCC expansion. Additionally, Table 3 shows that Phase 3 centres had higher probability of closure than the rest of the centres. Consequently, closures might be more prominent in less deprived areas.

In this section, we use survival analysis to predict the probability of a closure for each children's centre. The estimation equation is shown below:

$$cens_{it} = \beta Z_i^b + \gamma Z_{jt} + \eta_t + \varepsilon_{jt} \tag{B3}$$

The dependent variable is a binary variable which follows this logic: If centre i 's survival time is censored (i.e. if it is opened until the 31st March 2017, the last day we are examining), the binary dependent variable is equal to 0 for all of i 's time spells; if centre i 's survival time is not censored (i.e. it has closed down), the binary dependent variable is equal to 0 for all but the last of i 's spell. The rest of the variables follow the logic of SSCC expansion analysis, with the difference that the baseline year in this part of analysis is 2010.

Table B2 shows the results of the survival analysis. As expected, there is a very weak relationship between LA characteristics and the probability of closure of a children's centre. Nevertheless, the predictors of SSCC expansion seem to have the opposite effect on closures (e.g. deprivation is positively correlated with expansion and negatively correlated with closures), as expected, but the magnitude of the impact is very small.

Table B2: SSCC survival analysis

	(1)	(2)	(3)
IDACI Index (2010)	-0.0000841	-0.0000955	-0.000162+
	(0.0000858)	(0.0000862)	(0.0000916)
Teenage Preg. Rate (2010)	0.0000962	0.000206+	0.000142
	(0.000102)	(0.000106)	(0.000118)
LBW rate (2010)	-0.00783***	-0.00845***	-0.00886***
	(0.00161)	(0.00179)	(0.00184)
Gr. Fertility Rate (2010)		0.0000897	0.0000441
		(0.000132)	(0.000131)
School Exclusion Rate (2010)		-0.00281***	-0.00268***
		(0.000780)	(0.000781)
Lib Dem control			-0.0126
			(0.00839)
Labour control			0.00526+
			(0.00308)
No Overall Control			-0.00363
			(0.00276)
Constant	0.0234***	0.0280**	0.0348***
	(0.00566)	(0.00922)	(0.00946)
Year	Y	Y	Y
Observations	24,126	24,126	23,693

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The unit of observation is an open SSCC in a given year. The dependent variable is an indicator for the SSCC closing over the following 12 month period. For variable definitions, see Appendix A. Baseline variables are measured in 2010. Political control variables are time-varying.

B.3 Additional estimated coefficients

Table B3 presents further estimated coefficients from specifications (1) - (3) in Table 7.

Table B3: Alternative specifications: further estimated coefficients

	(1)	(2)	(3)
Black	-0.00935	-0.0168*	-0.00599
	(0.00855)	(0.00823)	(0.00547)
Asian	0.00175	-0.00371	-0.00219
	(0.00677)	(0.00696)	(0.00313)
Mixed-Other	-0.0129+	-0.0165*	0.00206
	(0.00709)	(0.00816)	(0.00634)
Ec.Act. Rate. 16+	-0.00708	-0.0221***	-0.0159**
	(0.00532)	(0.00522)	(0.00500)
Claimant Prop.	0.00325	0.0237	0.0179
	(0.0307)	(0.0272)	(0.0195)
High Qualified	0.00361	-0.00438	0.00272
	(0.00512)	(0.00515)	(0.00435)
Medium Qualified	-0.0000678	-0.00617	0.00872+
	(0.00441)	(0.00544)	(0.00452)
Income	-0.00000933	0.000000844	-0.0000135*
	(0.0000112)	(0.0000110)	(0.00000621)
Cons. Control	-0.00810	0.0211	0.0294
	(0.0590)	(0.0425)	(0.0419)
Lib. Dem. Control	-0.0361	0.0766	0.0501
	(0.0729)	(0.0505)	(0.0706)
No Overall Control	0.00319	0.0754*	0.0798*
	(0.0461)	(0.0360)	(0.0376)
IDACI (2004) × time		-0.000286	-0.000502
		(0.000359)	(0.000366)
Teenage Preg. Rate (2004) × time		0.000172	0.000192
		(0.000390)	(0.000393)
Gr. Fertility Rate (2004) × time		-0.00113	-0.00112
		(0.000691)	(0.000691)
Low Birth Weight (2004) × time		-0.000713	-0.00180
		(0.00276)	(0.00281)
Exclusions (2004) × time		-0.00389	-0.00384
		(0.00241)	(0.00241)
Inequality (2004) × time		-0.0397+	-0.0422+
		(0.0230)	(0.0240)
IDACI (2004)			0.00729*
			(0.00287)
Teenage Preg. Rate (2004)			0.0138***
			(0.00287)
Gr. Fertility Rate (2004)			0.000614
			(0.00560)
Low Birth Weight (2004)			0.0406
			(0.0249)
Exclusions (2004)			0.0164
			(0.0166)
Inequality (2004)			0.0829
			(0.142)
Observations	7,930	7,930	7,930

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The specifications in the current table correspond to specifications (1)-(3) in Table 7.