The Effect of Free Elderly Personal Care on Informal Caregiving and Labour Market Participation: Revisiting the Scottish Reform

Tilman F. Bretschneider and Joel Cheung

Abstract

We revisit the Scottish care reform of 2002 to estimate the effect of free formal personal care for the elderly on informal caregiving and labour market participation using a difference-in-differences approach. We find that previous studies' results suffer from bias due to violations of identifying assumptions. As a result, the effects on informal care are inconclusive. Nonetheless, we identify a causal effect of the reform on employment.

Keywords: Informal Care, Labour Supply, Difference-in-differences

NEKP01: Master's Essay II

Department of Economics

Supervisor: Ulf Gerdtham

Date of submission: 24/05/2023

Contents

1 Introduction

The winds of demographic change are increasingly straining health and social care systems of developed countries. This dynamic continues to grow due to increased life expectancy and decreased fertility rates (He et al., 2016). Final years' healthcare costs make up the lion's share of an individual's lifetime healthcare expenditures (Alemayehu & Warner, 2004). Besides increasing healthcare budgets, a growing share of the elderly population also results in a larger need for social care.

The rise in demand for elderly personal care causes costs for society as a whole. The demand is met either by informal care, typically provided by family or friends, or formal care provided by paid professional caregivers. Formal care entails direct costs. Costs of informal care are indirect, caregivers face opportunity costs of providing informal care through forgone earnings and leisure time (Karlsberg Schaffer, 2015; Schmitz & Westphal, 2017).

Governments may subsidise formal care in order to improve elders' health. However, positive effects may be undermined by the substitution of informal care. Conversely, if they act as complements, then the benefits of increasing formal care for elders may be two-fold. Additionally, a formal care policy might also be aimed at increasing the labour market participation of informal caregivers, through the substitution of informal care for labour. This might partially fund the formal care subsidy through higher income tax revenue. Desired policy outcomes thus rely on the rate of substitution between formal and informal care, and the subsequent substitution of informal care for labour.

We analyse the effect of the expansion of state-funded elderly care on informal care provision and labour market participation. Therefore, we exploit a "natural experiment" created by the Scottish elderly care reform of 2002. The reform was not enacted in the rest of the UK allowing the use of a difference-in-differences (DID) approach.

There exists extensive literature dealing with the substitution of formal and informal care. Recent contributions focus on the endogeneity of care decisions which can result from unobserved factors such as family ties and norms (Bolin et al., 2008a). Employing European panel data and using child characteristics as instrumental variables (IV), Bolin et al. (2008a) find home informal and formal care are substitutes with a specific nuance that results differ by the European north-south gradient implying heterogeneity of cultural norms. Bonsang (2009) who uses an IV approach also finds that formal and informal care

are substitutes, however, substitution effects decrease with increasing degrees of care-receiver disability.

Urwin et al. (2019) investigate if increases in informal care reduce formal care use. They use UK data and child characteristics as IVs finding that an increase in monthly informal care reduces the probability of using formal care. Interestingly, substitution effects were larger when formal care was state funded.

Also, the relationship between informal care and labour market participation has been subject to recent empirical studies. Bolin et al. (2008b) estimate the effect of informal care hours provided by caregivers on their employment probability, work hours, and wages. Using IVs, they find that providing informal care reduced employment both at the extensive margin (probability of employment) and the intensive margin (hours worked).

Using UK data and a pension reform for identification, Carrino et al. (2023) investigate the impact of increased labour hours on informal care provision. Their findings reveal that an increase in work hours reduced informal care. Notably, this reduction in informal care was not compensated by an increase in self-funded formal care, suggesting the substitution effects are not necessarily bidirectional.

Similarly to our study, studies using a formal care expansion reform as an exogenous source of variation in formal care, explore labour participation outcomes. Løken et al. (2017) follow a DID approach using a Norwegian formal care funding reform to explore how the expansion of formal care influences the labour participation decisions of informal caregivers. They measure a reduction in work absences among single daughters due to the expansion of publicly provided formal home-care services.

There are three papers that investigate the effect of the 2002 Scottish reform on informal caregiving and carer's labour market outcomes which all use a DID approach. Bell et al. (2007) and Karlsberg Schaffer (2015) both use the British Household Panel Survey (BHPS) while Hollingsworth et al. (2022) use the UK Family Resources Survey (UKFRS).

Using data from 1999 to 2003, Bell et al. (2007) find that the reform did not significantly affect informal care provision. Karlsberg Schaffer (2015) extend the time horizon (1996-2008) and conversely find a large increase in the likelihood of informal care provision, this effect is observed primarily among female caregivers. These results imply that informal and formal care act as complements in contrast to previous literature that finds substitution effects (see e.g. Bolin et al., 2008a; Urwin et al., 2019). The paper also employs a theoretical framework where the caregiver's utility is a function of leisure time and the care receiver's utility which in turn is a function of total care hours received. In their model, the formal care reform expands the production possibilities frontier which shows that increasing informal care could still maximise the caregiver's utility. However, the caregiver's utility function ignores labour income. It is plausible that the caregiver has incentives to expand labour supply to increase utility derived from income implying a reduction in either informal care provision or leisure time.

Hollingsworth et al. (2022) also investigate the Scottish reform using the UKFRS, a repeated cross-section data set, to explore informal care provision but also extends the investigation to caregivers' labour market participation. They find the reform decreased the likelihood of informal caregiving, indicating a substitution effect from informal to state-funded formal care, contradicting the findings of Karlsberg Schaffer (2015). Additionally, the study reveals that caregivers increase employment at both the intensive and extensive margins.

This paper revisits the Scottish care reform by examining its effect on informal caregiving and labour market participation. The focus is on understanding why the results of Karlsberg Schaffer (2015) and Hollingsworth et al. (2022) differ and to synthesise which estimates are reliable. We do this by investigating the parallel trends assumption which is key to any DID framework. Due to the longer time horizon, BHPS data is better suited for this task such that we focus on Karlsberg Schaffer (2015) while extending the study to investigate labour participation outcomes.

We are able to replicate the findings of Karlsberg Schaffer (2015) that informal care increased in response to the reform, however, we find strong evidence of a violation of the parallel trends assumption that suggests an upward bias of the positive results. Furthermore, we provide evidence that the reform caused an increase in employment.

2 Institutional Background

In the United Kingdom, formal personal care costs expose individuals to significant financial pressures, with those receiving care in England paying an average of £4,742 per year in 2001 (Hollingsworth et al., 2022).

In 1997, the UK government established the Royal Commission on Long-Term Care to evaluate the provision of care for the elderly and develop a sustainable funding system for long-term care. The resulting Sutherland report concluded that the existing system required reform and proposed a new system where personal care costs would be separated from living and housing costs and funded through general taxation. A key recommendation of the Sutherland report was the provision of free formal personal care for those aged 65 and over (Audit Scotland, 2008).

As part of the 1999 UK devolution arrangements, the UK government established the Scottish Parliament and granted Scotland the power to legislate on issues including education, health, and justice. The newly formed Scottish Parliament used these newfound powers to pass the Community Care and Health (Scotland) Act in 2002. The policy provisioned free personal care to the elderly achieved by abolishing means testing from formal care eligibility. Thus, subject to a needs assessment, the elderly became eligible for free personal care at home free of charge.

Personal care is a form of social care that is defined as physical and mental tasks related to day-to-day care such as assistance with personal hygiene, at mealtimes, immobility problems, medication and well-being (Audit Scotland, 2008).

The other UK nations did not implement the policy change, thus Scotland become the sole UK nation to provide home-based personal care to the elderly without means testing. This created a "natural experiment" whereby Scotland received the treatment while the rest of the UK nations (England, Wales and Northern Ireland) did not.

Prior to the reform, personal care provisions for the Scottish elderly varied depending on the individual's financial circumstances and their local authority's discretion. In many cases, individuals were required to contribute towards the cost of their care, either through direct payments or by liquidating assets such as their homes to pay for care costs. It is worth noting that before the development of the national eligibility criteria in 2008, there were no uniform needs assessments across local authorities, and services available to an individual differed depending on their place of residence.

The reform was a significant achievement for the newly devolved Scottish government and received political and media attention (e.g BBC, 2001; Inman, 2002). The policy saw significant uptake resulting in substantial costs (Karlsberg Schaffer, 2015).

3 Data

We utilise the BHPS, an annual survey consisting of a representative sample of around 5,500 households and 10,000 individuals (University of Essex, 2018). Our analysis covers the years 1991–2008, 11 years before and 6 years after the policy reform. The survey boosted sampling by 1,500 and 2,000 households for Scotland and Wales in 1999, and in Northern Ireland in 2001, observable table 1.

We have two groups of outcome variables. Variables measuring informal care and labour market participation outcomes at both the intensive and extensive margins. The BHPS includes a variable indicating if an individual provided care for a household member (residential care) and another variable indicating if they care for someone outside their household (extra-residential care). Similarly to Karlsberg Schaffer (2015), we create a third binary variable indicating if an individual provided either form of care. At the intensive margin, we employ an interval-coded variable indicating the number of weekly informal care hours provided.

The BHPS includes two labour market outcome variables; employment status and weekly working hours. We collapse employment status into a binary variable, aggregating all forms of paid employment. Weekly working hours are a cardinal scaled variable.

To replicate the findings of Karlsberg Schaffer (2015), we drop all individuals with children in the household to prevent caregivers of disabled children from being included in the care outcomes sample. Additionally, individuals under the age of 45 are dropped and the time period is restricted to 1996–2008.

The data reveal that caregivers under the age of 45 represent a substantial proportion of caregivers. Therefore, we create an additional sample including also those under 45 years old and expand the time horizon to 1991–2008. We use this sample to investigate the parallel trends assumption and analyses of residential care as the larger sample provides more efficient estimates. This sample for care outcomes is shown in table 1. We reason that child care-receivers are likely to live in the same household as their caregiving parents.

Table 1: Sample sizes for Scotland and the rest of UK.

	Full sample		Care outcomes		Labour outcomes	
	Rest of UK	Scotland	Rest of UK	Scotland	Rest of UK	Scotland
1991	9307	957	5200	539	5503	544
1992	8918	929	4953	508	4678	486
1993	8706	894	4830	482	4600	479
1994	8608	873	4723	478	4532	481
1995	8406	843	4650	464	4464	460
1996	8615	824	4772	439	4578	458
1997	10169	1061	5742	588	5385	556
1998	9978	984	5631	560	5325	522
1999	12240	3393	6989	1935	6511	1839
2000	12129	3548	6840	1959	6500	1929
2001	15501	3513	8713	1973	8377	1957
2002	13560	3254	7643	1840	7430	1832
2003	13309	3189	7586	1827	7303	1830
2004	13024	3097	7418	1771	7194	1814
2005	12862	2778	7320	1616	7039	1600
2006	12713	2706	7300	1617	6934	1545
2007	12320	2631	7122	1567	6734	1503
2008	11946	2473	6954	1485	6501	1413

Notes: Data comes from the British Household Panel Survey. From 1999 onwards the Scottish population is oversampled. The full sample consists of all individuals in the sample. Care outcomes use a sample of all individuals living in a household without children present. The labour outcomes sample consists of all individuals between 25 and 74 years of age. We drop retired individuals, students and long-term disabled/ill from that sample.

Therefore, we do not need to exclude households with children and thus we use the full sample when estimating results for extra-residential care.

We are interested in the reform's effect on labour market outcomes within the labour force. Thus our labour sample follows Hollingsworth et al. (2022) by selecting individuals aged 25–74, who work less than 60 hours weekly and are not retired, full-time students or permanently sick/disabled which we present in table 1. A summary of the different samples and which analysis they relate to is presented in table 6 of the appendix A.

The reform was implemented on the 1st of July 2002. In this year, the majority (77%) of Scottish interviews were conducted after the reform. Therefore, we categorise all Scottish observations from 2002 as treated leaving waves in calendar year format.

We report summary statistics in table 2. For residential care, levels of informal care increased in both regions and are higher in the rest of the UK compared to Scotland in

Table 2: Summary statistics for outcome variables before and after the reform by region.

	Scotland			Rest of UK			Difference-in-differences		
	Mean	SD	Observations	Mean	SD	Observations	Mean	SE	Observations
Residential care									
Before: 1991–2001	0.0585	0.2348	9,425	0.0647	0.2460	60,348			
After: 2002–2008	0.0781	0.2684	10,420	0.0800	0.2714	47,548			
Difference of means	0.0196	0.0036	19,845	0.0153	0.0016	107,896	0.0043	0.0039	127,741
Extra-residential care									
Before: 1991–2001	0.1056	0.3074	16,909	0.1086	0.3112	107,849			
After: 2002–2008	0.1231	0.3286	17,839	0.1124	0.3159	83,545			
Difference of means	0.0175	0.0034	34,748	0.0038	0.0015	191,394	0.0137***	0.0037	226,142
Employment									
Before: 1991–2001	0.8372	0.3692	9,526	0.8615	0.3454	60,066			
After: 2002–2008	0.8642	0.3426	10,966	0.8714	0.3347	48,545			
Difference of means	0.027	0.0050	20492	0.0099	0.0021	108,611	0.0171***	0.0054	129,103
Weekly working hours									
Before: 1991–2001	24.5325	16.9437	9,544	24.5726	17.2112	60,286			
After: 2002–2008	25.3461	16.4850	10,968	25.2937	16.7700	48,566			
Difference of means	0.8136	0.2342	20,512	0.7211	0.1035	108,852	0.0925	0.2561	129,364

Notes: We present summary statistics of outcome variables by treated region (Scotland) and control region (Rest of the UK) for the pre-and post-treatment period. SD stands for standard deviation and SE for standard error. *** indicates significance at the 1%-level.

both the pre and post-reform periods. We obtain 2×2 DID estimates by subtracting the difference of means of Scotland and the rest of the UK. Residential care increased slightly but insignificantly relative to the rest of the UK. For extra-residential care, both regions show similar pre-reform levels, however, Scottish extra-residential care grew significantly more than in the rest of the UK which leads to a significantly positive 2×2 DID estimate.

Scotland had lower pre-reform employment which remained below that of the rest of the UK post-reform but showed a greater increase. We calculate a significant 1.71 percentage point increase in employment. Both regions show similar pre-reform average working hours and increased at the same rate. The relative increase in working hours in Scotland is small and insignificant. With these preliminary benchmark estimates, we proceed to introduce our empirical strategy.

4 Method and Empirical Strategy

When estimating the causal effect of formal care on informal care and labour market participation, we are faced with endogeneity as care decisions are influenced by many unobserved characteristics (Bolin et al., 2008a). We use the Scottish care reform as it provides an exogenous increase in formal care supply. To estimate the treatment effect of a policy that provides the elderly with free personal care, one is generally confronted with the issue that individuals are observed either treated or untreated, but not the counterfactual. Comparing outcomes of the Scottish population before and after the reform in our case, would not distinguish between the treatment effect and the general trend aggregate outcomes follow over time. Comparing treated Scotland to the untreated rest of the UK would not account for differences in the population characteristics between these groups. In such a setting, a common approach to obtain causal estimates is the difference-in-differences method.

We estimate two different kinds of models. The first model is a simple DID model with a single coefficient of interest. The same model is used by Karlsberg Schaffer (2015) and Hollingsworth et al. (2022) such that we can easily compare our results to those of the existing literature. It is specified as follows:

$$y_{ist} = \alpha + \beta I_{st} + \lambda_t + \mu_s + X'_{ist}\Gamma + \epsilon_{ist} \tag{1}$$

The treatment indicator I_{st} takes the value one if an individual is living in Scotland and was interviewed in or after the policy introduction in the year 2002. λ_t is a vector of year fixed effects, μ_s indicates that the individual is living in Scotland and ϵ_{ist} is an error term. Assuming that Scotland and the rest of the UK would follow the same trend in the outcome variable y_{ist} if treatment had not occurred, we can estimate the average treatment effect of the treated (ATT), β. It is plausible that differences in the composition of the populations between Scotland and the rest of the UK lead to different dynamics in the demand for care as well as the supply of informal care and labour. As a consequence, we include a vector of individual level controls X_{ist} including age and sex which relaxes the identifying assumption to parallel trends conditioned on the vector of covariates X_{ist} and increases the precision of the estimates (Angrist & Pischke, 2009). Furthermore, we specify a model in which X_{ist} includes a state-specific time trend. The state-specific trend allows Scotland to follow a different linear trend over time and can potentially reveal different trends of treated and control regions if estimates would change significantly to the baseline model (Angrist & Pischke, 2009).

Other threats to the estimation of the ATT are anticipation of the reform and that the sample composition in treatment and control group is not stable over time (Hollingsworth et al., 2022). A non-stable sample composition could be caused by migration to Scotland in response to the reform (see e.g. Moffitt, 1992).

The second model that we estimate is a DID event study model which relies on the same identifying assumptions. We take advantage of the fact that our sample includes several pre and post-treatment years to estimate the average annual change in the outcome variables in Scotland relative to the average change in the rest of the UK both before and after the treatment. Let $D_{s\tau}$ denote the interaction between the binary indicator for the treated Scotland and a dummy variable for year τ. We specify the following equation:

$$y_{ist} = \alpha + \sum_{\tau=1991}^{2000} \gamma_\tau D_{s\tau} + \sum_{\tau=2002}^{2008} \beta_\tau D_{s\tau} + \lambda_t + \mu_s + X'_{ist}\Gamma + \epsilon_{ist} \tag{2}$$

We receive estimates of the treatment leads γ_τ and treatment lags β_τ which are relative to the year before the treatment introduction (2001) as it is left out as a baseline year. This strategy has two advantages over the simple DID model. Firstly, instead of a single DID estimate, we receive one coefficient for every year post-treatment such that we can analyse dynamic treatment effects. Secondly, estimates of treatment leads allow us to assess whether Scotland and the rest of the UK were comparable in their pre-treatment trends of outcome variables (Cunningham, 2021). However, it is important to note that this is no direct evidence for parallel counterfactual trends after treatment which is the identifying assumption (Kahn-Lang & Lang, 2020). Nevertheless, we test if estimates of lead coefficients are significantly different from zero and argue that if trends are not parallel before the treatment occurs, they would unlikely be parallel in the absence of treatment.

We estimate both models described above by ordinary least squares (OLS) for cardinal as well as binary outcomes. Linear probability models have the advantage that they are easy to interpret and were used in the previous literature studying the Scottish care reform such that we can compare estimated effect sizes easily to those of previous studies. We employ an interval regression estimated by maximum likelihood for interval-coded variables.

Estimating these models by OLS has the further advantage that we can treat the panel as a repeated cross-section. This is preferable in our case because the majority of individuals leave the panel study during the time horizon which could lead to non-random attrition-induced bias. Since the BHPS aims at maintaining a representative sample, OLS estimates on the unbalanced panel are preferred, furthermore, they are more precise (Lechner et al., 2016).

Treatment is introduced only in one region at one point in time, therefore concerns raised in the recent literature using the two-way fixed effects estimator when treatment is staggered do not apply (see e.g. Goodman-Bacon, 2021). A larger issue in our case with only two regions is inference. Bertrand et al. (2004) show that conventional standard errors are often downward biased due to large serial correlation of the outcome variable and only small variation of the treatment variable across regions and time. They suggest the use of block bootstrap standard errors as they are consistent for a sufficiently large number of groups. Furthermore, the use of cluster robust standard errors at the group level to which the treatment is assigned can provide consistent standard errors. However, neither option is viable in our application as the number of clusters is too small. There is no clear guidance in the literature on how to obtain consistent standard errors in this case except for using standard errors obtained from the 2×2 DID. Thus, we also consider our preliminary results from table 2. To account for individual-level auto-correlation from treating our unbalanced panel as repeated cross-sections, we simply cluster standard errors at the individual level.

5 Results

5.1 Informal care giving

Firstly, we replicate the main findings of Karlsberg Schaffer (2015) who use a DID approach to estimate the effect of the Scottish reform on informal care hours separately for male and female caregivers. Since the care-hours data is interval coded, they employ a participation distribution model which estimates probit models for a series of dummy variables indicating if an individual provided greater or equal to 5, 10, 20, 45, 50 or 100 hours of informal care. Our replication results are shown in table 7 and table 8 of the appendix B. Point estimates are similar to theirs but estimated with slightly lower precision. Furthermore, we find the same pattern as them, caregiving hours increase at the lower margin for females while hours decrease at the higher margin for males [1].

To further investigate caregiving at the intensive margin, we estimate an interval regression model using the care sample which includes all available time periods. Results shown in columns 4–6 of table 3 are similar to those of Hollingsworth et al. (2022), indicating an overall decrease in caregiving hours, significant at the 5%-level in all specifications.

[1] Results are almost identical when the large care sample **B** is used (not reported).

Table 3: Difference-in-differences estimates for informal care at the extensive and intensive margin.

	Informal Care			Weekly care hours		
	(1)	(2)	(3)	(4)	(5)	(6)
DD estimate						
	0.032***	0.032***	0.007	-4.440**	-4.927***	-5.136**
	(0.009)	(0.009)	(0.010)	(1.809)	(1.781)	(2.400)
Controls	-	✓	✓	-	✓	✓
State-specific trends	-	-	✓	-	-	✓
Constant	0.179***	0.094***	-0.961**	18.786***	-4.503**	-4.486**
	(0.005)	(0.008)	(0.411)	(1.153)	(2.215)	(2.216)
$\ln \sigma$						
				3.521***	3.510***	3.510***
				(0.018)	(0.018)	(0.018)
Trends test (F-stat)	5.440**	5.270**	0.420			
Observations	127499	127494	127494	22271	22269	22269

Notes: Standard errors are in parenthesis (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). Cluster robust standard errors allowing for individual-level correlation are used. Individual level controls for age and sex. Columns 1–6 use sample **B**.

Table 3 also includes extensive margin results for a DID linear probability model shown in columns 1–3. Results indicate that the reform increased the probability of an individual providing care by 3.2 percentage points which are in line with the findings of Karlsberg Schaffer (2015). However, these effects are undermined by the inclusion of state-specific trends as effects approach zero and become insignificant as presented in column 6.

We further investigate extensive margin results by differentiating between residential and extra-residential care with a DID linear probability model shown in table 4. Results in columns 1–3 indicate a 1 percentage point increase in residential care propensity from the reform, this is however insignificant. Those estimates seem plausible, due to personal ties cohabiting caregivers are unlikely to completely stop caregiving when formal care is made freely available. This contrasts the statistically significant decrease by 0.4 percentage points that Hollingsworth et al. (2022) estimate. Columns 4 and 5 reveal an increase in extra-residential care propensity by around 1.5 percentage points, contrasting Hollingsworth et al. (2022) who find no effects. We observe a similar change in estimates when including state-specific trends as in the aggregated care case in table 3. This is expected considering that the care outcome variable is an aggregation of residential and extra-residential care such that the effect first observed in table 3 is likely driven by the

TABLE 4: DIFFERENCE-IN-DIFFERENCES ESTIMATES FOR INFORMAL CARE AT THE EXTENSIVE MARGIN.

	Residential care			Extra-residential care		
	(1)	(2)	(3)	(4)	(5)	(6)
DD estimate						
	0.009	0.008	0.010	0.015***	0.014***	-0.011*
	(0.006)	(0.006)	(0.007)	(0.005)	(0.005)	(0.006)
Controls	-	✓	✓	-	✓	✓
State-specific trends	-	-	✓	-	-	✓
Constant	0.053***	-0.015***	0.056	0.110***	0.032***	-1.007***
	(0.003)	(0.005)	(0.255)	(0.003)	(0.004)	(0.244)
Trends test (F-stat)	0.836	0.000	0.180	14.300***	14.070***	0.800
Observations	127741	127736	127736	226142	226136	226136

Notes: Standard errors are in parenthesis (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). Cluster robust standard errors allowing for individual-level correlation are used. Individual level controls for age and sex. Columns 1–3 use sample **B**, columns 4–6 use sample **A**.

effect observed for extra-residential care. Interestingly, the coefficient for extra-residential care becomes negative when state-specific trends are included, suggesting a 1.1 percentage point decrease in response to the reform. The sensitivity of estimates to the inclusion of state-specific trends challenges the key identification assumption which we discuss extensively in section 5.3

5.2 Labour market participation

A policy reform introducing free formal care might have a desirable secondary effect of increasing employment. Caregivers may respond to this increase in formal care by decreasing their provision of informal care. Due to utility gains from additional income, individuals may shift their time use from informal care to paid employment. Therefore, we are interested in estimating the causal effect of the Scottish reform on employment and weekly working hours.

Table 5 presents DID estimates for the probability of employment in columns 1–3, while columns 4–6 show weekly working hours. In our baseline specification, we estimate that the reform increased employment by 1.8 percentage points. The effect is significant at the 5%-level and has a similar magnitude to the 2×2 DID estimates of table 2. When controlling for age and sex, the effect size and the precision of the estimate increase slightly. The point estimate is robust to the inclusion of state-specific trends and the effect

Table 5: Difference-in-differences estimates for labour market participation.

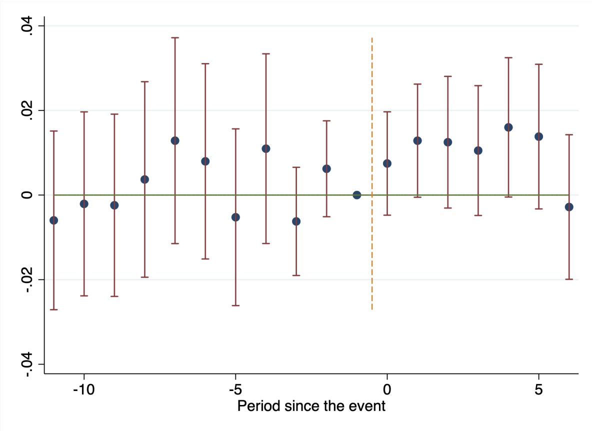
	Employed			Weekly working hours		
	(1)	(2)	(3)	(4)	(5)	(6)
DD estimate						
	0.018**	0.023***	0.017*	0.257	0.562	0.948**
	(0.008)	(0.008)	(0.009)	(0.397)	(0.387)	(0.392)
Controls	-	✓	✓	-	✓	✓
State-specific trends	-	-	✓	-	-	✓
Constant	0.782***	0.943***	0.655	21.594***	37.539***	54.227***
	(0.005)	(0.010)	(0.399)	(0.231)	(0.476)	(19.302)
Trends test (F-stat)	0.030	0.010	1.010	3.680*	2.460	4.550**
Observations	129103	127907	127907	129364	128168	128168

Notes: Standard errors are in parenthesis (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). Cluster robust standard errors allowing for individual-level correlation are used. Individual level controls for age and sex. Columns 1–6 use sample **D**.

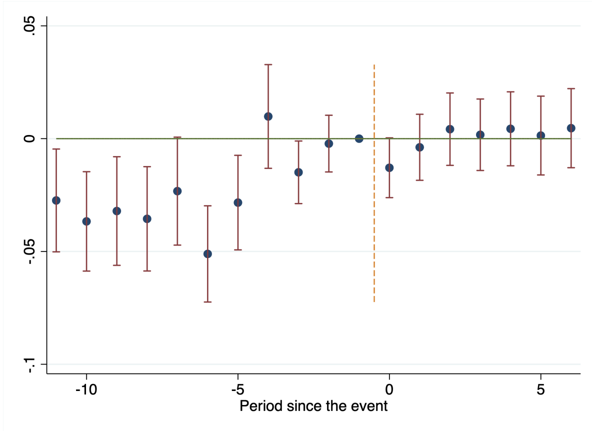
remains significant at the 10%-level. Hollingsworth et al. (2022) calculate DID estimates through mean differencing and regression with individual-level data. Interestingly, our point estimates are very similar to their 2×2 DID point estimates of 1.7 percentage points. However, our DID estimates are substantially larger than their regression DID estimates which show an insignificant increase of 0.7 percentage points.

We find that the reform increased weekly working hours on average by 0.26 hours in our baseline specification. Including controls the effect size roughly doubles to 0.57 hours. However, neither estimates are statistically significant. In column 6, additionally, we include state-specific trends which result in an estimate indicating an almost one-hour increase in working hours, significant at the 5%-level. We discuss the instability of the estimates in columns 4–6 in the following sub-section.

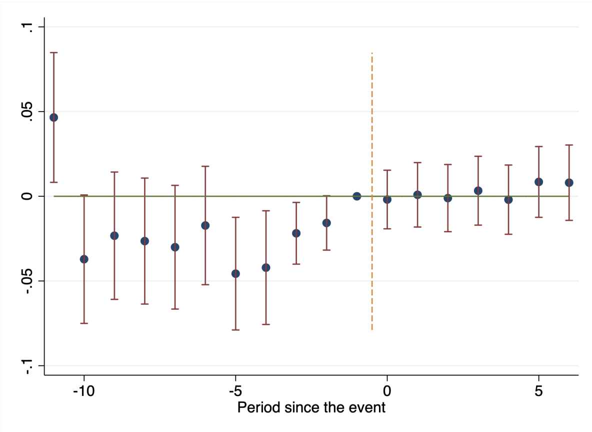
Hollingsworth et al. (2022) estimates a 0.41 hour increase in working hours, which falls between our coefficients in columns 4 and 5. Both their estimates and ours in table 5 use a sample including unemployed individuals. To examine if the increase in working hours is driven by the unemployed entering employment, we reduce the sample to employed individuals. We report these results in table 9 of the appendix B. Interestingly, the estimates are smaller and are roughly zero when including state-specific trends in the employed-only sample. This suggests that estimates for the entire workforce are likely driven by unemployed individuals becoming employed.



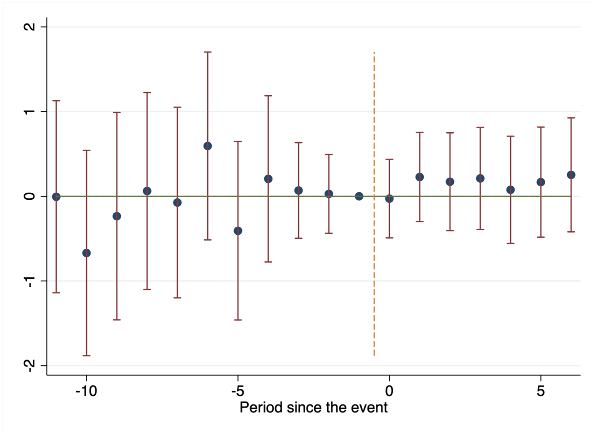
(A) Residential care



(B) Extra-residential care



(C) Employment probability



(D) Weekly working hours

Figure 1: Difference-in-differences event study estimates for informal care and labour market participation outcomes. Coefficients are plotted with 95% confidence intervals. Treatment is introduced in Scotland in 2002, the previous year is left out as a reference group.

5.3 Identification assumptions analysis

5.3.1 Parallel Trends

In the previous results, we observed that the coefficients for some of our specifications were affected drastically by the inclusion of state-specific trends. This inclusion caused the coefficient for extra-residential care to switch signs, see table 4. The same phenomenon, but weaker, is observed for the provision of any form of care in columns 1–3 of table 3. This is unsurprising given that the informal care outcome is an aggregation of residential and extra-residential care. Finally, the inclusion also causes the weekly working hours estimate to become significant in column 6 of table 5.

Drastic changes in estimates when including state-specific trends imply that the outcome variables actually follow different paths over time in Scotland compared to the rest of the UK. This challenges the identification assumption that trends would have been parallel in the absence of treatment. For instance, we might misinterpret a steeper trend in extra-residential care in Scotland as a positive treatment effect. The change in the coefficient when controlling for different linear trends is evidence for this phenomenon. This is further supported by testing for different linear trends as we can reject parallel trends only when state-specific trends are not included. The F-statistics for this test are reported in table 3. Although the common linear trends test might suggest common pre-trends, we cannot necessarily rely on estimates including linear state trends as the true functional form is unknown. In the case that actual trends are not linear, the model including linear trends would be miss-specified and thus biased (Kahn-Lang & Lang, 2020). Nevertheless, we interpret our findings as strong evidence that positive reform effects on extra-residential care may be biased. This has important implications as to whether formal and informal care are indeed substitutes as opposed to compliments as suggested by our baseline estimates and Karlsberg Schaffer (2015).

In contrast to extra-residential care, for weekly working hours, we can still reject the null hypothesis of parallel pre-treatment trends when state trends are included. This provides evidence that the estimates may be substantially biased as Scotland and England might follow different non-linear trends.

To further investigate the pre-treatment trends, we conduct placebo tests by testing if pre-treatment estimates are significantly different from zero. Therefore, we use event study plots shown in fig. 1 that show coefficients for every period and their respective confidence intervals. For residential care propensity, the lead coefficients of the event study plot are all insignificant but with large confidence bands. A large number of leads for extra-residential care are significantly different from zero indicating that pre-trends of treatment and control group are not parallel. This is consistent with the evidence from the inclusion of state-trends and the common linear trends test of table 4. In contrast to the previous findings when controlling for linear trends and of the linear trend test, the event study plot for employment probability reveals that some lead coefficients are significantly different from zero. For weekly working hours, we observe some lead coefficients that are relatively far from zero but insignificant as confidence bands are relatively large. That we cannot reject parallel pre-trends contrasts the evidence from the inclusion of state-trends and the common linear trends test in table 5.

5.3.2 Anticipation and reform-induced migration

Another possible threat to the identification of treatment effects is that individuals in Scotland were able to anticipate the reform and adjusted their behaviour before the reform took place. Indeed while the reform was enacted in July 2002, the intention to enact the policy was announced earlier, on 24 September 2001, receiving much media attention thereafter (BBC, 2001; Hollingsworth et al., 2022; Scottish Executive, 2005). This implies that any Scottish interview responses conducted between these two dates may be subject to anticipatory behaviours. Using interview date information, we see that around 80% of Scottish respondents in 2001 were interviewed after the policy announcement such that they could have possibly altered their behaviour.

To check if anticipation played an important role, we analyse if our results are robust to dropping the year 2001 from our sample. Figure 2 shows new event study plots excluding the year 2001. Pre-trends diagnostics for extra-residential care and weekly working hours are robust to the exclusion of the year 2001. Residential care now has a significant pre-treatment coefficient close to the reform year, undermining that pre-trends are parallel which challenges the counterfactual parallel trends assumption.

Comparing fig. 1c and fig. 2c where the year 2001 is omitted, we observe that employment lead coefficients become insignificant (except 1991) while lag coefficients increase and become significant towards the end of the time horizon. The results of this robustness analysis suggest that anticipatory behaviour exists. Results for pre-trends and treatment effects change because event study estimates are reported relative to the reference period. There are likely anticipation effects in 2001 masking treatment effects in the lags as they are reported relative to 2001. Furthermore, leads are significant relative to 2001. Loss of leads' significance when omitting 2001 and changing the reference year, suggests that deviations from a common pre-treatment trend observed in fig. 1c can be attributed to anticipatory behaviour. We find it plausible that individuals start job searches in advance with the knowledge that free formal care will become available in the future.

Another threat to identification is migration to Scotland induced by the reform. We investigate this threat, finding that less than 0.7% of the full sample migrated between the two regions during the time horizon. Furthermore, Ohinata and Picchio (2020) find with BHPS data that the reform had no impact on migration to Scotland from the rest of the UK.

5.3.3 Discussion

We investigated identification assumptions focusing on the parallel trends assumption by evaluating if Scotland and the rest of the UK followed similar trends before the treatment was introduced. Therefore, we included state-specific trends in our regressions as well as tests for common linear trends and pre-treatment placebo tests.

Results for employment are robust to the inclusion of state trends and we are not able to reject common linear trends. We observe some significant leads in the event study plot in fig. 1c which however are likely due to anticipatory behaviour as coefficients are reported relative to the baseline year before treatment. We interpret these findings as evidence that the reform likely caused an increase in employment.

The inclusion of state trends and the linear trends test for weekly working hours reveal non-linear trends. We are not able to reject parallel pre-trends with placebo tests. However, we do not interpret our estimates as causal as they are not robust to the inclusion of controls and state-specific trends.

Tests for parallel pre-trends in residential care do not provide evidence to challenge the parallel trends assumption. The estimates for residential care are insignificant such that we cannot rule out negative treatment effects. Conversely, extra-residential care linear trends tests suggest that estimates might pick up a steeper trend in Scotland as a treatment effect. Additionally, pre-treatment placebo tests reveal that Scotland and the rest of the UK do not follow a common trend before treatment was introduced. As a consequence, we cannot rely on these estimates since they are likely biased. Since Karlsberg Schaffer (2015) uses an aggregation of residential and extra-residential care with the same data her estimates of a positive treatment effect are likely to be strongly biased too. Moreover, the inclusion of state-specific trends shows that we cannot rule out that the actual treatment effect is negative.

We further investigate if it is plausible that formal and informal care are complements as our estimates in table 4 and those of Karlsberg Schaffer (2015) suggest. In such a case, the increase in employment caused by the reform together with caregivers' time constraints would imply a decrease in their leisure time. Our data allow us to test this hypothesis. Table 10 in appendix B reveals an increase in leisure time that is however insignificant. The simultaneous increase in employment, informal care, and leisure that we observe in our data are inconsistent with the caregivers' time constraints which makes an increase in informal care seem implausible.

5.4 Synthetic Difference-in-differences

The parallel trends assumption appears to be violated for the outcome variables extra-residential care and weekly working hours. It is unclear whether controlling for linear state-specific trends can solve the problem as trends could be non-linear. With the aim of achieving unbiased estimates, we use the new synthetic difference-in-differences (SDID) method developed by Arkhangelsky et al. (2021). The method can provide consistent estimates where the parallel trend assumption appears to be violated by re-weighting observations such that pre-trends are parallel.

The SDID method requires a strictly balanced panel. However, the BHPS is a large micro study running over a long time horizon, therefore there are many sample leavers and joiners. Reducing the sample to only individuals that are observed in all sample periods, shrinks the sample to about one-fifth of the original sample size, this poses the threat of non-random attrition bias.

We obtain SDID estimates reported in table 11 and trend plots in fig. 3 of appendix B. They indicate an increase in residential and extra-residential care. In contrast to our previous findings, they indicate a decrease in labour market participation both on the extensive and intensive margins. However, the results are insignificant for all outcomes. These results contradict our previous findings for labour market participation and the evidence we found for non-parallel trends, causing an upward bias of extra-residential care estimates.

Thus, we investigate if they are biased due to non-random panel attrition by performing DID estimates with the same sample as used for the SDID estimation. If DID estimates of the balanced sample are different compared to those of the main results in columns 1 and 4 of both table 4 and table 5, this suggests that non-random panel attrition biases them (Lechner et al., 2016). When comparing these, we also have positive, but substantially larger estimates for both residential and extra-residential care while for labour outcomes the signs are different. This provides strong evidence that the balanced sample indeed suffers from attrition bias. As a consequence, we conclude that the SDID estimates also suffer from the same attrition bias.

Because parallel trends were violated and a large number of control units were available, the SDID method was a promising approach. However, due to non-random attrition, we do not attain reliable estimates from it.

6 Conclusion

In this paper, we revisit the Scottish Community Care and Health Act of 2002 which provisioned free elderly personal care regardless of income or assets. Since the rest of the UK retained the existing system of means-tested eligibility, we use a DID approach to estimate the reform's effect on informal caregiving and labour market participation.

There have been two published studies investigating the reform finding contradicting results. Karlsberg Schaffer (2015) find that the reform increased informal caregiving suggesting that state-funded formal care and informal care are complements. In contrast, Hollingsworth et al. (2022) estimate a decrease in informal care in response to the reform. They extend the analysis for labour market participation and find an increase in both employment and weekly working hours.

We focus on understanding why the results of Karlsberg Schaffer (2015) differ from those of Hollingsworth et al. (2022) from which we synthesise which estimates are reliable. We replicate the findings of Karlsberg Schaffer (2015), however, find strong evidence that Scotland follows a steeper trend in extra-residential care that biases the estimates upwards. This finding suggests that formal and extra-residential informal care might be substitutes instead of compliments. We extend the analysis to include labour market participation and similar to Hollingsworth et al. (2022), find that the reform caused employment to increase by 1.7 percentage points. Furthermore, results for weekly working hours also indicate a positive treatment effect, however, this is mainly driven by the increase in employment. Additionally, we find evidence that Scotland and the rest of the UK follow different non-linear trends in weekly working hours which likely biases estimates.

The cost-effectiveness of providing state-funded formal elderly care rests on its effect on informal care provision and labour market participation. While the estimated increase in employment is a desirable effect of the reform, elderly well-being is also an essential factor in evaluating the policy's success. However, our results of the reform's effect on informal care are less conclusive. Since it is unclear if formal care crowds out or complements informal care, the net change in care received and thus the effect on elderly well-being is unclear.

Due to violations of the identifying assumptions, it is difficult to measure the reform's causal effect on informal caregiving. New synthetic difference-in-differences methods using BHPS data cannot help resolve this issue due to data limitations. Thus, we suggest that further research could focus on directly measuring benefits by estimating the reform's effect on elderly health.

References

Alemayehu, B., & Warner, K. E. (2004). The lifetime distribution of health care costs. *Health services research*, *39*(3), 627–642.

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review*, *111*(12), 4088–4118.

Audit Scotland. (2008). A review of free personal and nursing care [Available at: https://www.audit-scotland.gov.uk/uploads/docs/report/2008/nr_080201_free_personal_care.pdf].

BBC. (2001). Free care deal for elderly people. http://news.bbc.co.uk/1/hi/scotland/1559427.stm

Bell, D. N., Bowes, A., & Heitmueller, A. (2007). Did the introduction of free personal care in scotland result in a reduction of informal care? *WDA-HSG Discussion Paper*, (2007-3).

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly journal of economics*, *119*(1), 249–275.

Bolin, K., Lindgren, B., & Lundborg, P. (2008a). Informal and formal care among single-living elderly in europe. *Health economics*, *17*(3), 393–409.

Bolin, K., Lindgren, B., & Lundborg, P. (2008b). Your next of kin or your own career?: Caring and working among the 50+ of europe. *Journal of health economics*, *27*(3), 718–738.

Bonsang, E. (2009). Does informal care from children to their elderly parents substitute for formal care in europe? *Journal of health economics*, *28*(1), 143–154.

Carrino, L., Nafilyan, V., & Avendano, M. (2023). Should i care or should i work? the impact of work on informal care. *Journal of Policy Analysis and Management*, *42*(2), 424–455.

Cunningham, S. (2021). *Causal inference: The mixtape*. Yale university press.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277.

He, W., Goodkind, D., Kowal, P. R., et al. (2016). *An aging world: 2015*. United States Census Bureau Washington, DC.

Hollingsworth, B., Ohinata, A., Picchio, M., & Walker, I. (2022). The impacts of free universal elderly care on the supply of informal care and labour supply. *Oxford Bulletin of Economics and Statistics*, *84*(4), 933–960.

Inman, P. (2002). Free and easy for the scots. *The Guardian*. https://www.theguardian.com/society/2002/mar/23/longtermcare.housinginretirement3

Kahn-Lang, A., & Lang, K. (2020). The promise and pitfalls of differences-in-differences: Reflections on 16 and pregnant and other applications. *Journal of Business & Economic Statistics*, *38*(3), 613–620.

Karlsberg Schaffer, S. (2015). The effect of free personal care for the elderly on informal caregiving. *Health economics*, *24*, 104–117.

Lechner, M., Rodriguez-Planas, N., & Fernández Kranz, D. (2016). Difference-in-difference estimation by fe and ols when there is panel non-response. *Journal of Applied Statistics*, *43*(11), 2044–2052.

Løken, K. V., Lundberg, S., & Riise, J. (2017). Lifting the burden formal care of the elderly and labor supply of adult children. *Journal of Human Resources*, *52*(1), 247–271.

Moffitt, R. (1992). Incentive effects of the us welfare system: A review. *Journal of economic literature*, *30*(1), 1–61.

Ohinata, A., & Picchio, M. (2020). Financial support for long-term elderly care and household saving behaviour. *Oxford Economic Papers*, *72*(1), 247–268.

Schmitz, H., & Westphal, M. (2017). Informal care and long-term labor market outcomes. *Journal of health economics*, *56*, 1–18.

Scottish Executive. (2005). *Guidance on free personal and nursing care in scotland 2003* [Available at: https://www.gov.scot/publications/guidance-free-personal-nursing-care-scotland-2003/pages/2/.]. Scottish Government.

University of Essex. (2018). British Household Panel Survey: Waves 1-18, 1991-2009 [Institute for Social and Economic Research 8th Edition. UK Data Service. SN: 5151]. https://doi.org/10.5255/UKDA-SN-5151-1

Urwin, S., Lau, Y.-S., & Mason, T. (2019). Investigating the relationship between formal and informal care: An application using panel data for people living together. *Health Economics*, *28*(8), 984–997.

A Samples

Table 6: Overview of samples used and their restrictions.

A Full sample	Drop all observations with missing values in outcome or control variables
B Care outcomes	Include only observations with no children present in household
C Care outcomes (Karlsberg Schaffer (2015))	Include only individuals older than 44 and with no children present in household
D Labour outcomes	Use sample of labour force: drop individuals younger than 26 and older than 74, drop retirees, full-time students and long-term sick/disabled. Drop implausible values of weekly working hours (>60 hours)
E Labour outcomes (hours > 0)	Use sample D and keep only individuals with positive working hours

B Supplementary Results

Table 7: Marginal effects of the care reform on number of care hours of women.

	5+ hours	10+ hours	20+ hours	35+ hours	50+ hours	100+ hours
DD estimate	0.0228**	0.0110	0.0069	0.0010	0.0018	-0.0005
	(0.0114)	(0.0088)	(0.0067)	(0.0046)	(0.0040)	(0.0036)
Controls	✓	✓	✓	✓	✓	✓
Observations	42,867	42,867	43,195	42,753	42,753	42,753

Notes: Standard errors are in parenthesis (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). Cluster robust standard errors allowing for individual level correlation are used. Individual level controls for marital status, employment status, education level, household size and age. Columns 1–6 use sample **C**

Table 8: Marginal effects of the care reform on number of care hours of men.

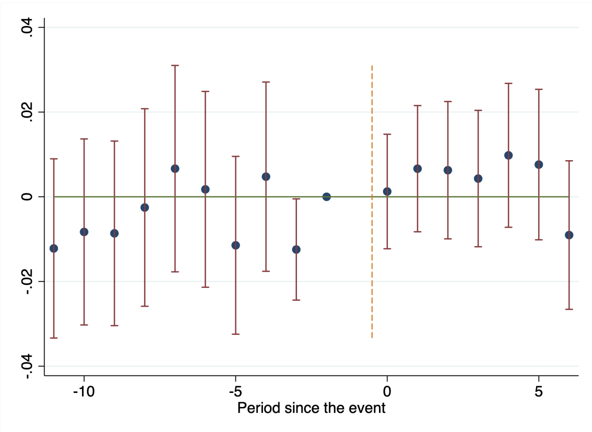
	5+ hours	10+ hours	20+ hours	35+ hours	50+ hours	100+ hours
DD estimate	0.0024	-0.0036	-0.0082	-0.0071	-0.0062	-0.0062*
	(0.0117)	(0.0095)	(0.0070)	(0.0047)	(0.0040)	(0.0035)
Controls	✓	✓	✓	✓	✓	✓
Observations	33,170	33,165	33,429	33,186	33,186	33,186

Notes: Standard errors are in parenthesis (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). Cluster robust standard errors allowing for individual level correlation are used. Individual level controls for marital status, employment status, education level, household size and age. Columns 1–6 use sample **C**.

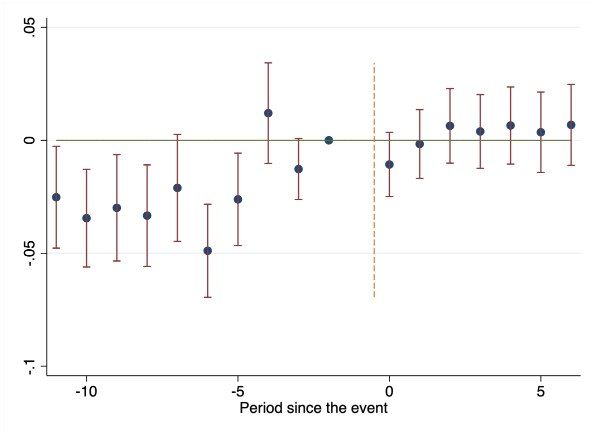
Table 9: Difference-in-differences estimates for labour market participation at the intensive margin.

	Weekly working hours			Weekly working hours (hours> 0)		
	(1)	(2)	(3)	(4)	(5)	(6)
DD estimate						
	0.257	0.562	0.948**	0.168	0.174	-0.026
	(0.397)	(0.387)	(0.392)	(0.271)	(0.240)	(0.249)
Controls	-	✓	✓	-	✓	✓
State-specific trends	-	-	✓	-	-	✓
Constant	21.594***	37.539***	54.227***	33.068***	43.051***	34.486***
	(0.231)	(0.476)	(19.302)	(0.169)	(0.288)	(12.453)
Trends test (F-stat)	3.680*	2.460	4.550**	0.170	0.280	0.000
Observations	129364	128168	128168	96090	95007	95007

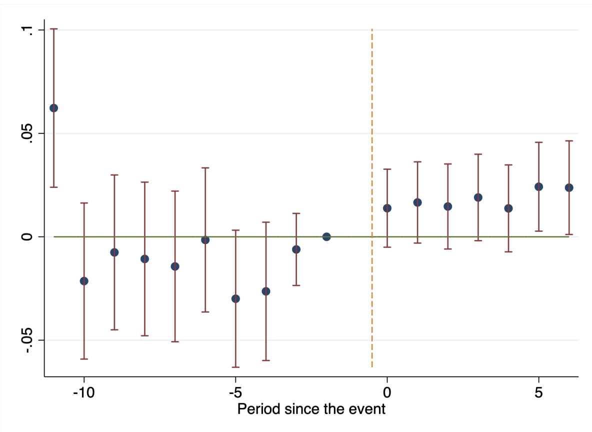
Notes: Standard errors are in parenthesis (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). Cluster robust standard errors allowing for individual level correlation are used. Individual level controls for age and sex. Columns 1–3 use sample **D**, columns 4–6 use **E**.



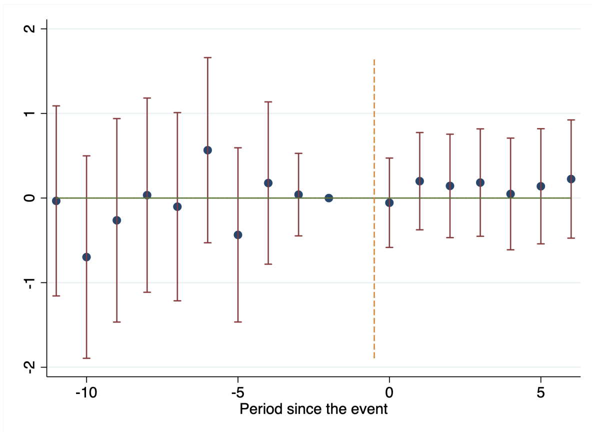
(a) Residential care



(b) Extra-residential care



(c) Employment probability



(d) Weekly working hours

Figure 2: Sensitivity analysis of anticipation effects: Difference-in-differences event study estimates for informal care and labour market participation outcomes. Coefficients are plotted with 95% confidence intervals. We drop the survey year 2001 from the sample to examine anticipation effects. Treatment is introduced in Scotland in 2002, the year 2000 is left out as reference group.

Table 10: Difference-in-differences estimates for leisure hours.

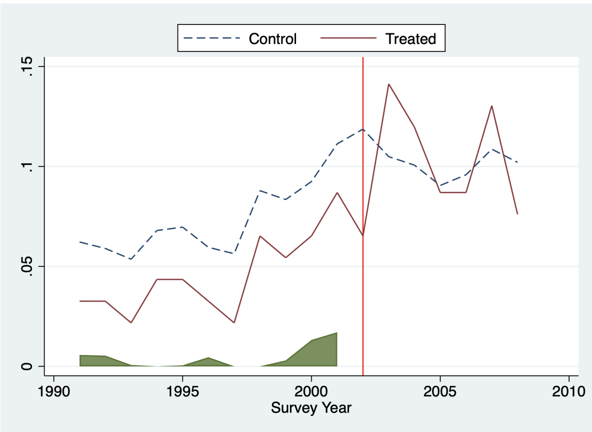
	(1)	(2)	(3)
DD estimate	0.587	0.809	1.234
	(0.869)	(0.853)	(0.844)
Controls	-	✓	✓
State-specific trends	-	-	✓
Constant	34.798***	72.221***	72.185***
	(0.409)	(0.661)	(0.662)
$\ln \sigma$	3.756***	3.690***	3.690***
	(0.005)	(0.005)	(0.005)
Observations	169813	169807	169807

Notes: Standard errors are in parenthesis (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). Cluster robust standard errors allowing for individual level correlation are used. Individual level controls for age and sex. Columns 1–3 use sample **A**. Leisure outcomes only available from 1997–2008.

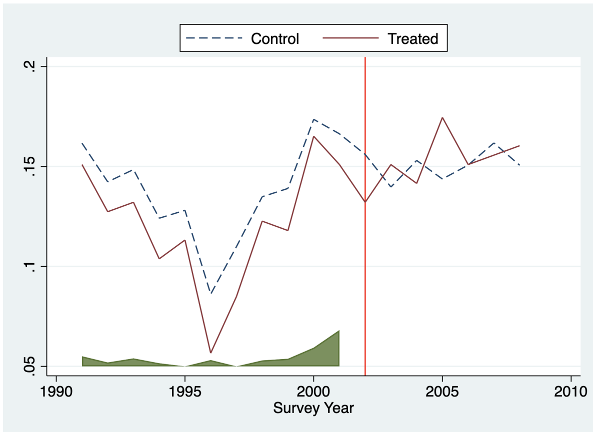
Table 11: Synthetic Difference-in-differences estimates.

	(1) Residential care	(2) Extra-residential care	(3) Employed	(4) Weekly working hours
SDID estimates	0.0245	0.0165	-0.0149	-1.719
	(0.0179)	(0.0194)	(0.0141)	(1.0820)
DID estimates	0.0312*	0.0309**	-0.00764	-2.739*
	(0.0185)	(0.0157)	(0.0164)	(1.6051)
Observations	25632	56124	16974	16974

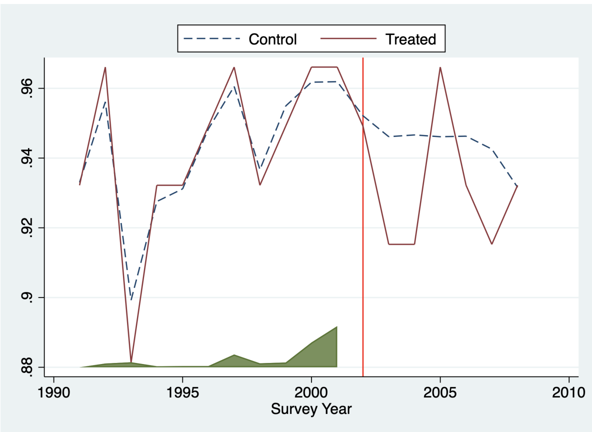
Notes: Standard errors are in parenthesis (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). Block bootstrap standard errors are used. Columns 1 uses sample **B**, columns 2 uses sample **A** and columns 3–4 use sample **D**. The samples consist of all individuals that responded in all 18 BHPS waves.



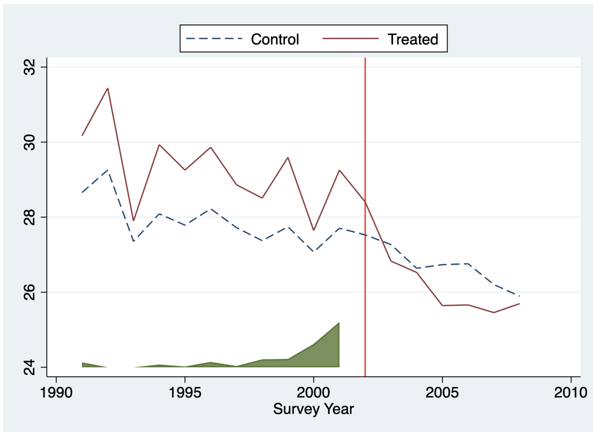
(A) Residential care



(B) Extra-residential care



(C) Employment probability



(D) Weekly working hours

Figure 3: Trend plots of Scotland's mean outcomes and the synthetic control unit which is a weighted average of the rest of UK.