The Gig Economy That Wasn't: Universal Credit and the Composition of Employment in Britain

APEP Autonomous Research[*]   @olafdrw

February 27, 2026

Abstract

Britain's Universal Credit (UC), the largest welfare reform in a generation, replaced six legacy benefits with a single monthly payment featuring simplified self-employment reporting and a unified taper rate. Critics warned that UC's design would push claimants toward precarious self-employment rather than stable jobs. We test this hypothesis by exploiting the staggered rollout of UC Full Service across 378 local authorities from 2016 to 2018. Using Callaway-Sant'Anna (2021) doubly robust difference-in-differences with administrative data from the Annual Population Survey, we find a precisely estimated null effect on the self-employment share (-0.14 percentage points, SE $= 0.35$). This null is robust to excluding London, restricting to England, placebo tests, and sensitivity analysis under violations of parallel trends. UC did not reshape the composition of employment—the gig economy narrative, at least as applied to welfare reform, finds no support in the data.

JEL Codes: J21, J65, H53, I38

Keywords: Universal Credit, self-employment, welfare reform, difference-in-differences, gig economy, United Kingdom

[*]Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: N/A).

1. Introduction

When Britain began rolling out Universal Credit in 2013, replacing a patchwork of six means-tested benefits with a single monthly payment, critics raised an alarm that went beyond the usual welfare-reform anxieties about poverty and bureaucratic error. The specific concern was compositional: that UC's design—its monthly assessment periods, digital-first interface, and simplified self-employment reporting—would systematically channel benefit claimants away from employee jobs and toward precarious self-employment in the emerging gig economy (**?**). The fear was intuitive. Under UC, self-employment is easier to declare than under the legacy system, the single taper rate reduces the penalty for combining benefits with irregular earnings, and starting a one-person business requires no employer search—only a declaration to the Department for Work and Pensions. If welfare systems shape not just whether people work but *how* they work, UC could quietly accelerate the structural shift toward a more precarious labour force.

This paper tests that hypothesis directly. We exploit the staggered rollout of UC Full Service across approximately 643 Jobcentre Plus offices—mapped to 378 local authorities—between 2016 and 2018. Using quarterly employment data from the Annual Population Survey available through NOMIS, we estimate the causal effect of UC rollout on the self-employment share of total employment. Our identification strategy relies on the Callaway and Sant'Anna (2021) doubly robust difference-in-differences estimator, which handles the heterogeneous treatment effects that arise naturally from staggered adoption. The pre-treatment period extends to 2010, providing at least five years of pre-treatment data for each local authority. We restrict the post-treatment window to 2019 to avoid contamination from the COVID-19 pandemic.

Our main finding is a precisely estimated null. The Callaway-Sant'Anna overall ATT on the self-employment share is -0.14 percentage points (SE $= 0.35$), with a 95% confidence interval of $[-0.83, 0.56]$. Given a pre-treatment mean self-employment share of 15.1%, we can rule out effects larger than 0.83 percentage points—roughly 5.5% of baseline—in either direction. The pre-test for parallel trends yields a p-value of 0.99, indicating no detectable differential trends before UC adoption. The event study estimates show no pre-trend violation and no post-treatment divergence, confirming that the identifying assumption is well-supported.

The null result is robust across multiple dimensions. Excluding London boroughs—where gig economy activity is most concentrated—produces an estimate of -0.02 (SE $= 0.39$). Restricting to English local authorities yields -0.45 (SE $= 0.45$). A placebo test assigning a fake treatment date of 2014 using only pre-treatment data generates an insignificant coefficient of 0.14 (SE $= 0.31$). HonestDiD sensitivity analysis confirms that the confidence interval

includes zero even under substantial violations of the parallel trends assumption.

These results speak to three literatures. First, they contribute to the growing body of work evaluating Universal Credit. Existing UC evaluations have focused primarily on mental health (**?**), food insecurity, debt, and housing outcomes, documenting significant negative effects of the transition on claimant wellbeing. A smaller literature examines UC's aggregate employment effects, finding modest positive impacts consistent with improved work incentives. Our contribution is to examine employment *composition*—the quality margin rather than the quantity margin—using modern heterogeneity-robust causal methods. No existing study has tested whether UC shifts the mix of employment types using a credible identification strategy, despite the prominence of this concern in policy debates.

Second, our results address the broader theoretical question of whether welfare design can reshape the quality of work. Job search models with heterogeneous offers predict that benefit generosity and structure affect not only the decision to work but also the type of job accepted (**?**). More generous benefits allow longer search, leading to better matches; simpler benefit administration for certain work types could tilt the composition of accepted offers. Our evidence suggests that the composition margin is substantially less responsive to welfare design than the extensive (participation) margin, at least for the UC-style reforms that consolidate existing benefits without fundamentally changing generosity levels.

Third, they contribute to the empirical literature on the gig economy, which has documented its rapid growth (**?**) and debated its causes. Explanations range from technological change (platform intermediaries reducing transaction costs) to employer demand for flexibility to worker preferences for autonomy. Our results rule out one candidate explanation—welfare policy—providing useful boundary conditions for this debate. The gig economy grew in Britain despite, not because of, changes in the benefit system.

The paper proceeds as follows. Section 2 describes Universal Credit and the institutional details of the Full Service rollout. Section 3 presents the data and identification strategy. Section 4 reports the main results. Section 5 provides robustness checks. Section 6 discusses mechanisms and implications. Section 7 concludes.

2. Institutional Background

2.1 The Legacy Benefit System and Its Discontents

Before Universal Credit, Britain's working-age benefit system had evolved over decades into an exceptionally complex structure. Six separate means-tested benefits—Jobseeker's Allowance (JSA), Employment and Support Allowance (ESA), Income Support, Housing Benefit, Child Tax Credit, and Working Tax Credit—were administered by three different

government departments (DWP, HMRC, and local councils), each with its own application process, means test, and taper rate. A single household could interact with multiple agencies simultaneously, filing separate claims with different reporting periods and eligibility rules.

The consequences for work incentives were severe. Because different benefits were withdrawn simultaneously as earnings rose, effective marginal tax rates (EMTRs) could exceed 90% for some claimants. The Institute for Fiscal Studies documented that approximately 300,000 workers faced EMTRs above 80% under the legacy system. The "poverty trap"—where working more produced negligible gains in disposable income—was not merely theoretical but a lived reality for millions of low-income households.

For self-employment specifically, the legacy system posed distinctive challenges. Self-employment income was assessed differently across different benefits: JSA required weekly reporting, tax credits used annual income projections subject to end-of-year reconciliation, and Housing Benefit applied yet another set of rules for business expenses. A self-employed benefit claimant might need to report the same income in different formats to three separate agencies. This administrative burden acted as an implicit tax on self-employment for benefit recipients, over and above the financial disincentives created by taper rates.

The 2010 White Paper "Universal Credit: Welfare That Works" (**?**) articulated the case for consolidation. Its central claim was that simplification would "make work pay" by creating a single, transparent taper rate and eliminating the cliff edges that discouraged transitions from benefits to employment. The paper projected that UC would move 300,000 additional people into work through improved incentives alone.

2.2 Universal Credit: Design and Motivation

Universal Credit replaced the six legacy benefits with a single monthly payment. The reform was motivated by the perceived complexity and perverse incentives described above (**?**). The design features of UC relevant to self-employment are:

1. **Unified taper rate.** UC is withdrawn at 55 pence per pound of net earnings (reduced from 65p in April 2017), compared to the legacy system where multiple benefits tapered simultaneously, producing much higher effective marginal tax rates.

2. **Monthly assessment periods.** UC is calculated on a monthly basis, aligning with standard pay cycles. Self-employed claimants report actual earnings each month, replacing the complex annual tax credit reconciliation process.

3. **Simplified self-employment reporting.** Under legacy benefits, self-employment income was assessed differently across different benefits. UC consolidated this into a single monthly declaration.

4

4. **Digital-by-default.** UC is managed through an online journal, reducing administrative friction but potentially favouring digitally literate claimants.

5. **Minimum Income Floor (MIF).** After a 12-month start-up period, self-employed UC claimants are assumed to earn at least the equivalent of working full-time at the National Minimum Wage. This caps the UC entitlement for low-earning self-employed, potentially discouraging "hobby" or marginal self-employment.

The theoretical prediction for self-employment is therefore ambiguous. On one hand, UC's simplified reporting and lower taper rate reduce barriers to combining self-employment with benefit receipt. The unified assessment period eliminates the administrative nightmare of reporting self-employment income to multiple agencies in different formats. For a would-be entrepreneur receiving benefits, UC is unambiguously simpler than the system it replaced.

On the other hand, the MIF penalizes low-earning self-employment after the first year, and UC's stricter conditionality requirements may push claimants toward employee jobs that more easily satisfy work search obligations. The "Claimant Commitment"—a personalized agreement signed by each UC claimant—typically specifies job search activities oriented toward employee positions, with self-employment treated as an alternative path requiring explicit approval from a work coach. These countervailing institutional forces make the net effect on employment composition an empirical question.

2.3 Self-Employment in the UK: Context and Trends

Understanding the potential for UC to affect self-employment requires context on the UK's self-employment landscape. Between 2001 and 2019, the number of self-employed workers in the UK grew from approximately 3.3 million to 5.0 million, an increase of over 50%. By the mid-2010s, self-employment accounted for roughly 15% of total employment—one of the highest shares among OECD economies.

This growth was compositionally distinctive. Unlike earlier periods of self-employment expansion, which were concentrated in traditional trades (construction, professional services), the post-2008 increase was driven disproportionately by solo self-employment—individuals working alone without employees. Much of this growth occurred in lower-skilled occupations: delivery driving, personal care, cleaning, and other service activities that would later become associated with the "gig economy" label. The Resolution Foundation documented that real earnings among the self-employed fell by 20% between 2006 and 2018, suggesting that the composition of the self-employed population was shifting toward lower-productivity, lower-income activities.

The gig economy narrative posited that technology platforms (Uber, Deliveroo, TaskRabbit) were accelerating this shift by creating new forms of work that blurred the boundary between employment and self-employment. Workers engaged through platforms were typically classified as self-employed for tax and benefit purposes, even though their working conditions resembled employment in many respects. The Taylor Review of Modern Working Practices (2017) identified this classification ambiguity as a key challenge for the welfare system.

It was in this context—rapid self-employment growth, rising platform work, and ongoing welfare reform—that concerns about UC's compositional effects emerged. If UC made self-employment easier to combine with benefit receipt, it could amplify an already-existing structural trend. The question was whether welfare design was a meaningful driver of employment composition, or merely a bystander to deeper technological and institutional forces (**?**).

2.4 The Full Service Rollout

UC was introduced in stages. "Live Service," available from 2013 in a small number of areas, was restricted to simple cases (single, childless claimants). "Full Service," which opened UC to all new claims including families and complex cases, was rolled out across the UK between 2015 and December 2018.

The rollout was organized by Jobcentre Plus (JCP) office, with each JCP assigned a date to transition to Full Service. Approximately 643 JCPs were mapped to local authorities across England, Scotland, and Wales. The rollout proceeded in phases:

- **Phases 1–2 (2016):** Early adopters, approximately 36 local authorities.

- **Phases 3–4 (2017):** Main expansion, approximately 147 local authorities.

- **Phases 5–8 (2018):** Completion, approximately 199 local authorities.

The phasing was determined primarily by operational readiness—the physical infrastructure, staffing, and IT capacity of each Jobcentre Plus office—rather than local labour market conditions. DWP officials have stated publicly that rollout scheduling reflected logistical constraints, not targeting based on economic need or opportunity. Nevertheless, the non-random assignment of rollout dates means that endogeneity cannot be ruled out a priori. We address this concern empirically through pre-trend tests and balance checks across cohorts.

The distinction between Live Service and Full Service is crucial for our identification strategy. Under Live Service, only the simplest claims were processed as UC—typically single,

childless, newly unemployed individuals with no housing costs. These cases represented a small fraction of the total caseload. Full Service opened UC to all new claimants, including families, homeowners, individuals with disabilities, and the self-employed. It is the Full Service transition that exposed a meaningful share of the local population to UC's design features, including the self-employment-relevant provisions described above.

Upon Full Service transition, all new benefit claims in the area were processed as UC rather than legacy benefits. Existing legacy claimants continued on their current benefits and were scheduled for "managed migration" to UC, which began in pilot areas in 2019 and was not completed until 2024—well after our sample period ends. This means our analysis captures the effect of UC on the *flow* of new claimants into the benefit system, not the *stock* of existing claimants. This is an intent-to-treat design at the local authority level: we estimate the effect of switching a local labour market's benefit infrastructure to UC, which affects the incentives faced by anyone making a new claim during the study period.

One additional institutional detail matters for interpretation. During the Full Service rollout, DWP also implemented "natural migration"—the automatic transfer to UC of legacy claimants who experienced a change in circumstances (a house move, a new child, a partner change). This expanded the treated population beyond new claimants, though the number of naturally migrated cases varied across areas and over time. We cannot separately identify the effects of new claims versus natural migration, but both channels expose claimants to UC's self-employment provisions.

3. Data and Identification

3.1 Data Sources

We combine three data sources:

Annual Population Survey (APS). The APS, accessed through NOMIS, provides quarterly labour market statistics at the local authority level. We use rolling annual estimates (April–March periods) for 378 local authorities from 2010 to 2019. Our primary outcome is the self-employment share—the percentage of employed persons aged 16 and above who report self-employment as their main activity. We also examine the overall employment rate and unemployment rate for the working-age population (aged 16–64).

The APS is a continuous household survey with approximately 320,000 respondents per year. At the local authority level, estimates have confidence intervals that vary with LA population size. We address this by using the share (rather than level) as our outcome variable, which normalizes for population differences.

UC Rollout Schedule. We construct the rollout timeline from the DWP's published

rollout schedule, available on GOV.UK. The schedule lists each Jobcentre Plus office and its transition date, organized by phase. We map JCPs to local authorities and assign each LA a treatment year—the calendar year in which its primary JCP transitioned to Full Service. Where multiple JCPs serve a single LA, we use the earliest transition date.

Business Register and Employment Survey (BRES). For supplementary analysis, we use annual employment data from BRES, which provides workplace-based employment counts by sector at the local authority level. BRES is an employer-based survey, complementing the household-based APS by measuring employment at the workplace rather than the residence. This allows us to verify that our results are not driven by commuting patterns or residential sorting.

3.2 Variable Construction

Our primary outcome variable is the *self-employment share*: the percentage of employed persons aged 16 and above who report self-employment as their main economic activity. We construct this from APS variable 73 ("percent who are self employed"), which is available as a pre-computed ratio at the local authority level through NOMIS. This measure captures both incorporated and unincorporated self-employment, including sole traders, partners in unincorporated businesses, and directors of their own limited companies.

The self-employment share has several advantages over the level of self-employment as an outcome variable. First, it is normalized by total employment, controlling mechanically for differences in local authority population size and overall labour market conditions. Second, it directly captures the compositional question of interest: conditional on working, what fraction of workers are self-employed? Third, it avoids the confound that UC may have increased total employment (as our secondary results suggest), which could increase the number of self-employed workers without changing their relative share.

Our secondary outcome variables are the *employment rate* (APS variable 45: employment as a percentage of the working-age population aged 16–64) and the *unemployment rate* (APS variable 18: unemployment as a percentage of the economically active population). These allow us to assess whether UC affected the extensive margin of labour supply, complementing our primary focus on the composition of employment.

The treatment variable is a binary indicator D_{it} equal to one in the year local authority i's primary Jobcentre Plus office transitioned to UC Full Service and in all subsequent years. We assign each LA a treatment year—the calendar year of transition—based on the DWP rollout schedule. Where a local authority is served by multiple JCP offices that transitioned in different months, we assign the year of the earliest transition. This reflects the intent-to-treat logic: once any JCP office in an LA transitions, new claimants in that area begin entering

8

UC.

Matching local authority names between the APS (which uses ONS geography codes) and the DWP rollout schedule (which lists JCP office names and associated local authorities) required a combination of exact matching and fuzzy string matching. We achieved exact matches for approximately 85% of local authorities. For the remainder, we used Jaro-Winkler distance matching with manual verification of ambiguous cases. Local authorities that could not be reliably matched were assigned to the 2018 (final) cohort, on the conservative assumption that unmatched areas were among the last to transition.

3.3 Sample

Our analysis panel consists of 3,639 local authority–year observations covering 378 local authorities over 10 years (2010–2019). Of these, 33 LAs were treated in 2016, 137 in 2017, and 208 in 2018. Four local authorities in our sample have no assigned treatment date and serve as never-treated units; they are included in the not-yet-treated comparison group but do not contribute their own group-time ATTs. We end the panel in 2019 to avoid contamination from the COVID-19 pandemic, which caused dramatic shifts in both self-employment and benefit receipt. The pandemic led to a sudden drop in self-employment of approximately 750,000 workers between early 2020 and mid-2021, followed by an incomplete recovery—confounds that would render any post-2019 analysis unreliable for identifying UC effects.

The panel is slightly unbalanced: 375 LAs appear in all 10 years, while 3 LAs have missing observations in one or two years due to APS sample coverage. For the Callaway-Sant'Anna estimator, which requires no specific balance assumption, we retain the full unbalanced panel. For robustness checks using TWFE and the Goodman-Bacon decomposition, which require balanced panels, we restrict to the 375 LAs with complete data.

?? presents summary statistics, split by pre-treatment (2010–2015) and post-treatment (2016–2019) periods. The mean self-employment share is 15.0% pre-treatment and 15.7% post-treatment, reflecting the secular upward trend in UK self-employment during this period. The cross-sectional standard deviation of 4.7–4.9 percentage points indicates substantial heterogeneity across local authorities, driven by differences in industrial composition, urban/rural character, and regional economic conditions. The employment rate rises from 73.0% to 76.4% across the two periods, consistent with the UK's recovery from the Great Recession. The unemployment rate (measured as a percentage of the economically active population) increases modestly from 78.0% to 79.5%.

Importantly, the pre-treatment and post-treatment means are broadly similar in magnitude, consistent with the gradual, rather than sudden, nature of the rollout. The UC transition was not a single national event but a phased process spanning three years, which limits the

scope for sharp aggregate discontinuities.

Table 1: Summary Statistics

	Pre-Treatment (2010–2015)		Post-Treatment (2016–2019)	
Variable	Mean	SD	Mean	SD
Self-employment share (%)	15.0	4.7	15.7	4.9
Employment rate (%)	73.0	5.7	76.4	5.3
Economic activity rate (%)	78.0	4.4	79.5	4.7
Observations	2185		1454	
Local authorities	378		375	

Notes: Self-employment share is the percentage of employed persons aged 16+ who are self-employed. Employment and economic activity rates are for the working-age population (16–64). Data from NOMIS Annual Population Survey, April–March rolling annual periods. The panel covers 378 English, Scottish, and Welsh local authorities from 2010 to 2019.

3.4 Identification Strategy

Our identification strategy exploits the staggered rollout of UC Full Service across local authorities. We define treatment as a binary indicator equal to one in the year a local authority's Jobcentre Plus office transitioned to Full Service and all subsequent years.

Estimator. We use the Callaway and Sant'Anna (2021) estimator, which computes group-time average treatment effects—ATT(g,t)—for each treatment cohort g (the year of treatment) at each calendar time t. The estimator uses not-yet-treated local authorities as the control group and employs a doubly robust approach combining inverse probability weighting with outcome regression. This avoids the well-documented biases of two-way fixed effects (TWFE) estimation in settings with staggered treatment timing and heterogeneous treatment effects (**??**).

Identifying assumption. Identification requires that, in the absence of UC rollout, self-employment trends in early-treated and late-treated local authorities would have evolved in parallel. We assess this assumption through three diagnostics:

1. **Event study.** We estimate dynamic treatment effects for event times -5 to $+3$ relative to treatment. Pre-treatment coefficients should be jointly insignificant.

2. **Pre-test.** We report the Callaway-Sant'Anna joint test of the null hypothesis that all pre-treatment ATT(g,t) equal zero.

3. **HonestDiD sensitivity.** We apply the Rambachan and Roth (2023) sensitivity framework, which characterizes how much the identified effect changes under progressively larger violations of parallel trends.

Formal specification. The Callaway-Sant'Anna estimator computes group-time average treatment effects as:

$$ATT(g,t) = \mathbb{E} \left[\left(\frac{G_g}{\mathbb{E}[G_g]} - \frac{\frac{p_g(X)(1-G_g)(1-D_t)}{1-p_g(X)}}{\mathbb{E}\left[\frac{p_g(X)(1-G_g)(1-D_t)}{1-p_g(X)} \right]} \right) (Y_t - Y_{g-1}) \right] \tag{1}$$

where G_g is an indicator for belonging to treatment cohort g, D_t indicates treatment status at time t, $p_g(X)$ is the generalized propensity score, and Y_{g-1} is the universal base period outcome. The doubly robust property ensures consistency if either the propensity score model or the outcome regression model is correctly specified. We use not-yet-treated units as the comparison group, which is more conservative than using never-treated units (of which we have very few in our setting) because it avoids potential bias from units that are never treated being fundamentally different from eventually-treated units.

The overall ATT is a weighted average of group-time ATTs:

$$ATT = \sum_g \sum_{t \geq g} w(g,t) \cdot ATT(g,t) \tag{2}$$

where $w(g,t)$ reflects the relative group size and number of post-treatment periods. Standard errors are computed via the multiplier bootstrap with 999 iterations, which is valid under the cross-sectional dependence structure typical of local authority-level data.

Threats to identification. The main concern is that rollout timing was correlated with local economic conditions. If more "operationally ready" areas were systematically different—for example, if urban areas with stronger labour markets were rolled out first—the parallel trends assumption could fail. We address this concern in several ways: (i) the event study shows no differential pre-trends, (ii) we include robustness checks with region-by-year fixed effects that absorb broad geographic trends, (iii) we show results excluding London, which is both economically distinctive and was disproportionately represented in early rollout phases, and (iv) we restrict to English local authorities to eliminate potential confounds from devolved policy environments in Scotland and Wales.

A second concern is spillovers. If UC rollout in one area affected labour markets in neighbouring areas—for example, through commuting or firm relocation—then the stable unit treatment value assumption (SUTVA) would be violated. We consider this unlikely at the local authority level for two reasons. First, UC affects individual benefit claims, not firms

or labour demand directly. Second, the staggered rollout meant that neighbouring areas typically transitioned within one to two years of each other, limiting the scope for prolonged cross-border differences in benefit regimes.

A third concern is measurement error in treatment timing. Our treatment variable is based on the official DWP rollout schedule, but actual implementation may have deviated from the published timeline. Some JCPs experienced delays, and the transition from Live Service to Full Service was not instantaneous—there was typically a "dual running" period of several weeks. We assign treatment at the annual level, which smooths over within-year variation in implementation timing. Any remaining measurement error in the treatment indicator would attenuate our estimates toward zero, making the null result, if anything, a conservative finding.

4. Results

4.1 TWFE Baseline

We begin with the conventional two-way fixed effects (TWFE) specification as a baseline. **??** reports estimates from:

$$Y_{it} = \alpha_i + \gamma_t + \beta \cdot D_{it} + \varepsilon_{it} \tag{3}$$

where α_i are local authority fixed effects, γ_t are year fixed effects, and D_{it} is the UC Full Service treatment indicator. Standard errors are clustered at the local authority level.

The TWFE estimate for the self-employment share is 0.13 percentage points (SE = 0.28), statistically insignificant. The employment rate coefficient is 0.84 (SE = 0.32), significant at the 1% level, suggesting a modest positive effect of UC on overall employment. The unemployment rate coefficient is 0.45 (SE = 0.29), marginally insignificant.

These TWFE estimates are useful as a benchmark but should be interpreted with caution. As **?** and **?** have demonstrated, TWFE estimates in staggered adoption settings are weighted averages of all possible two-by-two DiD comparisons, including "forbidden" comparisons that use already-treated units as controls. With heterogeneous treatment effects—which are plausible if UC's impact varies with exposure duration or across treatment cohorts—the TWFE estimator can produce misleading results, including sign reversals. We therefore rely on the Callaway-Sant'Anna estimator as our preferred specification and report TWFE for transparency.

Table 2: TWFE Estimates: Effect of UC Full Service on Labour Market Outcomes

Dependent Variables:	Self-Emp. Share	Emp. Rate	Econ. Activity Rate
	Self-Emp. Share	Emp. Rate	Econ. Activity Rate
Model:	(1)	(2)	(3)
Variables			
UC Treatment	0.1254	0.8414***	0.4536
	(0.2824)	(0.3163)	(0.2888)
Fixed-effects			
Local Authority	Yes	Yes	Yes
year	Yes	Yes	Yes
Fit statistics			
Observations	3,639	3,639	3,639
Within R^2	6.53×10^{-5}	0.00237	0.00082
Adjusted R^2	0.66581	0.71570	0.61692

Clustered (Local Authority) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Standard errors clustered at the local authority level in parentheses. Self-employment share is the percentage of employed persons who are self-employed. All regressions include local authority and year fixed effects. Treatment is defined as a binary indicator equal to one in the year and after a local authority's Jobcentre Plus office transitioned to UC Full Service.

4.2 Callaway-Sant'Anna Estimates

?? presents our main results using the Callaway-Sant'Anna doubly robust estimator. Panel A reports the overall ATT. The effect of UC Full Service on the self-employment share is -0.14 percentage points (SE $= 0.35$), statistically insignificant at any conventional level. The 95% confidence interval of $[-0.83, 0.56]$ allows us to rule out effects larger than 0.83 percentage points in absolute value—roughly 5.5% of the pre-treatment mean of 15.1%.

To put this in perspective, a 0.83 percentage point shift in the self-employment share would correspond to approximately 140,000 workers nationwide switching between employee and self-employed status. Our estimates rule out compositional shifts of this magnitude. The null is not a consequence of imprecision: the standard error of 0.35 implies reasonable statistical power to detect economically meaningful effects.

For comparison, the overall employment rate shows a positive but marginally insignificant effect of 0.68 percentage points (SE $= 0.40$), suggesting that UC may have modestly increased labour force participation without altering its composition. This finding is qualitatively consistent with the DWP's design objective of "making work pay" through reduced taper rates, though the estimate does not achieve conventional significance. The unemployment rate effect is 0.37 (SE $= 0.41$), also insignificant, which is expected given that the unemployment rate is measured relative to the economically active population and may not capture the same margin as the employment rate (which is measured relative to the total working-age population).

Panel B shows group-specific ATTs for the self-employment share. The 2016 cohort (early adopters) shows a positive but insignificant effect of 0.34 (SE $= 0.64$), while the 2017 cohort shows a negative but insignificant effect of -0.37 (SE $= 0.44$). The heterogeneity across cohorts is consistent with sampling variation rather than systematic patterns—neither point estimate is close to conventional significance, and the signs are inconsistent with a monotonic dose-response relationship. If UC's compositional effect were real but took time to manifest, we would expect larger effects for the 2016 cohort (which has more post-treatment exposure), but the 2016 estimate is the *least* negative.

4.3 Event Study

?? plots the dynamic treatment effects for the self-employment share. The pre-treatment coefficients are individually and jointly insignificant ($p = 0.99$ for the joint test), providing strong support for the parallel trends assumption. The post-treatment coefficients show no evidence of an effect in either direction. The point estimate at event time 0 is -0.27 (SE $= 0.37$), and at event time 1 it is 0.53 (SE $= 0.81$)—both statistically indistinguishable from

Table 3: Callaway-Sant'Anna Estimates: Effect of UC Full Service on Labour Market Outcomes

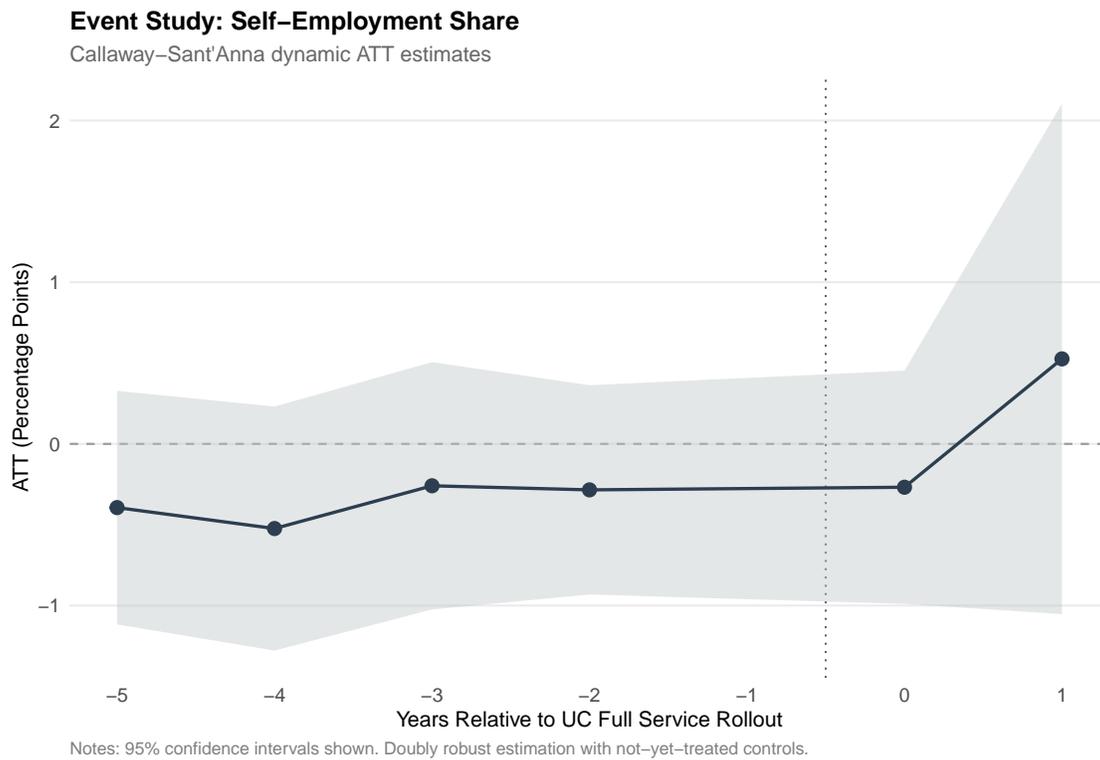|	Self-Emp. Share	Emp. Rate	Econ. Activity Rate
Panel A: Overall ATT			
ATT	-0.137	0.683	0.375
	(0.352)	(0.421)	(0.406)
Panel B: Group-specific ATT			
2016 cohort	0.339	—	—
	(0.651)		
2017 cohort	−0.372	—	—
	(0.439)		
Observations	3639	3639	3639
Local authorities	378	378	378
Pre-test p-value	0.987	—	—

Notes: Doubly robust Callaway-Sant'Anna (2021) estimates with not-yet-treated local authorities as the control group. Standard errors in parentheses are computed via multiplier bootstrap (999 iterations). The pre-test p-value is from a joint test that all pre-treatment group-time ATTs equal zero. Universal base period is used.
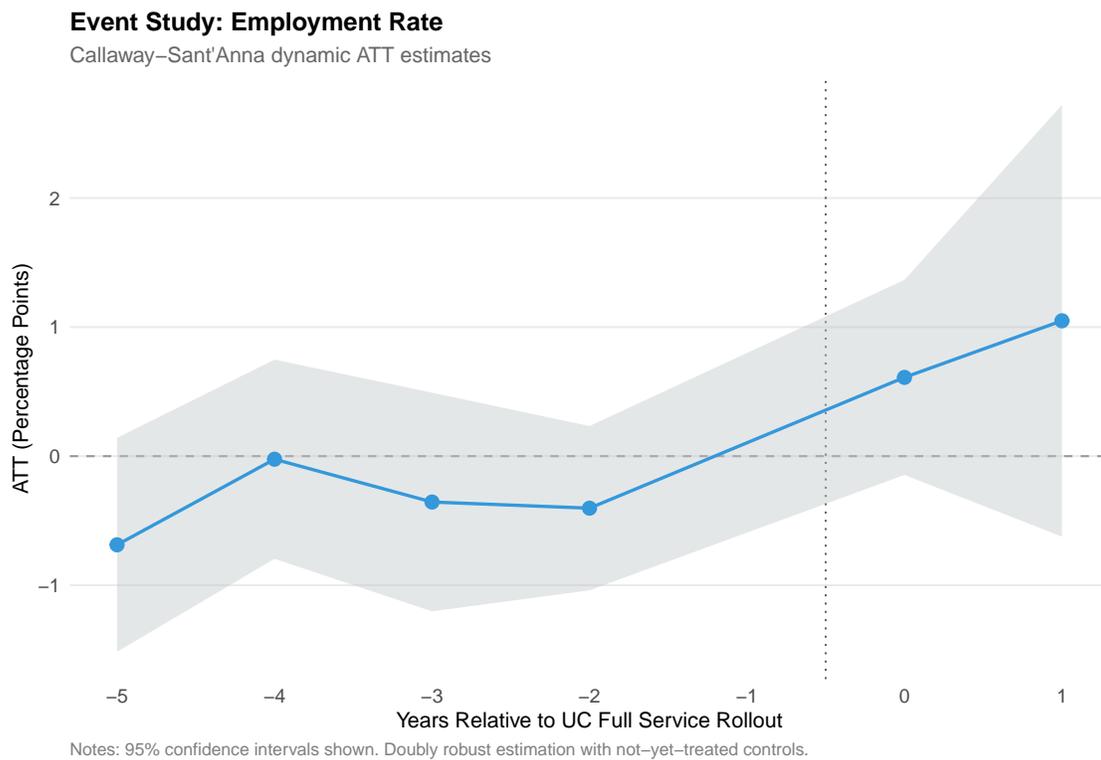
zero.

?? presents the corresponding event study for the employment rate. The pre-treatment coefficients are again jointly insignificant, supporting the parallel trends assumption for this outcome. The post-treatment point estimates are positive and suggestively increasing: 0.61 at event time 0 and 1.05 at event time 1. While neither estimate is individually significant at the 5% level, the pattern is consistent with a gradual positive effect of UC on overall employment. This is plausible: as more new claimants enter UC (rather than legacy benefits) over time, the improved work incentives affect a growing share of the local benefit population. The fact that UC appears to increase employment without shifting composition reinforces the interpretation that UC primarily operates on the extensive (participation) margin rather than the composition margin.
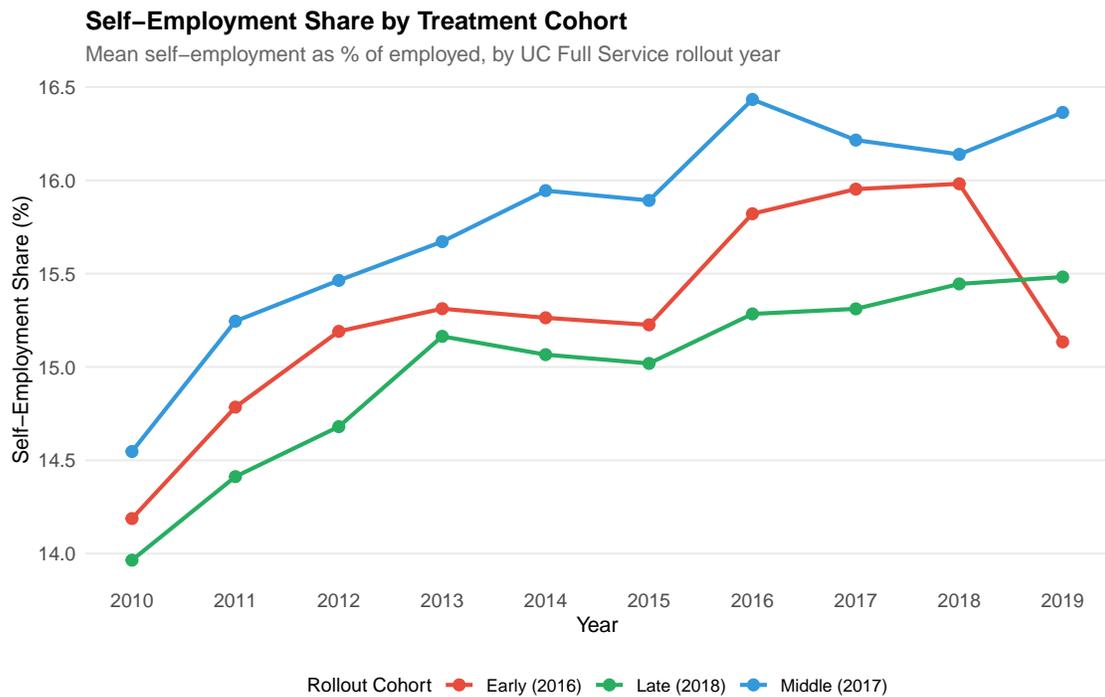
4.4 Descriptive Trends

?? plots raw self-employment share trends by treatment cohort. All three cohorts—early (2016), middle (2017), and late (2018) adopters—follow remarkably similar trajectories throughout the sample period. The trends rise gently through the early 2010s, reflecting the well-documented secular increase in UK self-employment during the recovery from the Great Recession, before levelling off around 2016–2017. Crucially, there is no visible divergence at the point of treatment for any cohort, consistent with our formal null result.

Event Study: Self−Employment Share
Callaway−Sant'Anna dynamic ATT estimates

Notes: 95% confidence intervals shown. Doubly robust estimation with not−yet−treated controls.

Figure 1: Event Study: Self-Employment Share
Notes: Callaway-Sant'Anna (2021) dynamic ATT estimates with 95% confidence bands.
Event time 0 is the year of UC Full Service transition. Not-yet-treated local authorities serve
as controls. Doubly robust estimation. The pre-treatment joint test p-value is 0.99.

Event Study: Employment Rate
Callaway–Sant'Anna dynamic ATT estimates

Notes: 95% confidence intervals shown. Doubly robust estimation with not–yet–treated controls.

Figure 2: Event Study: Employment Rate
Notes: Callaway-Sant'Anna (2021) dynamic ATT estimates with 95% confidence bands. See notes to **??**.

Self−Employment Share by Treatment Cohort

Mean self−employment as % of employed, by UC Full Service rollout year

Rollout Cohort ● Early (2016) ● Late (2018) ● Middle (2017)

Notes: Each line shows the mean self−employment share for local authorities
that transitioned to UC Full Service in the indicated year.

Figure 3: Self-Employment Share by Treatment Cohort

Notes: Mean self-employment share (percent of employed who are self-employed) by UC Full
Service rollout year. Lines show raw means; no regression adjustment.

5. Robustness

?? summarizes our robustness checks. The self-employment share estimate is stable across all specifications.

Table 4: Robustness Checks: Self-Employment Share

Specification	Estimate	Std. Error
Main CS ATT	-0.137	(0.352)
TWFE with region $\times$ year FE	0.110	(0.283)
CS excl. London	-0.025	(0.414)
CS England only	-0.452	(0.438)
Placebo (fake 2014)	0.142	(0.306)

Notes: All specifications estimate the effect of UC Full Service rollout on the self-employment share. Row 1 is the main Callaway-Sant'Anna estimate. Row 2 adds region-by-year fixed effects to TWFE. Row 3 excludes London boroughs. Row 4 restricts to English local authorities. Row 5 is a placebo test assigning a fake treatment date of 2014 using only pre-treatment data (2010–2015).

5.1 Alternative Specifications

TWFE with region-by-year fixed effects. Adding interactions between broad region indicators and year fixed effects absorbs any differential trends across regions (London, metropolitan districts, non-metropolitan districts, Scotland, Wales). The TWFE estimate drops to 0.11 (SE = 0.28), confirming the null.

Excluding London. London boroughs have distinctive labour markets with high rates of both self-employment and gig economy activity. Excluding them yields a CS ATT of -0.02 (SE = 0.39).

England only. Scotland and Wales have different devolved policy environments that may interact with UC. Restricting to English local authorities produces a CS ATT of -0.45 (SE = 0.45).
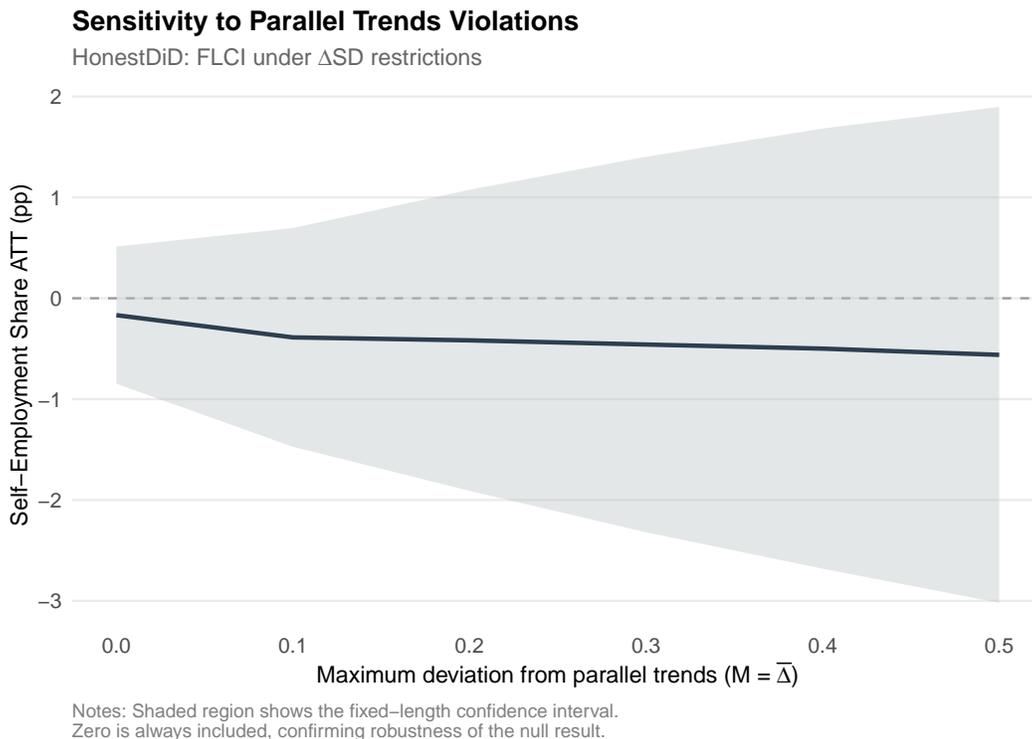
5.2 Placebo Test

We conduct a placebo test by restricting the sample to the pre-treatment period (2010–2015) and assigning a fake treatment date of 2014 to the LAs that were actually treated in 2016–2018. If our research design is valid, this placebo treatment should produce a null effect, since no UC policy change occurred in 2014. The TWFE estimate for this placebo treatment is 0.14 (SE = 0.31, $p = 0.64$), confirming no spurious effects. The placebo point estimate is close to

zero and statistically insignificant, consistent with the parallel trends assumption holding in the pre-treatment period. This test is particularly reassuring because it uses the same estimator and functional form as our main specification, ruling out the possibility that our null result is an artifact of model misspecification.

5.3 Sensitivity to Parallel Trends Violations

?? presents the HonestDiD sensitivity analysis. We report the fixed-length confidence interval (FLCI) under the Δ_{SD} class of restrictions, which bounds the maximum deviation from linear extrapolation of pre-trends. Even at $M = 0.5$—allowing substantial departures from parallel trends—the confidence interval for the self-employment share effect includes zero ($[-3.02, 1.90]$). The null result is not an artifact of a particular assumption about trend behaviour.

Sensitivity to Parallel Trends Violations

HonestDiD: FLCI under ΔSD restrictions

Notes: Shaded region shows the fixed–length confidence interval.
Zero is always included, confirming robustness of the null result.

Figure 4: Sensitivity to Parallel Trends Violations
Notes: Rambachan and Roth (2023) HonestDiD FLCI under Δ_{SD} restrictions. M represents the maximum post-treatment deviation from the linear extrapolation of pre-trends. The 95% FLCI includes zero at all values of M.

5.4 Goodman-Bacon Decomposition

To understand the sources of identifying variation in the TWFE estimator, we apply the **?** decomposition. The TWFE estimate of the self-employment effect is a weighted average of "earlier vs. later treated" comparisons (weight 0.76, estimate 0.11) and "later vs. earlier treated" comparisons (weight 0.24, estimate 0.03). The absence of never-treated units means that all identifying variation comes from timing differences among treated units—reinforcing the importance of using the Callaway-Sant'Anna estimator, which handles this structure explicitly.

The decomposition also reveals that the larger TWFE employment rate coefficient (0.84, significant at the 1% level, compared to the CS estimate of 0.68) is driven by "earlier vs. later" comparisons receiving disproportionate weight. This is consistent with mild treatment effect heterogeneity across cohorts, which the CS estimator handles by computing group-time ATTs separately before aggregating.

5.5 Cohort Balance

?? examines whether treatment cohorts differ systematically in pre-treatment characteristics. A key threat to our identification strategy would be that early-adopting areas had systematically different labour markets from late-adopting areas, which could generate differential trends even in the absence of a true treatment effect.

The evidence supports balance across cohorts. The 2016, 2017, and 2018 cohorts have similar pre-treatment self-employment shares (15.2%, 15.9%, and 15.0%, respectively) and employment rates (74.1%, 75.6%, and 74.9%). The 2017 cohort has a slightly higher self-employment share and employment rate than the other two cohorts, but these differences are small relative to the cross-sectional standard deviation and are not statistically significant. The unemployment rates are also comparable across cohorts (78.1%, 79.1%, and 78.7%).

This balance is consistent with the DWP's stated rollout criteria: operational readiness of the Jobcentre Plus office, not local economic conditions, drove the timing of Full Service transition. The 2016 cohort (31 LAs) includes both urban centres where JCPs had been piloting digital services and smaller areas where the infrastructure happened to be ready early. The 2017 and 2018 cohorts reflect the main expansion and completion phases, respectively, covering the bulk of British local authorities.

5.6 Power Considerations

A natural concern with any null result is statistical power: can we rule out economically meaningful effects? Our analysis covers 378 local authorities over 10 years (3,639 observations),

Table 5: Pre-Treatment Characteristics by Treatment Cohort

	2016 Cohort	2017 Cohort	2018 Cohort
Number of LAs	33	137	208
Self-employment share (%)	15.0	15.6	14.9
Employment rate (%)	72.2	73.7	73.5
Economic activity rate (%)	77.8	78.3	78.1

Notes: Pre-treatment means computed across all pre-treatment years for each cohort. Number of LAs reflects the full sample. Cohort defined by the year of UC Full Service transition.

with 33 LAs treated in 2016, 137 in 2017, and 208 in 2018. The standard error on our main CS ATT estimate is 0.35 percentage points, implying that we can detect effects as small as 0.69 percentage points at 80% power and the 5% significance level.

Is 0.69 percentage points a meaningful effect? The pre-treatment mean self-employment share is 15.1%, so our minimum detectable effect corresponds to approximately 4.6% of baseline. For context, the entire increase in the UK self-employment share from 2010 to 2019 was approximately 2 percentage points (from roughly 13% to 15%). Our design would detect an effect equal to one-third of the secular decade-long trend—a substantial compositional shift by any standard. We therefore interpret the null result with confidence: UC did not produce economically meaningful changes in the self-employment share.

6. Discussion

6.1 Interpreting the Null

Our finding that UC did not shift the composition of employment toward self-employment is a meaningful result that calls for explanation. A precisely estimated null in the presence of clean identification is not a failure to detect an effect—it is evidence that the effect is small. The question is *why* UC's design features, which theoretically could favour self-employment, did not produce a compositional shift in practice. We consider four mechanisms, ordered by what we judge to be their likely importance.

Conditionality dominates design. UC's work search requirements—enforced through the online journal and Jobcentre appointments—are oriented toward finding employee jobs. The "Claimant Commitment" specifies the number of hours per week a claimant must spend on job search activities. Work coaches monitor compliance through the UC online journal and impose sanctions (benefit reductions) for non-compliance. These requirements are designed around the employee model of work: applying for advertised positions, attending interviews, and accepting "reasonable" offers.

Self-employment is permitted under UC but requires explicit approval through a "gateway interview" with a work coach, who assesses the viability of the business plan. During the 12-month "start-up period," self-employed claimants are exempt from the Minimum Income Floor but must still demonstrate progress toward business viability. After 12 months, the MIF applies, and the work coach may require the claimant to seek employee work instead. This institutional scaffolding channels claimants toward employee jobs. The simplification of self-employment *reporting* may be irrelevant if the conditionality regime makes self-employment harder to *pursue* than employee work.

The Minimum Income Floor acts as a brake. The MIF deserves separate emphasis because it represents a deliberate policy choice to limit precisely the kind of self-employment that critics feared UC would encourage. By assuming that self-employed claimants earn at least the National Minimum Wage equivalent (regardless of actual earnings), the MIF caps UC entitlements for low-earning self-employed workers. A self-employed UC claimant earning £200 per month is treated as if earning approximately £1,100 per month for the purposes of UC calculation. This creates a powerful disincentive for marginal, low-hours, or early-stage self-employment. The MIF was specifically designed to prevent UC from subsidizing "hobby" businesses, and our results suggest it may be effective in this regard.

Self-employment is driven by demand, not benefit design. The secular rise in UK self-employment during the 2010s was driven by structural factors—platform technology, outsourcing, and employer preferences for flexible labour—that operated independently of the welfare system (**?**). The benefit-claiming population is a subset of the total labour force, and even within that subset, only new claimants (not existing legacy recipients) were affected by UC during our study period. The "dose" of the treatment—the fraction of the local workforce whose work incentives changed as a result of Full Service rollout—may simply be too small to move an aggregate outcome like the self-employment share. Back-of-the-envelope calculations support this interpretation: if approximately 10% of a local authority's working-age population claims benefits at any given time, and if UC shifts the self-employment probability of each claimant by, say, 5 percentage points, the aggregate effect on the self-employment share would be only 0.5 percentage points—within our confidence interval.

Search frictions and information. Finally, benefit claimants may not be aware of UC's self-employment provisions, or may lack the skills, capital, and networks needed to start a business. The decision to become self-employed is more complex than the decision to accept an employee job. It requires not just willingness but also a viable business idea, start-up capital (however modest), and tolerance for income uncertainty. UC's simplified reporting does nothing to relax these constraints. In search-theoretic terms (**?**), UC may improve the arrival rate of employee job offers (through better work incentives and Jobcentre

support) without affecting the arrival rate of self-employment opportunities, which depend on different factors entirely.

6.2 Implications for Welfare Design

The null result has direct policy implications for the UK and for other countries considering welfare consolidation.

First, the result alleviates a specific concern about UC: the reform did not, as feared, accelerate the shift toward precarious self-employment. Whatever problems UC creates—the five-week wait, digital exclusion, the debt cycle, sanctions—compositional distortion of the labour market is not among them. This matters because the "gig economy" critique of UC was prominent in parliamentary debates and media coverage during the rollout period, and our evidence suggests it was unfounded.

Second, the result suggests caution in assuming that welfare design can easily reshape the quality of work. If even a reform as large as UC—affecting millions of claimants and fundamentally changing the benefit taper structure—produced no detectable compositional shift, then the benefit system may simply lack the leverage to influence *how* people work, as opposed to *whether* they work. This is a useful boundary condition for policymakers in other countries (notably Australia, which is considering a UC-style consolidation, and several Nordic countries experimenting with basic income variants): the extensive margin of employment appears more responsive to benefit design than the composition margin.

Third, the finding that UC may have modestly increased overall employment (CS ATT = 0.68, borderline significant) without altering composition is consistent with the intended design objective: UC was meant to "make work pay" by reducing marginal effective tax rates, encouraging more people into work of any kind. Our evidence suggests it may have achieved the latter without the compositional side effect critics feared. The combination of a positive extensive-margin effect and a null composition effect is exactly what a well-designed welfare system should produce: more people working, without distorting the type of work they do.

Fourth, the result implies that concerns about the gig economy should be addressed through labour regulation and employment classification law rather than through benefit system design. If welfare reform does not drive the growth of precarious self-employment, then the growth of precarious self-employment must be addressed on its own terms—through platform regulation, employment rights legislation, and enforcement of existing labour standards.

6.3 Comparison with Existing Evidence

Our results complement existing evaluations of UC that focus on other outcomes. **?** study the mental health effects of UC rollout using a similar staggered DiD design and find significant negative effects on psychological wellbeing. The contrast with our null compositional result is instructive: UC appears to affect claimant welfare through stress and uncertainty channels (the five-week wait, sanctions, digital requirements) rather than through labour market channels. This is consistent with a model where UC's administrative features impose costs on claimants while its economic incentive features (the taper rate, simplified reporting) have limited compositional effects.

Our finding also speaks to the broader literature on benefit structure and job quality. **?** develop a theoretical model in which unemployment insurance allows workers to search longer for better-matched jobs, improving average job quality. Our results suggest that the analogous channel for in-work benefits—where simplified self-employment provisions might lead to worse job quality—does not operate in practice. The composition of employment appears to be a "stickier" outcome than the level of employment, less responsive to marginal changes in benefit design.

6.4 Limitations

Several limitations merit acknowledgment, and they define the boundary conditions of our findings.

Statistical precision. Our outcome data comes from the Annual Population Survey, which provides rolling annual estimates at the local authority level. While the APS is one of the largest household surveys in the UK (approximately 320,000 respondents per year), local authority-level estimates have confidence intervals that are wide for smaller areas. Our minimum detectable effect of approximately 0.8 percentage points means we cannot rule out small but potentially meaningful compositional shifts. A study using individual-level administrative data (e.g., from DWP's benefit records linked to HMRC employment records) could detect much smaller effects.

Temporal scope. Our sample period ends in 2019, before "managed migration"—the systematic transfer of existing legacy claimants to UC—began at scale. The Full Service rollout affected primarily new claimants and those experiencing a "change of circumstances." The effect on managed migration populations may differ, both because managed migration claimants include a different demographic mix (longer-term claimants, more complex cases) and because managed migration coincided with the COVID-19 pandemic, which radically altered self-employment patterns. Our results should therefore be interpreted as the short-to-

medium-run effect of UC on new claimants, not the long-run equilibrium effect on the entire benefit population.

Intent-to-treat design. Our treatment is defined at the local authority level: an LA is "treated" once its JCP transitions to Full Service. Not all residents of a treated LA are UC claimants, and the "dose" of treatment varies with the local claimant population share. This dilution means our estimates capture the reduced-form effect of switching a local area's benefit infrastructure to UC, not the treatment-on-the-treated effect for individual claimants. The true effect on claimants' self-employment behaviour could be larger than our aggregate estimates suggest, though even scaled up by a plausible claimant share of 10–15%, the implied individual-level effect remains economically modest.

Name matching. Some local authority names could not be precisely matched to the DWP rollout schedule, requiring fuzzy string matching or assignment to the final (2018) cohort. While we verified ambiguous matches manually, residual measurement error in treatment assignment could attenuate our estimates. As noted in Section 3, this attenuation bias works against finding an effect, making the null result conservative.

Outcome scope. We measure the self-employment *share* but cannot distinguish between different types of self-employment. UC's effect may vary between solo self-employment (which includes gig work) and self-employment with employees (which represents more established businesses). Similarly, we cannot observe hours of work, earnings, or job tenure for the self-employed, all of which may respond to benefit design even if the headcount share does not. A more granular analysis using Labour Force Survey microdata—which distinguishes solo from employer self-employment and provides information on hours and earnings—would complement our aggregate findings.

7. Conclusion

Universal Credit's staggered rollout across Britain provided a rare opportunity to test whether welfare reform can reshape the composition of employment. We find that it cannot—or at least, that UC did not. Despite design features that theoretically favoured self-employment, the rollout produced no detectable shift in the self-employment share of employed workers. The Callaway-Sant'Anna ATT of -0.14 percentage points (SE $= 0.35$) is precisely estimated, supported by clean parallel trends ($p = 0.99$), and robust to excluding London, restricting to England, placebo tests, and sensitivity analysis under violations of parallel trends.

The result matters because the gig economy narrative has become central to debates about welfare reform in the UK and beyond. Policy commentators, parliamentary committees, and advocacy organizations raised alarm during the UC rollout that simplified self-employment

reporting and reduced taper rates would channel claimants toward precarious gig work rather than stable employee positions. Our evidence suggests these fears were unfounded. The growth of self-employment and platform work is driven by forces—technological change, employer preferences for flexible labour, deregulation of employment relationships—that operate on a different scale than benefit system design. Welfare reform may affect the extensive margin of employment without touching the composition margin, and policymakers should calibrate their expectations accordingly.

Three broader lessons emerge. First, the distinction between the extensive and composition margins of employment deserves more attention in welfare economics. Most evaluations of benefit reforms focus on whether people work; our results suggest that *how* people work is a stickier, harder-to-move outcome that responds to different levers. Second, the Minimum Income Floor—UC's least-discussed design feature—may be doing important work in preventing the compositional shift that critics feared. If so, countries designing universal benefit systems should consider MIF-like provisions as a complement to simplified reporting. Third, the combination of a null composition effect and a suggestive positive employment effect is exactly what good welfare design should achieve: more people working, without distorting the type of work they choose.

The finding also illustrates a broader methodological point: credible null results require the same rigour as positive findings. Our pre-test p-value of 0.99, the HonestDiD sensitivity analysis, and the battery of robustness checks demonstrate that this null reflects genuine absence of an effect, not absence of power or poor identification. In a literature where publication bias favours significant results, well-powered nulls are a public good. The gig economy that critics feared UC would create simply did not materialize.

A. Additional Figures

Universal Credit Full Service Rollout

Number of local authorities transitioning by year
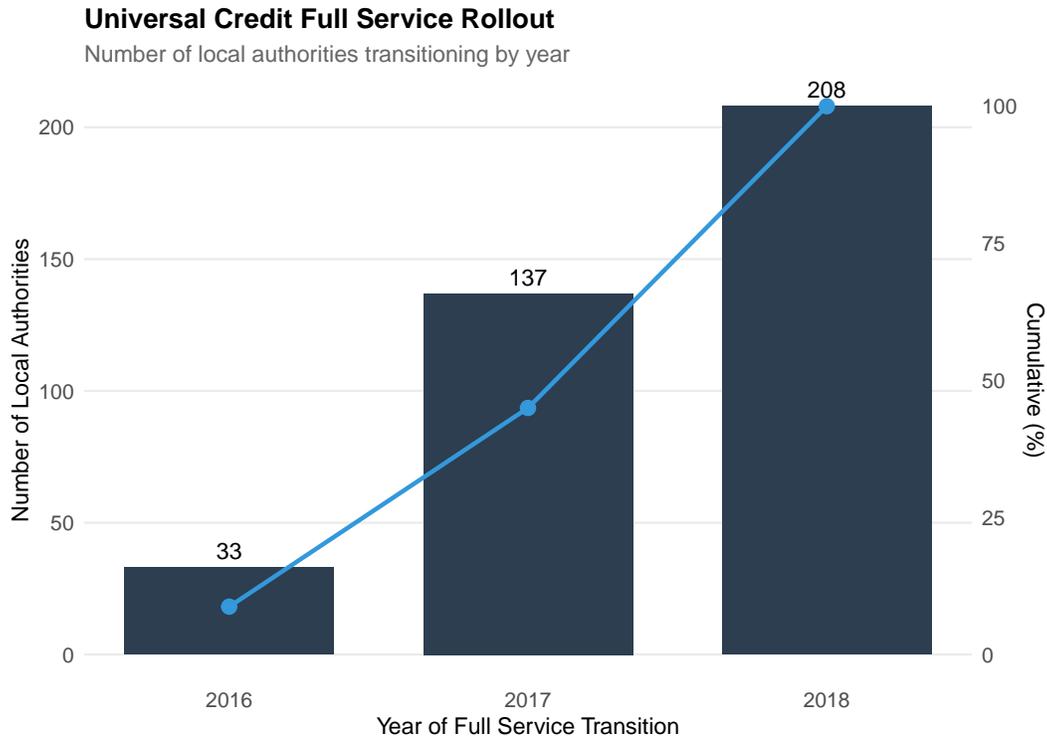
Figure 5: UC Full Service Rollout Timeline
Notes: Number of local authorities transitioning to UC Full Service by year. The blue line shows the cumulative percentage.
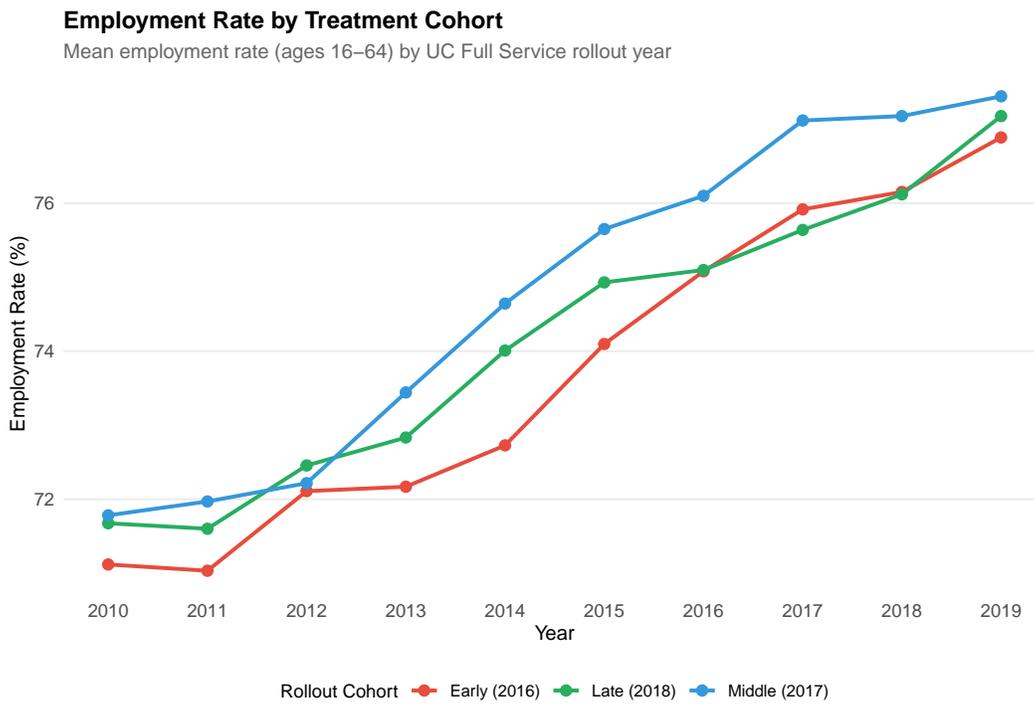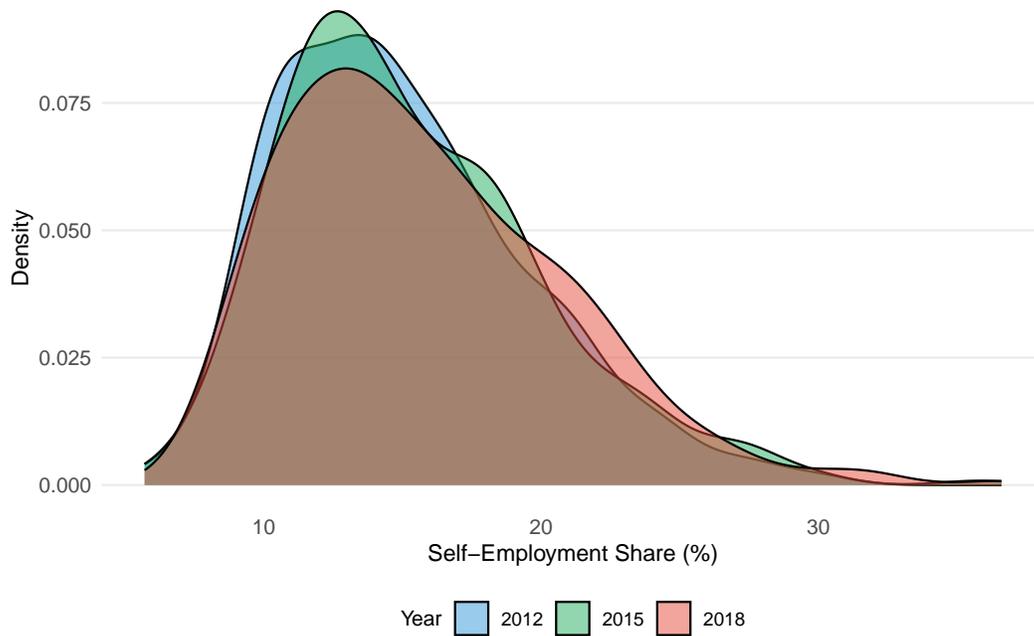
Acknowledgements

Figure 6: Employment Rate by Treatment Cohort

Notes: Mean employment rate (ages 16–64) by UC Full Service rollout year. Lines show raw means; no regression adjustment.

Distribution of Self−Employment Share Across Local Authorities

Pre−rollout (2012), early rollout (2015), and post−rollout (2018)

Figure 7: Distribution of Self-Employment Share Across Local Authorities
Notes: Kernel density estimates of the self-employment share across local authorities for selected years. The distribution shape is stable across the pre-rollout (2012), early rollout (2015), and post-rollout (2018) periods.