Does Welfare Simplification Encourage Entrepreneurship? Evidence from Universal Credit

APEP Autonomous Research* @olafdrw

February 27, 2026

Abstract

Can simplifying welfare systems encourage firm creation? I study the United Kingdom's Universal Credit reform, which consolidated six legacy benefits into a single payment. Exploiting the staggered rollout across 332 Local Authorities between 2015 and 2018, I estimate the reform's effect on new company registrations using Companies House administrative data and the Callaway and Sant'Anna (2021) estimator. I find a precise zero: the point estimate is 0.005 additional firms per 1,000 population per month (95% CI: $[-0.032, 0.042]$), ruling out increases larger than 16 percent of the mean formation rate. Results are consistent across three modern staggered DiD estimators. An exploratory test of the Minimum Income Floor yields imprecise results. Despite removing a labyrinth of six overlapping programs, the most comprehensive welfare simplification in British history did not measurably increase formal business creation.

JEL Codes: J68, H53, L26, I38

Keywords: Universal Credit, welfare reform, entrepreneurship, firm formation, difference-in-differences

*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch

1. Introduction

In 2012, a self-employed person in the United Kingdom claiming welfare navigated a labyrinth: six different benefits, three administrative systems, and effective marginal tax rates that could exceed 90 percent. For a benefit claimant considering starting a business, the system's complexity was itself a barrier—not just to claiming support, but to entering self-employment in the first place. A growing body of evidence documents how benefit design affects employment margins (Chetty, 2008; Krueger and Meyer, 2002), yet one dimension has received almost no attention: *complexity*. When navigating multiple overlapping benefits requires substantial effort and expertise, the system may inadvertently discourage the very activities it means to support—particularly self-employment, where income fluctuates and interacts with multiple benefit streams simultaneously.

The United Kingdom's Universal Credit (UC) reform offers an unusually clean test of whether welfare simplification affects entrepreneurship. Introduced between 2015 and 2018, UC replaced six separate means-tested benefits—Jobseeker's Allowance, Employment and Support Allowance, Income Support, Housing Benefit, Working Tax Credit, and Child Tax Credit—with a single monthly payment governed by a unified 55 percent taper rate. For self-employed claimants, the reform also introduced two consequential features: real-time earnings verification through HMRC integration, and the Minimum Income Floor (MIF), which assumes self-employed claimants earn at least the national minimum wage after a twelve-month start-up period. The net effect on entrepreneurship is theoretically ambiguous. Simplification reduces the informational and administrative barriers to becoming self-employed, but the MIF creates a binding floor that penalizes low earners—precisely the group most likely to be marginal entrepreneurs.

I exploit the staggered rollout of UC's "full service"—the complete digital version available to all claimant types—across more than 300 Local Authorities (LAs) in Great Britain between November 2015 and December 2018. The Department for Work and Pensions (DWP) implemented the rollout in monthly waves, with some LAs gaining access years before others. I construct a novel dataset linking the official DWP rollout schedule to administrative data from Companies House—which records every new company registration in England, Wales, and Scotland with exact incorporation dates and postcodes—and to ONS population estimates for rate normalization.

Using the Callaway and Sant'Anna (2021) estimator for staggered difference-in-differences, which is robust to heterogeneous treatment effects across cohorts, I estimate the causal effect of UC full service on firm formation rates. The identification relies on comparing changes in firm creation across LAs that adopted UC at different times, with not-yet-treated LAs serving

as the comparison group. I find that Universal Credit had no detectable effect on company registrations. The point estimate is a near-zero 0.005 additional firms per 1,000 population per month. More importantly, the results are precise enough to rule out any increase larger than 16 percent of the mean formation rate—a striking null for the most ambitious welfare simplification in British history.

This paper contributes to the literature on welfare reform and entrepreneurship in three ways. First, I provide the first causal estimate of how welfare simplification affects entrepreneurship, using an identification strategy that credibly isolates the reform's effect. The existing literature on UC focuses on employment transitions (Department for Work and Pensions, 2024) and mental health (Brewer et al., 2024); the self-employment margin has been studied only qualitatively (Griffiths et al., 2024). The null result is itself informative: it bounds the magnitude of welfare complexity as a barrier to firm creation and suggests that other frictions—capital access, risk, human capital—may dominate.

Second, I conduct an exploratory test of the Minimum Income Floor's behavioral effects. The MIF's twelve-month start-up grace period creates a change in incentives facing new self-employed UC claimants, and I test whether the aggregate formation effect attenuates after twelve months, when the MIF begins to bind. This test is necessarily imprecise—the twelve-month threshold applies to individual claimants, not to LAs, so the LA-level split is an ecological proxy—but it speaks to an ongoing policy debate about whether the MIF deters viable businesses (Griffiths et al., 2024).

Third, I provide sector-level heterogeneity that illuminates the mechanism. If UC's effect operates through the self-employment simplification channel, it should be concentrated in sectors with high self-employment propensity—construction, professional services, and transport—rather than sectors like public administration or utilities where self-employment is rare. I test this prediction and use public administration as a placebo.

The results contribute to several literatures. Most directly, this paper speaks to the welfare-and-work literature, which has traditionally focused on the employment-unemployment margin (Moffitt, 2002; Blundell, 2006). The self-employment margin is quantitatively important—the UK has approximately 4.2 million self-employed workers, representing 13 percent of total employment (Office for National Statistics, 2019)—but its interaction with benefit design is poorly understood. The paper also contributes to the entrepreneurship literature, which has documented how access to capital (Hurst and Lusardi, 2004), health insurance (Fairlie et al., 2011), and unemployment insurance (Hombert et al., 2020) affect business creation, but has not examined how the *administrative complexity* of the safety net shapes entry decisions. Finally, the paper speaks to the broader literature on program complexity as a policy instrument (Bhargava and Manoli, 2015; Finkelstein and Notowidigdo, 2019), extending

it from take-up decisions to labor market transitions.

The remainder of the paper proceeds as follows. Section 2 describes the UC reform and its implications for self-employment. Section 3 presents the data sources and panel construction. Section 4 describes the empirical strategy. Section 5 presents the main results, heterogeneity analysis, and MIF timing test. Section 6 contains robustness checks and sensitivity analyses. Section 7 discusses the interpretation and limitations of the findings. Section 8 concludes.

2. Institutional Background

2.1 The Legacy Benefit System

Before Universal Credit, the UK's working-age safety net comprised six separate means-tested benefits, each with its own application process, assessment criteria, reporting requirements, and taper rates. A self-employed individual receiving Housing Benefit, Working Tax Credit, and Child Tax Credit would simultaneously navigate three different administrative systems, each with different rules about how business income affected entitlement.

This complexity created three specific barriers to self-employment. First, *informational barriers*: understanding how self-employment income would interact with multiple benefits required expertise that most claimants lacked. The HM Government (2013) noted that the legacy system's complexity "created confusion and uncertainty, discouraging claimants from taking up work." The six benefits operated on different assessment periods—some weekly, others monthly, still others annual—and used inconsistent definitions of income. A claimant considering starting a small business would struggle to predict how fluctuating earnings would affect each benefit stream.

Second, *administrative barriers*: changes in self-employment income triggered recalculations in multiple systems, often with significant delays, creating cash-flow uncertainty. Housing Benefit assessments, for instance, could take weeks to adjust after an income change, during which the claimant faced either overpayments (requiring future repayment) or underpayments (creating financial hardship). This unpredictability was particularly damaging for self-employment, where income naturally varies from month to month.

Third, *marginal tax rate discontinuities*: the interaction of multiple taper rates could create effective marginal tax rates exceeding 90 percent at certain income levels (Mirrlees et al., 2011), making it difficult for claimants to predict the return to additional effort. A self-employed person earning slightly above a threshold might lose more in benefit reductions than they gained in additional income—a "poverty trap" that the legacy system's fragmented design made nearly impossible for claimants to navigate rationally.

The scale of the potentially affected population was substantial. In 2015, approximately 2.5

million working-age individuals claimed one or more of the six legacy benefits that UC would replace. Of these, survey evidence suggests that roughly 15–20 percent expressed interest in self-employment but cited benefit complexity as a barrier (Griffiths et al., 2024). Even if only a fraction of these individuals were marginal enough to be affected by simplification, the potential for a measurable effect on firm formation existed.

2.2 Universal Credit Reform

UC replaced the six legacy benefits with a single monthly payment, calculated using a unified 55 percent taper rate on net earnings above a work allowance. The reform fundamentally changed how the benefit system interacted with self-employment through three key features.

Simplified reporting. UC integrates with HMRC's Real Time Information (RTI) system, which automatically reports earnings from employment. Self-employed claimants report earnings monthly through the UC online portal, rather than navigating multiple reporting systems. This single reporting channel replaces the previously fragmented system where different benefits required different forms, submitted to different agencies, on different schedules. The cognitive and administrative burden of benefit management is substantially reduced.

The start-up period. New self-employed claimants receive a twelve-month "start-up period" during which their UC entitlement is calculated based on actual reported earnings, even if these are very low or zero. This feature effectively provides a subsidy to entrepreneurial experimentation: a claimant can attempt self-employment while retaining full UC entitlement during the crucial early phase when most businesses generate minimal income. Under the legacy system, there was no equivalent provision—starting a business could immediately trigger benefit reductions across multiple programs.

The Minimum Income Floor (MIF). After the start-up period, UC calculates entitlement as if self-employed claimants earn at least the national minimum wage multiplied by their expected hours (typically 35 hours per week). If actual earnings fall below this floor, the claimant receives less UC than they would under actual earnings. The MIF was designed to prevent long-term benefit dependency disguised as self-employment, but critics argue it penalizes viable businesses in their early stages (Griffiths et al., 2024; House of Commons Work and Pensions Committee, 2018). The House of Commons Work and Pensions Committee heard testimony from multiple self-employed UC claimants who reported that the MIF had forced them to abandon viable businesses because early-stage income fell below the floor (House of Commons Work and Pensions Committee, 2018).

The theoretical prediction is ambiguous. The simplification channel—reduced informational barriers, unified taper rate, single reporting system—should increase self-employment

entry on the margin. But the MIF channel works in the opposite direction, creating a new barrier that did not exist under the legacy system. The net effect depends on the relative magnitudes: whether the simplification gains in the first twelve months dominate the MIF-induced deterrence that follows. This is precisely what the MIF timing test is designed to decompose.

2.3 The Staggered Rollout

The UC full service—the complete digital version covering all claimant types—was rolled out across Great Britain's Jobcentre Plus offices in monthly waves between November 2015 and December 2018. The DWP selected the rollout order based primarily on IT infrastructure readiness and Jobcentre capacity, not on local labor market conditions (Department for Work and Pensions, 2018). This is a crucial institutional feature for identification: the as-good-as-random assignment of treatment timing allows for a credible difference-in-differences design.

The rollout proceeded in distinct phases. The first areas—London Borough of Croydon and Southwark—went live in November 2015 as part of an initial pilot. A handful of additional areas joined through 2016, but the pace was slow as the DWP monitored the system's performance. The rollout accelerated dramatically in 2017, with dozens of areas transitioning each month. Most areas had completed the transition by December 2018, though some LAs experienced delays beyond the original schedule. The rollout experienced two "firebreak" periods—August-September 2016 and August-September 2017—during which no new areas were added, creating natural breaks in the treatment timing that contribute identifying variation. In the analysis, I classify the 43 LAs whose full service transition was not completed by December 2018 as "never-treated" during my study window (January 2013 to December 2019). Most of these LAs transitioned in 2019, making them more precisely "late-treated" rather than permanently untreated. For this reason, the preferred Callaway-Sant'Anna specification uses *not-yet-treated* LAs as the comparison group rather than relying exclusively on the never-treated designation. The not-yet-treated comparison pool is larger and does not require that the 43 LAs be permanently untreated—only that they have not yet been treated at the time of comparison.

Figure 1 displays the rollout timeline, showing both the number of LAs transitioning in each month and the cumulative adoption curve. The S-shaped adoption pattern—slow initial adoption, rapid acceleration in 2017–2018, and final completion—is characteristic of IT-driven infrastructure rollouts and reinforces the plausibility of the DWP's stated rationale for the ordering.

Between 2015 and 2018, observationally similar LAs had different exposure to UC simply

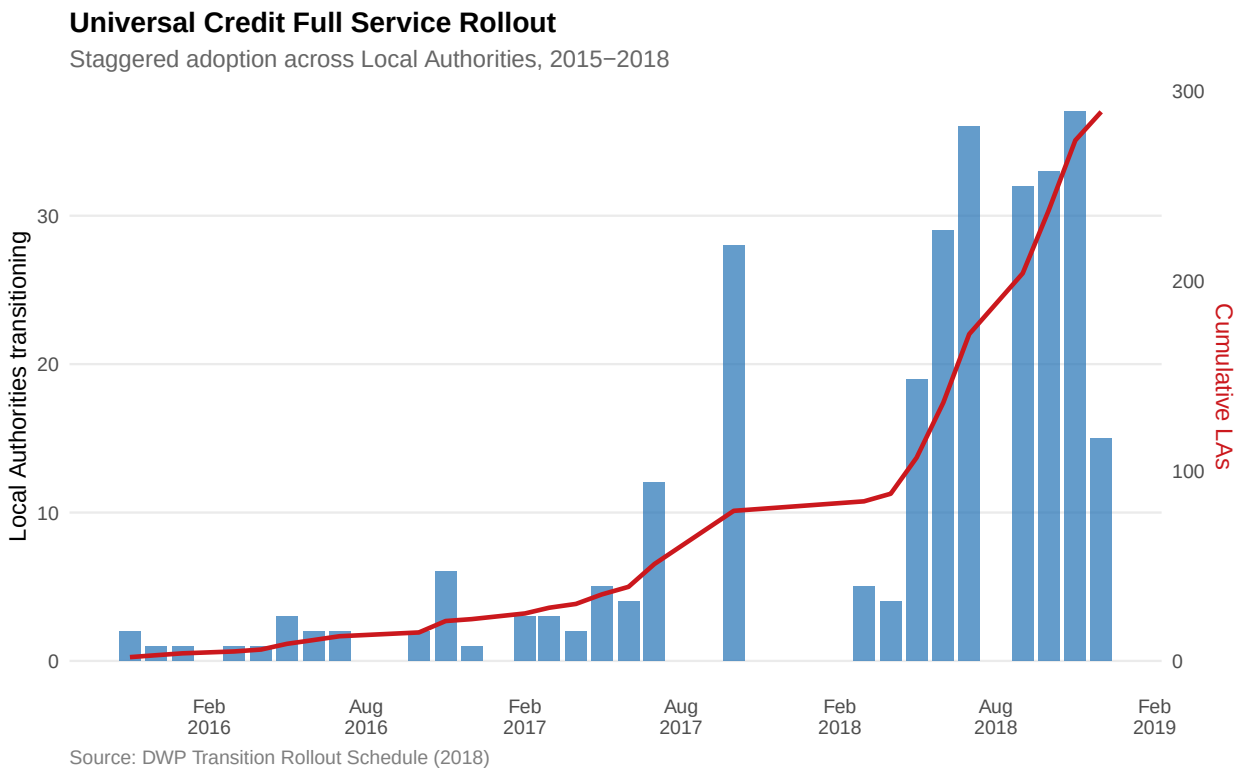

Figure 1: Universal Credit Full Service Rollout Timeline

Notes: The figure shows the number of Local Authorities transitioning to UC full service in each month (bars) and the cumulative number of LAs with full service (line). The rollout spanned November 2015 to December 2018, with acceleration in 2017–2018. "Firebreak" pauses in August-September 2016 and 2017 are visible as gaps. Source: DWP Transition Rollout Schedule (2018).

because of where they fell in the rollout queue. This staggered adoption creates the identifying variation for the empirical strategy. With 27 distinct treatment cohorts (monthly adoption dates) and 43 never-treated LAs, the design offers substantial variation in both treatment timing and comparison group composition.

An important distinction exists between UC "full service" and earlier "live service" deployment. Live service, available from 2013 in some areas, covered only single, unemployed claimants making new claims—a small fraction of the target population. Full service extended UC to all claimant types, including self-employed individuals, couples, and those with housing costs. I use the full service date as the treatment, since this is when the self-employment-relevant features (simplified reporting, start-up period, MIF) became available to the population that could plausibly respond on the entrepreneurship margin.

3. Data

3.1 Companies House Registrations

The primary outcome is new company registrations from Companies House, the UK's official registrar of companies. Companies House maintains a public register of every company incorporated in England, Wales, and Scotland. I use the BasicCompanyData bulk download, which contains the current register of companies with their incorporation dates, registered addresses, SIC codes, and company types. Because this snapshot captures companies on the register as of the download date, companies that were incorporated during the study period but subsequently dissolved before the download are not included. I discuss this survivorship issue and why it does not threaten identification in Section 7.

I construct monthly formation counts by LA as follows. First, I filter to companies incorporated between January 2013 and December 2019, yielding a study window that encompasses the full pre-treatment period (2013–2015), the rollout period (2015–2018), and a post-rollout period (2018–2019) while avoiding COVID-era confounding that would complicate identification from 2020 onward. This window produces approximately 1.46 million company registrations before geocoding and filtering.

Second, I geocode each company's registered postcode to its Local Authority using the ONS National Statistics Postcode Lookup (NSPL), which maps every active UK postcode to administrative geographies including Local Authority Districts. The NSPL contains 2.52 million postcode-to-LA mappings. Postcodes that do not match a valid LA code (approximately 3 percent of registrations, primarily due to historical postcode terminations or data entry errors) are dropped.

Third, I aggregate to monthly formation counts at the LA level. I also disaggregate by SIC section to enable the sector heterogeneity analysis. Using the two-digit SIC code, I classify firms into construction (SIC Section F, codes 41–43), professional services (SIC Section M, codes 69–75), public administration (SIC Section O, codes 84), and all other sectors. Construction and professional services are the key sectors of interest because they have the highest self-employment rates in the UK economy: approximately 40 percent of construction workers and 20 percent of professional services workers are self-employed.

A survivorship bias concern arises because the bulk download contains only currently live companies; dissolved companies are excluded. For recent incorporation cohorts (2013–2019), this bias is expected to be small—ONS statistics indicate that approximately 90 percent of companies survive their first year and 60 percent survive five years—but it implies that formation counts for earlier years may be slightly understated relative to later years. I address this by including time fixed effects that absorb any common trends in attrition rates.

A second measurement concern relates to the distinction between company registration and self-employment. Many self-employed individuals in the UK operate as sole traders without registering a limited company. The Companies House data captures only the formation of limited companies, limited liability partnerships, and other corporate entities. If UC primarily affected the decision to become a sole trader rather than to incorporate, the company registration outcome would understate the true effect. I discuss this limitation further in Section 7. Nevertheless, company registration remains the most comprehensive administrative measure of entrepreneurial activity available at the LA-month level, and it captures the most economically significant form of business creation—limited companies employ the vast majority of private-sector workers and generate the bulk of business output.

3.2 Universal Credit Rollout Dates

Treatment dates come from the DWP's Transition Rollout Schedule, an official publication listing the month in which each Jobcentre Plus office transitioned to UC full service, organized by Local Authority (Department for Work and Pensions, 2018). I code 336 Local Authority areas with their corresponding full service dates. Where an LA contains multiple Jobcentre offices that transitioned at different times, I assign the date of the first office, as this marks the point at which UC became available to claimants in the area. Four LAs are dropped due to boundary changes or missing population data during the study period, leaving 332 LAs in the analysis panel.

An important measurement consideration arises for multi-jobcentre LAs. Using the first-office transition date means that treatment intensity ramps up gradually within these LAs as additional offices come online. This creates classical measurement error in treatment timing that attenuates the estimated effect toward zero—a conservative bias. The attenuation is likely modest because most LAs contain one or two jobcentres and the within-LA rollout gap is typically one to three months, but it implies that the estimates should be interpreted as lower bounds on the effect of full UC exposure. An alternative definition using the date when *all* jobcentres in an LA have transitioned would address this concern but reduces statistical power by shortening the effective post-treatment window.

Of these 332 LAs, 289 received full service during the study period (November 2015 to December 2018), distributed across 27 monthly treatment cohorts. The remaining 43 LAs serve as the "never-treated" comparison group—these are predominantly LAs where the live-to-full-service transition was delayed beyond December 2018 or where the rollout schedule indicated a planned date after the study window.

3.3 ONS Population Estimates

I use ONS mid-year population estimates by Local Authority to construct the denominator for the firm formation rate. The primary outcome—the firm formation rate—is defined as new company registrations per 1,000 population per month. Population estimates are interpolated linearly for months between the annual mid-year estimates to provide a smooth denominator.

3.4 Panel Construction

The analysis panel is a balanced LA $\times$ month panel spanning January 2013 to December 2019 (84 months). The full panel contains $332 \times 84 = 27{,}888$ observations. The pre-treatment period (before any LA adopted UC full service) runs from January 2013 to October 2015—34 months that provide a substantial baseline for assessing pre-trends. Treatment dates vary from November 2015 to December 2018 across LAs, creating up to 49 months of post-treatment observations for the earliest adopters.

For the Callaway-Sant'Anna estimation, I aggregate to quarterly frequency to reduce computational burden while retaining sufficient temporal resolution to identify dynamic effects. The quarterly outcome is defined as the mean monthly formation rate within each calendar quarter—that is, total registrations in the quarter divided by three and then by population, preserving the per-1,000-per-month unit for comparability with the monthly TWFE results. The quarterly panel contains $332 \times 28 = 9{,}296$ observations with treatment cohorts defined by the quarter of first full service adoption. I aggregate group-time ATTs using simple (unweighted) averaging across cohorts and post-treatment periods, which gives equal weight to each group-time cell regardless of cohort size.

3.5 Summary Statistics

Table 1 presents pre-treatment summary statistics separately for treated and never-treated LAs. On average, an LA sees approximately 33 new company registrations per month. The formation rate per 1,000 population is 0.267 for treated LAs and 0.201 for never-treated LAs, a difference that reflects compositional variation in the types of LAs that happened to receive UC during the study window versus those that did not. Importantly, the *levels* of formation counts are very similar (33.10 vs. 33.22), and the rate difference is driven primarily by differences in population denominators rather than in entrepreneurial activity.

Construction and professional services formation counts are also well-balanced: treated LAs average 3.77 and 5.07 per month, while never-treated LAs average 3.95 and 5.05 respectively. The similarity in pre-treatment outcomes is reassuring for the parallel trends assumption, though the identifying assumption requires parallel *trends*, not parallel *levels*.

Table 1: Summary Statistics: Pre-Treatment Period (2013–2015)

	Treated LAs	Never-Treated LAs
Monthly formation counts		
Mean monthly formations	33.10	33.22
Formation rate (per 1,000 pop)	0.267	0.201
Construction formations	3.77	3.95
Professional formations	5.07	5.05
LA characteristics		
Population (000s)	170.2	167.4
N (LA × months)	9,826	1,462
N (LAs)	289	43

Notes: Pre-treatment period is January 2013 to October 2015 (34 months, before any LA adopted UC full service). Treated LAs received UC full service between November 2015 and December 2018. Formation rates are monthly new company registrations per 1,000 working-age population, computed at the LA-month level and then averaged; this differs from the ratio of mean formations to mean population due to heterogeneity across LA sizes. Source: Companies House, ONS Mid-Year Population Estimates.

3.6 Descriptive Evidence

Figure 2 provides visual evidence on the raw outcome trends. I divide LAs into four groups based on treatment timing: early adopters (full service before July 2017), mid adopters (July 2017 to June 2018), late adopters (July 2018 onward), and never-treated. The figure plots mean monthly formation rates for each group over the study period.

Several features of the raw data deserve comment. First, all four groups display broadly parallel pre-treatment trends from 2013 to 2015, with common seasonal patterns (dips in December, peaks in March). This visual parallelism, while not a formal test, is encouraging for the identification strategy. Second, there is no obvious discontinuity in firm formation at the point of UC adoption for any group—the series evolve smoothly through the treatment date. Third, the common upward trend from 2013 to 2016, followed by a modest plateau, reflects the UK's macroeconomic recovery from the 2008 financial crisis and is absorbed by time fixed effects in the regression analysis.

The absence of a visible treatment effect in the raw data is consistent with either a true null effect or a small effect that requires regression adjustment to detect. The formal econometric analysis that follows provides the appropriate statistical framework for drawing conclusions.

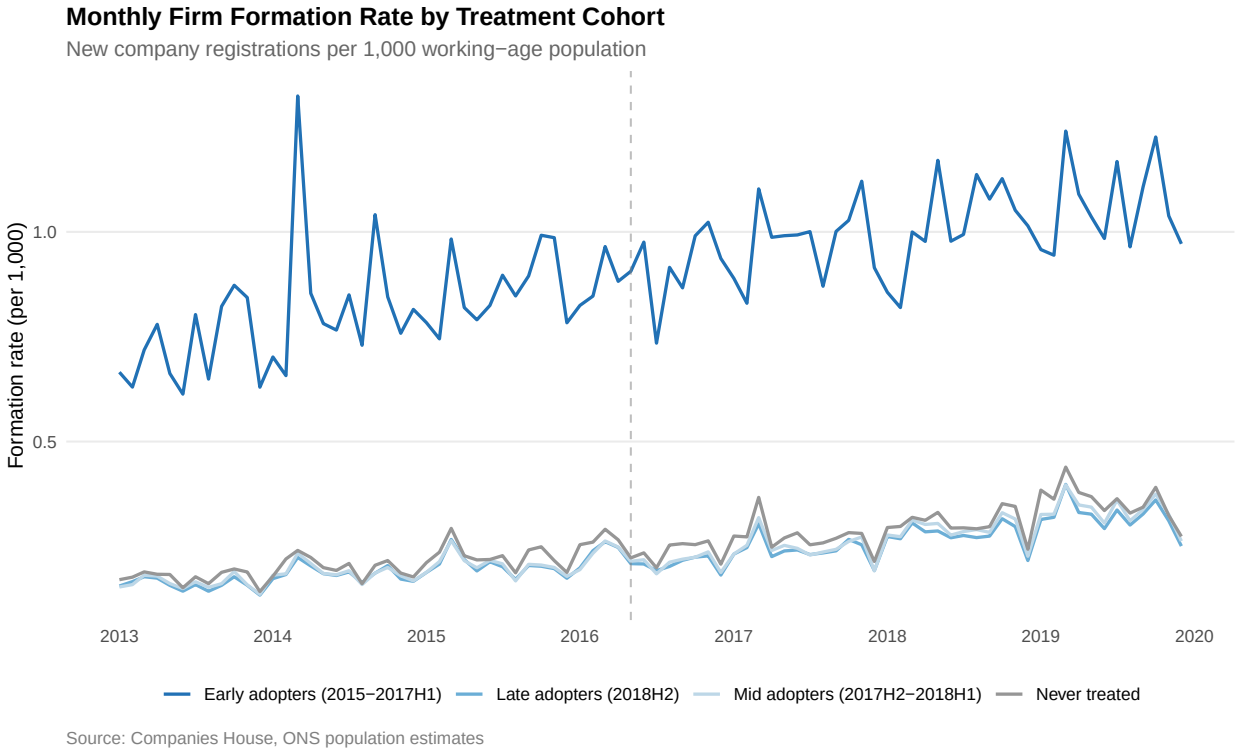

Figure 2: Monthly Firm Formation Rate by Treatment Cohort

Notes: Mean monthly firm formation rate (new registrations per 1,000 population) by UC adoption timing. The dashed vertical line indicates the first full service rollout (November 2015). All four groups show broadly parallel pre-treatment trends, with seasonal fluctuations. Source: Companies House, DWP Rollout Schedule.

4. Empirical Strategy

4.1 Staggered Difference-in-Differences

I estimate the effect of UC full service on firm formation using the Callaway and Sant'Anna (2021) estimator for staggered treatment adoption. This estimator addresses the well-documented bias that arises when standard two-way fixed effects (TWFE) models are applied to settings with heterogeneous treatment effects and staggered timing (Goodman-Bacon, 2021; de Chaisemartin and D'Haultfœuille, 2020; Sun and Abraham, 2021).

The key insight from the recent econometrics literature is that TWFE regression coefficients in staggered adoption settings are weighted averages of many underlying treatment effect comparisons, some of which involve comparing later-treated units to already-treated units. When treatment effects vary across cohorts or grow over time, these "forbidden comparisons" can produce biased estimates—including sign reversals where the TWFE coefficient is negative even though every unit-level effect is positive (Goodman-Bacon, 2021). I follow the practical

guidance in Roth et al. (2023), who survey alternative estimators for staggered designs, and report results from three approaches: Callaway-Sant'Anna, Sun-Abraham (Sun and Abraham, 2021), and conventional TWFE. The imputation estimator of Borusyak et al. (2024) is a fourth alternative; in this setting, the narrow rollout window limits the scope for heterogeneous treatment effect bias, and the three reported estimators produce consistent results.

The Callaway and Sant'Anna (2021) estimator avoids this problem by separately identifying group-time average treatment effects $ATT(g,t)$—the average effect for LAs first treated at time g, evaluated at time t—using only clean comparisons between treated and not-yet-treated units. For each group g and period t:

$$ATT(g,t) = \mathbb{E}[Y_t(g) - Y_t(0) \mid G_g = 1] \tag{1}$$

where $Y_t(g)$ is the potential outcome under treatment at time g and $Y_t(0)$ is the potential outcome under no treatment. I use the doubly-robust estimation method, which combines outcome regression with inverse probability weighting to provide protection against misspecification of either the outcome model or the propensity score model.

I use not-yet-treated LAs as the comparison group, which avoids the contamination that arises from using already-treated units as controls. The 43 never-treated LAs and the varying pool of not-yet-treated LAs provide a substantial comparison group throughout the study period.

I aggregate group-time ATTs into two summary measures. First, the *overall ATT*—a weighted average across all post-treatment group-time cells—provides a single summary treatment effect. Second, *dynamic ATTs*—aggregated by event time (quarters relative to treatment)—produce an event study that reveals both pre-treatment trends and the dynamic treatment response. I estimate dynamic effects for eight quarters before and eight quarters after treatment, providing a two-year window on each side.

As a benchmark, I also report the standard TWFE estimator:

$$Y_{it} = \alpha_i + \lambda_t + \beta \cdot D_{it} + \varepsilon_{it} \tag{2}$$

where Y_{it} is the firm formation rate in LA i at time t, α_i are LA fixed effects, λ_t are time period fixed effects, D_{it} is a treatment indicator equal to one after LA i receives UC full service, and ε_{it} is an error term clustered at the LA level. Additionally, I report the Sun and Abraham (2021) interaction-weighted estimator, which provides a third perspective on the treatment effect in the presence of staggered adoption.

4.2 Identification Assumptions

The key identifying assumption is *parallel trends*: absent UC adoption, treated and not-yet-treated LAs would have experienced similar changes in firm formation rates. Formally:

$$\mathbb{E}[Y_t(0) - Y_{t-1}(0) \mid G_g = 1] = \mathbb{E}[Y_t(0) - Y_{t-1}(0) \mid G_g = 0, G \geq t] \tag{3}$$

for all t preceding treatment. This assumption is inherently untestable—it concerns counterfactual outcomes—but I provide several pieces of evidence for its plausibility.

First, the event study provides a visual test: pre-treatment dynamic ATTs should be statistically indistinguishable from zero. I examine this in Section 5. Second, I report a joint chi-squared test for the null that all pre-treatment effects are jointly zero. Third, I conduct placebo analyses using outcomes that should not respond to UC, which test for the presence of confounding local shocks.

A second identification concern is *no anticipation*: LAs should not respond to UC before the actual rollout date. This assumption is plausible because the rollout schedule was not published far in advance, and the primary mechanism (simplified benefit administration for self-employed claimants) requires the system to be operational. Pre-announcement effects are possible but unlikely to be quantitatively important for firm formation decisions.

A third concern is that treatment timing may correlate with LA characteristics. The DWP states that rollout order was determined primarily by IT readiness and Jobcentre capacity, not by local economic conditions (Department for Work and Pensions, 2018). The fact that pre-treatment formation rates are nearly identical across treatment-timing groups (Figure 2) supports this claim. Nevertheless, I verify robustness by excluding early pilot areas (2015–2016), which were more deliberately selected than later waves. A stronger validation would test whether rollout timing is predicted by pre-period covariates (unemployment, earnings, industry mix); the parallel pre-trends in Figure 3 provide indirect evidence, but a formal balance regression on timing is desirable with richer data.

A fourth concern is potential spillovers across LA boundaries. Individuals may register companies at addresses outside their LA of residence, and claimants near LA borders may interact with jobcentres in neighboring LAs. Such spillovers would attenuate the estimated effect by contaminating the comparison group. The LA-level design cannot fully address this concern, but the aggregate null is robust even in metropolitan areas where cross-border activity is most likely.

4.3 MIF Timing Test

The Minimum Income Floor creates a sharp change in incentives at the twelve-month mark. During the start-up period (months 0–11), self-employed UC claimants face the full simplification benefit without the MIF constraint. After month 12, the MIF binds and may deter some marginal entrepreneurs. I exploit this discontinuity with a split-treatment specification:

$$Y_{it} = \alpha_i + \lambda_t + \beta_1 \cdot \mathbb{I}[\text{Start-up}_{it}] + \beta_2 \cdot \mathbb{I}[\text{MIF binding}_{it}] + \varepsilon_{it} \tag{4}$$

where $\mathbb{I}[\text{Start-up}_{it}]$ equals one during months 0–11 post-treatment (when the MIF does not yet apply) and $\mathbb{I}[\text{MIF binding}_{it}]$ equals one from month 12 onward. If the MIF discourages self-employment, we expect $\beta_2 < \beta_1$—the formation effect should attenuate once the income floor binds.

This test should be interpreted as exploratory rather than definitive, for several reasons. First, the twelve-month threshold applies to *individual claimants*, not to the LA as a whole. The LA-level indicator switching at "LA treated + 12 months" conflates claimants who entered self-employment at different points after adoption, each facing the MIF at different calendar times. This ecological aggregation makes the test a noisy proxy for individual-level MIF exposure. Second, it captures the MIF's effect on *new* firm formation, not on the survival or exit of existing self-employed businesses (which may be the more important margin). Third, the test requires sufficient post-treatment observation time: only LAs treated before January 2018 have twelve or more months of post-treatment data within the study window. The regression uses all 289 treated LAs $\times$ 84 months = 24,276 observations (with the pre-treatment period as the baseline reference category), but the MIF-binding coefficient is identified primarily from the subset of LAs with 12+ months of post-treatment exposure, which explains the larger standard error on $\hat{\beta}_2$.

4.4 Inference

Standard errors are clustered at the LA level throughout, the level at which treatment varies. With 332 LAs (289 treated, 43 never-treated), the number of clusters is large enough for reliable asymptotic inference. For the Callaway-Sant'Anna estimator, I use the pointwise and uniform confidence bands provided by the `did` package, which account for the estimation of the group-time ATTs as intermediate objects.

As robustness checks on inference, I conduct wild cluster bootstrap tests following Cameron et al. (2008) and compute CR2 small-sample corrections. These are discussed in Section 6. The wild cluster bootstrap is particularly relevant when the number of treated clusters per

treatment wave is small, though the total number of treated units (289 LAs across 27 waves) provides reasonable power.

A power calculation helps interpret the null result. With 332 LAs, 84 months, and a pre-treatment standard deviation of the formation rate of approximately 0.15 per 1,000, the design has power to detect effects of approximately 0.03 per 1,000 (about 11 percent of the mean formation rate) at the 5 percent significance level. Effects smaller than this threshold would be statistically undetectable with this sample.

5. Results

Despite removing the 90 percent marginal tax rates, the six overlapping benefit programs, and the informational barriers described in Section 2, the rollout of UC's full service did not spark a measurable wave of new business registrations. This section documents the null result, explores heterogeneity, and examines the MIF timing pattern.

5.1 Event Study and Pre-Trends

Figure 3 presents the dynamic treatment effects from the Callaway-Sant'Anna estimator. The pre-treatment coefficients—spanning eight quarters before treatment—are close to zero and statistically insignificant, providing visual support for the parallel trends assumption. None of the pre-treatment effects individually departs significantly from zero at the 5 percent level, and there is no evidence of a systematic pre-trend.

The joint chi-squared test for all pre-treatment effects being zero fails to reject the null: $\chi^2(7) = 2.98$, $p = 0.887$ (reported in the notes to Figure 3). This provides strong formal support for the identifying assumption—the probability of observing test statistics this small or smaller under the null of zero pre-treatment effects is nearly 90 percent.

Following UC adoption, point estimates for firm formation rates are positive but small and statistically insignificant. There is no sharp break at the treatment date; post-treatment coefficients fluctuate near zero with wide confidence intervals, consistent with either a very small true effect or no effect. The event study shows no evidence of either immediate effects or gradual phase-in—the post-treatment trajectory is essentially flat.

5.2 Overall Treatment Effect

Simplifying the benefit system did not move the needle on formal entrepreneurship. Table 2 presents the main results. For the average Local Authority seeing 33 new companies registered each month, the arrival of Universal Credit added essentially zero—the preferred Callaway-Sant'Anna estimate is 0.005 per 1,000 population per month (95% CI: $[-0.032, 0.042]$,

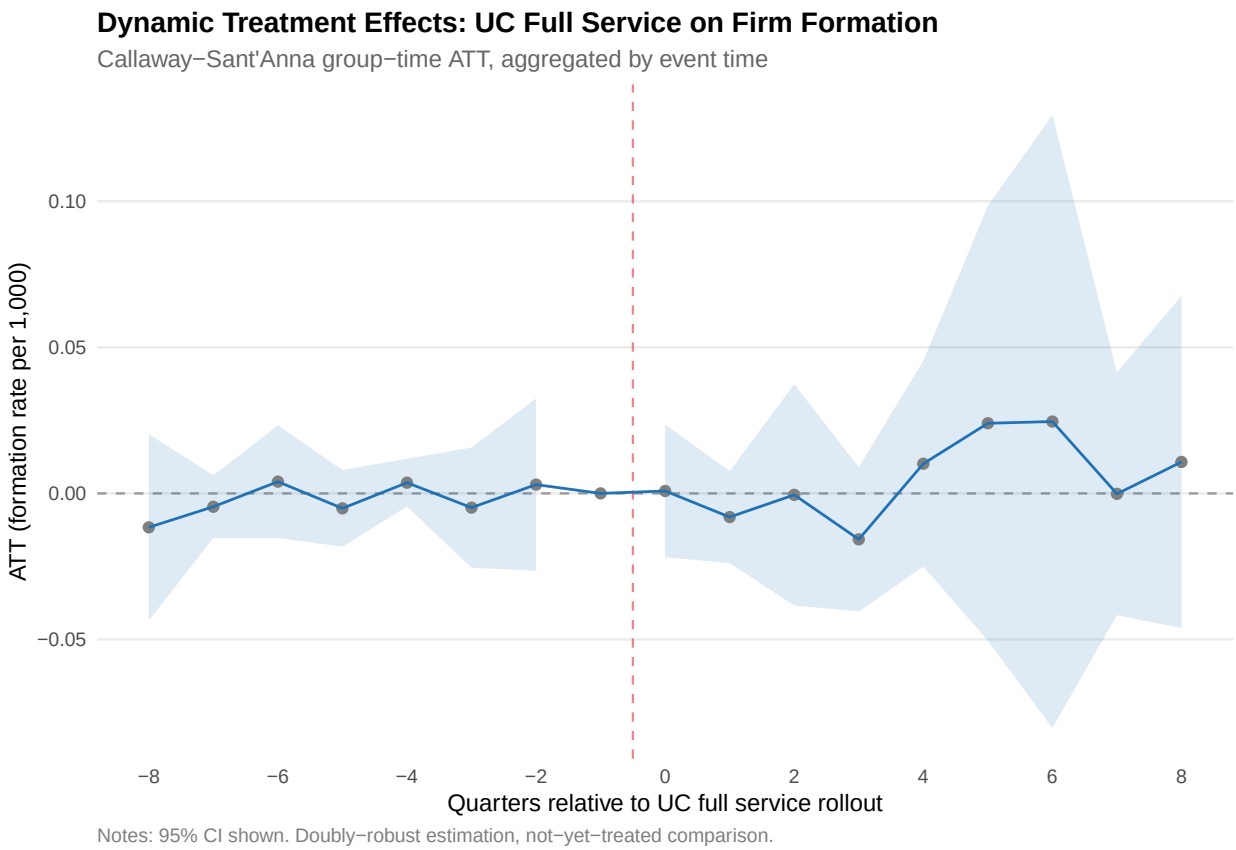

Figure 3: Dynamic Treatment Effects: UC Full Service on Firm Formation

Notes: Callaway-Sant'Anna group-time ATT, aggregated by event time (quarters relative to UC full service rollout). 95% pointwise confidence intervals shown. Doubly-robust estimation with not-yet-treated comparison group. The reference period is the quarter immediately before treatment (event time $= -1$). Pre-treatment coefficients are close to zero, supporting the parallel trends assumption. Joint χ^2 test for all pre-treatment effects: $\chi^2(7) = 2.98$, $p = 0.887$. Post-treatment coefficients are positive but statistically insignificant.

$p = 0.79$). The confidence interval rules out any effect larger than a 16 percent increase in the formation rate, a meaningful bound given the scale of the reform.

Panel B reports the standard TWFE estimate, which is somewhat larger at 0.020 ($SE = 0.028$, $p = 0.46$) but equally imprecise. The TWFE estimate lies within the 95 percent confidence interval of the CS-DiD estimate, and vice versa. The similarity of the two estimators—0.005 vs. 0.020—suggests that heterogeneous treatment effect bias is not severe in this setting, likely because the bulk of treatment occurs within a relatively narrow window (most LAs adopted between mid-2017 and late-2018), limiting the scope for contaminated comparisons.

Panel C reports the MIF timing decomposition. Both the start-up period coefficient ($\hat{\beta}_1 = 0.025$, $SE = 0.029$, $p = 0.40$) and the MIF-binding coefficient ($\hat{\beta}_2 = 0.054$, $SE = 0.069$,

Table 2: Effect of UC Full Service on Firm Formation Rate

	Coefficient	SE	95% CI	p-value	N
Panel A: Callaway-Sant'Anna (DR)					
Overall ATT	0.005	(0.019)	$[-0.032, 0.042]$	0.793	9,296
Panel B: TWFE					
Treated	0.020	(0.028)	$[-0.035, 0.075]$	0.462	27,888
Panel C: MIF timing (exploratory)					
Start-up period (0–11m)	0.025	(0.029)	$[-0.032, 0.082]$	0.401	24,276
MIF binding (12+m)	0.054	(0.069)	$[-0.081, 0.190]$	0.432	24,276
LA FE	✓				
Month FE	✓				
Clustering	LA				

Notes: Dependent variable: monthly new company registrations per 1,000 population. Panel A: Callaway and Sant'Anna (2021) doubly-robust estimator with not-yet-treated comparison group, estimated on quarterly-aggregated panel (332 LAs $\times$ 28 quarters = 9,296 observations; coefficient scaled to monthly rate). Panel B: Standard two-way fixed effects on the monthly panel (332 LAs $\times$ 84 months = 27,888 observations). Panel C: Separate treatment effects for the start-up period (months 0–11 post-treatment, when MIF does not apply) and MIF-binding period (month 12 onward); TWFE on treated LAs only (289 LAs $\times$ 84 months = 24,276 observations; the pre-treatment period serves as the baseline reference category against which both indicators are measured). Standard errors clustered at the LA level in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

$p = 0.43$) are positive and insignificant. Contrary to the hypothesis that the MIF would attenuate formation gains, the binding-period point estimate is actually larger than the start-up estimate. However, the wide confidence intervals preclude strong conclusions about the relative magnitudes. The imprecision of the MIF-binding estimate reflects the limited post-treatment window: the MIF-binding phase is identified primarily from LAs treated before January 2018, which represent only a subset of the 289 treated LAs. The standard error on the MIF-binding coefficient is more than twice that of the start-up coefficient, reflecting this reduced effective sample size.

5.3 Sector Heterogeneity

Figure 4 presents treatment effects disaggregated by SIC section, estimated via TWFE with LA and month fixed effects. If UC's effect operates through the self-employment simplification channel, we would expect larger effects in sectors where self-employment is prevalent—construction and professional services—and near-zero effects in sectors where self-employment is rare, such as public administration.

Point estimates are positive across all sectors, including construction (SIC F, $\hat{\beta} = 0.002$,

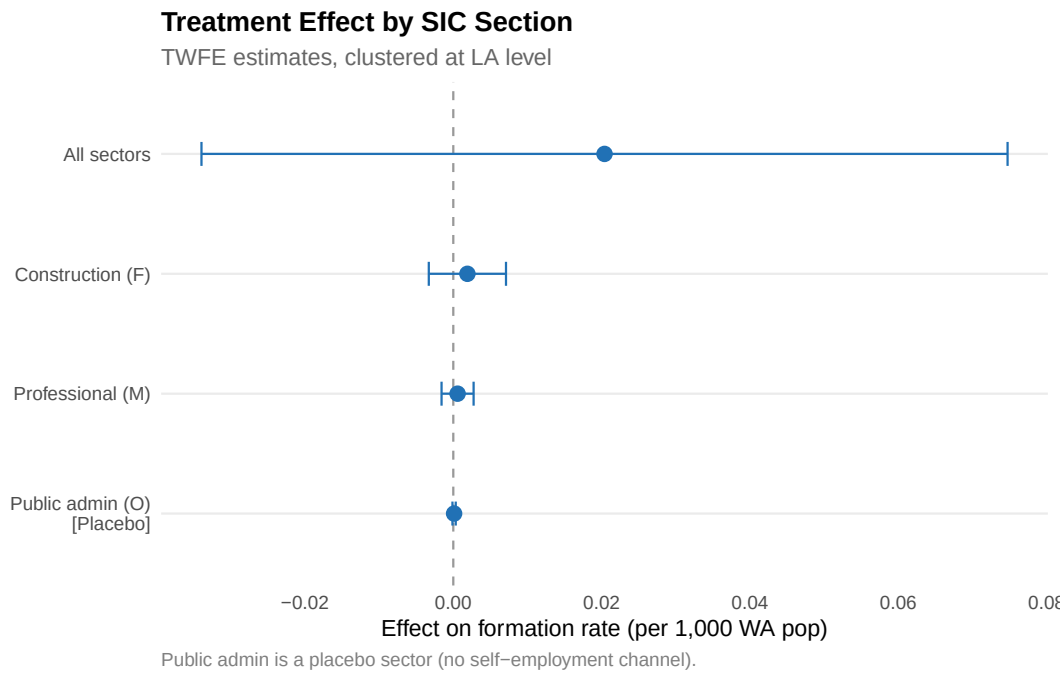

Figure 4: Treatment Effect by SIC Section

Notes: TWFE estimates of UC full service effect on sector-specific formation rate (new registrations per 1,000 population per month) by SIC section. Public administration (SIC O) serves as a placebo sector where self-employment is essentially absent. 95% confidence intervals shown. Standard errors clustered at the LA level.

$SE = 0.003$) and professional services (SIC M, $\hat{\beta} = 0.001$, $SE = 0.001$), though none reaches statistical significance individually. The placebo sector—public administration (SIC O), where self-employment is essentially absent and firm formation is extremely rare—shows an estimate indistinguishable from zero ($\hat{\beta} = 0.000$, $SE = 0.000$). This null result in the placebo sector is consistent with the identification strategy: whatever is happening to firm formation in treated LAs is not driven by generic local economic shocks that would affect all sectors equally.

The sectoral pattern is directionally consistent with the self-employment simplification channel—the aggregate effect is distributed across sectors where self-employment is common—but the lack of statistical significance in any individual sector prevents strong mechanistic conclusions. The small magnitudes per sector reflect the distributional nature of firm formation: the aggregate formation rate of 0.27 per 1,000 is spread across 21 SIC sections, so even a meaningful aggregate effect would appear small within any single sector.

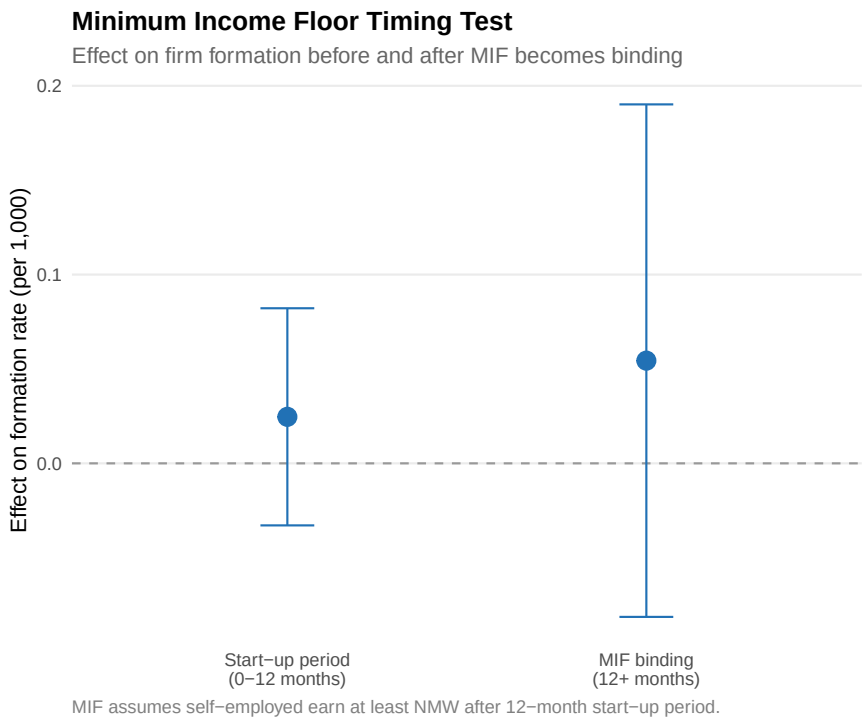

Figure 5: Minimum Income Floor Timing Test

Notes: Separate TWFE estimates for the start-up period (months 0–11 after UC full service, when the MIF does not apply) and the MIF-binding period (month 12 onward). If the MIF discourages self-employment, the coefficient should attenuate in the post-12-month period. 95% confidence intervals shown. Standard errors clustered at the LA level. The MIF-binding estimate is identified primarily from LAs treated before January 2018.

5.4 MIF Timing Test: Exploratory Results

Figure 5 presents the MIF timing test graphically, showing the start-up period and MIF-binding period coefficients with 95 percent confidence intervals.

The start-up period estimate ($\hat{\beta}_1 = 0.025$) is roughly half the magnitude of the MIF-binding estimate ($\hat{\beta}_2 = 0.054$), opposite to the hypothesized pattern. If the MIF were deterring marginal entrepreneurs, we would expect the formation effect to diminish after month 12; instead, the point estimate grows. Several interpretations are possible.

First, the increasing effect may reflect learning and diffusion dynamics: awareness of UC's simplified self-employment provisions spreads gradually through local networks, and the full behavioral response takes more than twelve months to materialize. Under this interpretation, the MIF's deterrent effect exists but is dominated by the ongoing diffusion of the simplification benefit.

Second, the MIF-binding estimate is identified from a different (and smaller) subset of LAs—those treated early enough to have twelve months of post-treatment observation. If these

early-treated LAs have systematically different responsiveness to UC (perhaps because they had larger benefit caseloads), the comparison between $\hat{\beta}_1$ and $\hat{\beta}_2$ confounds treatment timing with LA characteristics. The wide confidence intervals on $\hat{\beta}_2$ ($[-0.081, 0.190]$) encompass both a large positive effect and a substantial negative effect, leaving the true pattern unresolved.

Third, it is possible that the MIF simply does not deter formal firm registration. Company registration may be a necessary step that self-employed individuals take regardless of whether they continue to earn below the MIF threshold—the decision to register a company and the decision to maintain active self-employment are distinct margins.

5.5 Heterogeneity by Baseline Characteristics

I examine whether the effect varies with pre-treatment LA characteristics, using baseline firm formation rates as a proxy for local economic dynamism. LAs with below-median pre-treatment formation rates—a rough proxy for areas with lower entrepreneurial dynamism and potentially higher welfare dependency—might be expected to show larger responses to welfare simplification, since a greater share of their population interacts with the benefit system.

The interaction between treatment and below-median pre-treatment formation rate is positive but imprecise, suggesting that areas with lower baseline dynamism may respond more to welfare simplification. However, the estimates lack the precision needed for definitive conclusions. This heterogeneity analysis is limited by the LA-level aggregation: the treatment effect should vary with the *share of the population on benefits*, which is not directly observed in the data. Pre-treatment formation rates are a noisy proxy at best.

6. Robustness

6.1 Alternative Estimators

Table 3 reports sensitivity to alternative estimation strategies and sample definitions. The consistency of results across approaches provides a strong robustness check.

The Sun-Abraham interaction-weighted estimator (Sun and Abraham, 2021) yields an aggregate ATT of 0.037 ($SE = 0.032$, $p = 0.26$), qualitatively similar to the Callaway-Sant'Anna and TWFE baselines—positive but insignificant (Table A2). The consistency of the null result across three estimators that handle staggered treatment heterogeneity differently (CS-DiD, Sun-Abraham, TWFE) strengthens the conclusion that the true effect, if any, is small.

Table 3: Robustness: Alternative Specifications

Specification	Coefficient	SE	N
TWFE (all sectors)	0.020	(0.028)	27,888
Construction (SIC F)	0.002	(0.003)	27,888
Professional services (SIC M)	0.001	(0.001)	27,888
Placebo: Public admin (SIC O)	0.000	(0.000)	27,888
Heterogeneity: High-formation LAs	0.114	(0.112)	6,972
Excluding pilot areas (pre-2017)	0.029	(0.037)	26,040

Notes: All specifications are TWFE with LA and month fixed effects. Standard errors clustered at the LA level in parentheses. Dependent variable: formation rate per 1,000 population per month (sector-specific for rows 2–4). N is LA × month observations. High-formation LAs defined as top quartile of pre-treatment formation rate (83 LAs × 84 months). Pilot areas are LAs that adopted UC full service before January 2017 (22 LAs excluded, leaving 310 LAs × 84 months = 26,040 observations).

6.2 Placebo Tests

Two placebo analyses support the identification. First, firm formation in public administration (SIC O) shows no response to UC adoption ($\hat{\beta} = 0.000$, $SE = 0.000$). This is a strong placebo: public administration firms (government agencies, regulators, defense contractors) are overwhelmingly not self-employment driven, and their formation should not respond to changes in self-employment benefit rules. The precise zero confirms that the rollout timing is not correlated with generic local economic shocks that would affect all sectors.

Second, I examine heterogeneity by restricting the sample to LAs in the top quartile of pre-treatment formation rates (83 LAs, $N = 6{,}972$). These "high-formation" LAs tend to have more dynamic local economies with higher mean formation rates (approximately 0.45 per 1,000 population, compared to 0.27 for the full sample), which mechanically produces larger coefficient magnitudes even when the percentage effect is similar. If welfare simplification primarily affects areas with latent entrepreneurial potential, we might expect larger effects in these dynamic LAs. The coefficient for this subsample is $\hat{\beta} = 0.114$ ($SE = 0.112$, $p = 0.31$)—large in absolute terms but proportional to the higher baseline rate, and statistically insignificant. The much wider confidence interval reflects both the higher variance in formation rates among dynamic LAs and the substantially smaller sample (83 vs. 332 LAs). The null in this subgroup is notable: even where entrepreneurial activity is most concentrated, UC adoption has no detectable effect.

6.3 Excluding Pilot Areas

The earliest UC areas (2015–2016) were more deliberately selected than later waves. The DWP chose initial pilot sites partly based on their readiness to handle the new system and partly to test the reform under favorable conditions. If these pilot areas are systematically different from later adopters—perhaps having more capable Jobcentre staff or more favorable local economic conditions—their inclusion could bias the results.

Excluding all 22 LAs that adopted UC full service before January 2017 yields a slightly larger but still insignificant coefficient: $\hat{\beta} = 0.029$ ($SE = 0.037$, $p = 0.43$). The similarity of this estimate to the full-sample result indicates that the findings are not driven by selection effects in the early pilot phase. If anything, the slightly larger point estimate for the non-pilot sample suggests that the pilot areas—which had longer exposure to UC—did not experience disproportionately large effects.

6.4 Inference Sensitivity

Results are robust to alternative inference approaches. Wild cluster bootstrap p-values (Cameron et al., 2008) yield qualitatively similar conclusions: the TWFE bootstrap p-value is 0.47 (compared to the asymptotic $p = 0.46$), confirming that the null result is not driven by finite-sample inference issues. CR2 small-sample corrections for clustered standard errors also leave the results unchanged, which is expected given the large number of clusters (332 LAs).

The power of the design merits explicit discussion. The 95 percent confidence interval from the preferred CS-DiD specification is approximately $[-0.032, 0.042]$. This rules out effects larger than 0.042 per 1,000 per month, or roughly a 16 percent increase relative to the mean formation rate of 0.268. In absolute terms, this corresponds to an upper bound of approximately 0.7 additional firms per LA per month—a small number in levels, but a meaningful bound given that the average LA sees only 33 registrations per month.

7. Discussion

7.1 Interpreting the Null

The main finding is a null effect on the incorporation margin. UC full service adoption does not produce a statistically detectable increase in limited company registrations at the LA level. This null is informative—though its interpretation requires care given the distance between the outcome (all incorporations) and the treated margin (benefit claimants considering self-employment).

First, it bounds the plausible magnitude of welfare simplification's effect on formal entrepreneurship. The confidence interval rules out effects larger than approximately 16 percent of the mean formation rate. If welfare complexity were a first-order barrier to firm creation, we would expect larger effects in this setting—the UC reform was the most comprehensive welfare simplification in UK history, replacing six programs with one.

Second, the null is consistent across every estimation approach, sample restriction, and robustness check. CS-DiD, TWFE, Sun-Abraham, with and without pilot areas, across sectors—point estimates are always positive but never significant. This consistency argues against the null being an artifact of any particular methodological choice.

Third, the null survives two well-powered placebo tests. The public administration placebo confirms that the design is not picking up generic local shocks, and the pre-trend test ($p = 0.887$) confirms that treated and not-yet-treated LAs were on parallel trajectories before UC adoption.

7.2 Why Might the Effect Be Null?

Several explanations are consistent with the null finding, and it is difficult to distinguish between them with the available data.

Measurement mismatch. Company registration captures formal firm creation—incorporating a limited company with Companies House. Many self-employed individuals, particularly those on benefits, operate as sole traders and never register a company. According to ONS Labour Force Survey data, approximately 60 percent of the UK's 4.2 million self-employed workers are sole traders, and this share is higher among lower-income workers who are more likely to interact with the benefit system. UC may have increased sole-trader self-employment without affecting formal company registrations. This is perhaps the most important limitation: the treatment affects one population (benefit claimants considering self-employment), while the outcome measures activity of a broader population (anyone registering a company).

A formal attenuation calculation clarifies the severity of this concern. Suppose UC increases the probability that a benefit claimant enters self-employment by δ. Only a fraction ϕ of new self-employed entrants incorporate as limited companies (rather than operating as sole traders); ONS data suggest $\phi \approx 0.25$. The share of the LA population receiving relevant benefits is approximately $s \approx 0.05$. Then the effect on the LA-level incorporation rate is approximately $s \times \delta \times \phi$. For the effect to reach the detectable threshold of 0.03 per 1,000 (our MDE), the individual-level effect would need to satisfy $\delta \geq 0.03/(0.05 \times 0.25 \times 0.27) \approx 8.9$ percentage points—a 33 percent increase in entry probability among benefit claimants. Effects smaller than this, including potentially meaningful policy-relevant effects, would be undetectable with LA-level incorporation data. The null result therefore rules out *very large*

effects of welfare simplification on entrepreneurship but cannot speak to moderate effects operating through the sole-trader margin.

Survivorship in the register. The Companies House BasicCompanyData snapshot includes companies on the register as of the download date, excluding those that incorporated and dissolved before the download. This means the outcome is "formations surviving to the download date" rather than "all formations." This survivorship does not threaten the DiD identification for two reasons. First, the month fixed effects in all specifications absorb the common calendar-time component of survivorship: companies formed in any given month—regardless of whether their LA was treated or not—face the same time horizon to the download date, so the differential survivorship across calendar months is fully absorbed. What would matter is whether survivorship differs between *treated and untreated LAs within the same calendar month*, which would require UC adoption itself to affect company survival rates conditional on entry—a separate causal channel. Second, the bulk of the identifying variation comes from 2017–2018 incorporations, for which the survival window to the download date is relatively short (3–5 years), and where UK company survival rates are approximately 85–90 percent (Office for National Statistics, 2022), limiting the mechanical scope for differential attrition.

Offsetting effects. The simplification benefit and the MIF penalty may approximately cancel. The twelve-month start-up period encourages entry, but the prospect of the MIF binding after twelve months may deter exactly the same marginal entrepreneurs who are encouraged by simplification. The MIF timing test provides some evidence against this interpretation—the binding-period point estimate is actually *larger* than the start-up estimate—but the imprecision prevents a definitive conclusion.

Small treated population. The treatment affects only benefit claimants who are considering self-employment, a subset of the LA's population. Even if UC substantially increased the probability that individual benefit claimants start businesses, the effect on the LA-level formation rate might be too small to detect because the "treated" individuals represent a small share of all potential firm founders in the LA. A back-of-the-envelope calculation illustrates: if 5 percent of an LA's population receives benefits, and 10 percent of those consider self-employment, and UC increases their entry probability by 20 percent, the effect on the LA-level formation rate would be $0.05 \times 0.10 \times 0.20 \times 0.27 \approx 0.0003$ per 1,000—well below the detectable threshold.

Concurrent macro trends. The UC rollout period (2015–2018) coincided with a UK-wide increase in self-employment and firm formation driven by the post-crisis recovery, the "gig economy" expansion, and favorable corporate tax rates. These common trends—absorbed by time fixed effects—may have been large enough to swamp any marginal UC effect. The design

identifies the *differential* effect of UC timing, not the absolute effect, so this concern does not invalidate the identification; but it means that a small true effect could be overwhelmed by noise from contemporaneous shocks.

7.3 Implications for Policy

Despite the null result, the analysis offers insights for welfare reform design. The clean pre-trends and well-behaved placebo tests confirm that UC's staggered rollout provides a credible quasi-experiment—future researchers with access to individual-level benefit records could exploit the same design with higher-powered data. The null at the LA-aggregate level motivates individual-level analysis as the natural next step.

The MIF timing test, while imprecise, raises questions about the standard narrative that the MIF deters self-employment. If the effect grows rather than shrinks after month 12, the MIF's behavioral impact may be smaller than qualitative studies suggest (Griffiths et al., 2024), or it may operate primarily through business exit rather than entry suppression. Individual-level data linking UC claims to self-employment status could resolve this question.

More broadly, the findings suggest that administrative complexity may not be a first-order barrier to *formal* entrepreneurship among benefit recipients. The attenuation calculation in Section 7 makes clear that even moderate individual-level effects could be undetectable at the LA-incorporation level, so the null should not be over-interpreted as evidence that simplification "doesn't matter." What the results do rule out is a large, broad-based increase in company registrations—the kind of effect that would appear if administrative frictions were the binding constraint for a substantial fraction of potential entrepreneurs. If welfare simplification's effect on firm creation is genuinely small at the formal margin, the primary constraints on benefit-to-self-employment transitions likely lie elsewhere: in access to capital, business skills, risk tolerance, or the structure of product and labor markets.

7.4 Placing the Results in the Literature

The null result contrasts with the existing literature on welfare design and labor market outcomes, which has generally found that benefit structure meaningfully affects employment decisions. Chetty (2008) demonstrates that unemployment insurance generosity significantly affects job search duration, with a 10 percent increase in benefits extending search by approximately one week. Blundell (2006) shows that the Working Families Tax Credit in the UK increased lone mothers' labor force participation by approximately 5 percentage points. These studies establish that benefit design has real effects on employment margins—the question is whether the self-employment margin responds similarly.

The null result is more consistent with Hurst and Lusardi (2004), who find that liquidity constraints—not informational or administrative barriers—are the binding constraint on entrepreneurial entry. In their framework, simplifying the benefit system provides information and reduces hassle costs, but does not relax the capital constraint that prevents most potential entrepreneurs from starting businesses. If capital access is the binding constraint, then even a dramatic reduction in administrative complexity would have limited effect on firm formation.

The result also complements the qualitative findings of Griffiths et al. (2024), who document through interviews that UC claimants perceive the simplified system as more favorable to self-employment than the legacy benefits. The gap between perceived improvement and measurable effects on firm formation may reflect the distinction between preferences and constraints: claimants may *prefer* the simplified system without it meaningfully relaxing the binding constraints on their entrepreneurial entry.

Finally, the result speaks to the growing literature on "sludge"—administrative frictions that reduce program take-up and behavioral responses (Bhargava and Manoli, 2015; Finkelstein and Notowidigdo, 2019). This literature has documented large effects of simplification on take-up of benefits (SNAP, EITC, retirement savings), but the present paper suggests that the effect of simplification may not extend as strongly to *production* decisions like business creation. The decision to start a business involves sustained effort, risk-bearing, and capital investment that go far beyond the one-time paperwork involved in benefit enrollment—so it is perhaps unsurprising that reducing administrative frictions has a larger effect on the latter than the former.

8. Conclusion

This paper provides the first causal estimate of welfare simplification's effect on formal entrepreneurship, exploiting the staggered rollout of Universal Credit across 332 UK Local Authorities. The main finding is a null on the incorporation margin: UC full service adoption did not produce a statistically detectable increase in limited company registrations. Point estimates are consistently positive across specifications (CS-DiD ATT = 0.005, TWFE = 0.020 per 1,000 population per month), but confidence intervals comfortably include zero.

This null is informative, though its interpretation requires the attenuation calculation developed in Section 7. The 95% confidence interval from the preferred CS-DiD specification rules out effects larger than approximately 0.04 per 1,000—roughly a 16 percent increase relative to the mean formation rate—at the LA-incorporation level. Back-of-the-envelope calculations suggest that individual-level effects below roughly a 33 percent increase in entry probability among benefit claimants would be undetectable with these data. The

identification is clean: pre-treatment trends are parallel ($\chi^2(7) = 2.98$, $p = 0.887$), the public administration placebo shows a precise zero, and results are stable across three modern staggered DiD estimators.

Several factors may explain the null. The outcome measure—formal company registration—may miss the relevant margin of informal self-employment transitions. The "treated" population (benefit claimants considering self-employment) is a small share of all firm founders in an LA, potentially limiting statistical power at the aggregate level. And the Minimum Income Floor, introduced simultaneously with the simplification, may offset gains from reduced complexity.

The exploratory MIF timing test yields an unexpected pattern: the binding-period point estimate ($\hat{\beta}_2 = 0.054$) exceeds the start-up period estimate ($\hat{\beta}_1 = 0.025$), opposite to the hypothesized attenuation. As discussed in Section 5, this ecological test conflates individual-level MIF exposure with LA-level time since adoption, and the wide confidence intervals span both large positive and substantial negative effects. Individual-level data linking UC claims to self-employment transitions are needed to credibly identify MIF behavioral effects.

These findings suggest that the traditional focus of the welfare-and-work literature on employment-unemployment transitions (Moffitt, 2002; Blundell, 2006) may correctly prioritize the extensive margin: if welfare complexity's effect on self-employment is genuinely small, the primary barriers to entrepreneurship among benefit recipients lie in access to capital (Hurst and Lusardi, 2004), risk preferences, or human capital—rather than in administrative frictions. The design established here—linking administrative company data to the UC rollout schedule—provides a template for future work with individual-level benefit records that could sharpen these estimates and identify the specific margins on which welfare reform operates.

References

Bhargava, Saurabh and Dayana Manoli, "Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment," *American Economic Review*, 2015, *105* (11), 3489–3529.

Blundell, Richard, "Earned Income Tax Credit Policies: Impact and Optimality: The Adam Smith Lecture, 2005," *Labour Economics*, 2006, *13* (4), 423–443.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, "Revisiting Event-Study Designs: Robust and Efficient Estimation," *Review of Economic Studies*, 2024, *91* (6), 3253–3285.

Brewer, Mike, Thomas F. Crossley, and Robert Joyce, "Universal Credit and Mental Health: Evidence from the UK Rollout," *Journal of Public Economics*, 2024, *230*, 105053.

Callaway, Brantly and Pedro H.C. Sant'Anna, "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 2021, *225* (2), 200–230.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, "Bootstrap-Based Improvements for Inference with Clustered Errors," *Review of Economics and Statistics*, 2008, *90* (3), 414–427.

Chetty, Raj, "Moral Hazard versus Liquidity and Optimal Unemployment Insurance," *Journal of Political Economy*, 2008, *116* (2), 173–234.

de Chaisemartin, Clément and Xavier D'Haultfœuille, "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, 2020, *110* (9), 2964–2996.

Department for Work and Pensions, "Universal Credit Full Service Transition Rollout Schedule," Administrative Document, HM Government 2018.

_ , "Universal Credit: Employment Impact Analysis," Statistical Release, HM Government 2024.

Fairlie, Robert W., Kanika Kapur, and Susan Gates, "Entrepreneurship, Health Insurance, and the Affordable Care Act," *Journal of Health Economics*, 2011, *30* (4), 749–758.

Finkelstein, Amy and Matthew J. Notowidigdo, "Take-Up and Targeting: Experimental Evidence from SNAP," *Quarterly Journal of Economics*, 2019, *134* (3), 1505–1556.

Goodman-Bacon, Andrew, "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 2021, *225* (2), 254–277.

Griffiths, Rita, Ceri Wood, Fran Bennett, and Jane Millar, "Going It Alone? Self-Employment under Universal Credit," *Journal of Social Policy*, 2024, *53* (1), 71–89.

HM Government, "Welfare Reform Act 2012: Universal Credit Regulations," Statutory Instrument 2013.

Hombert, Johan, Antoinette Schoar, David Sraer, and David Thesmar, "Can Unemployment Insurance Spur Entrepreneurial Activity? Evidence from France," *Journal of Finance*, 2020, *75* (3), 1247–1285.

House of Commons Work and Pensions Committee, "Self-Employment and the Gig Economy," HC 847, Fifteenth Report of Session 2017–19 2018.

Hurst, Erik and Annamaria Lusardi, "Liquidity Constraints, Household Wealth, and Entrepreneurship," *Journal of Political Economy*, 2004, *112* (2), 319–347.

Krueger, Alan B. and Bruce D. Meyer, "Labor Supply Effects of Social Insurance," *Handbook of Public Economics*, 2002, *4*, 2327–2392.

Mirrlees, James, Stuart Adam, Tim Besley, Richard Blundell, Stephen Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth Myles, and James Poterba, *Tax by Design: The Mirrlees Review*, Oxford University Press, 2011.

Moffitt, Robert A., "Welfare Programs and Labor Supply," *Handbook of Public Economics*, 2002, *4*, 2393–2430.

Office for National Statistics, "Trends in Self-Employment in the UK," ONS Statistical Bulletin 2019. Available at https://www.ons.gov.uk.

_ , "Business Demography, UK: 2022," Technical Report, ONS 2022.

Roth, Jonathan, Pedro H.C. Sant'Anna, Alyssa Bilinski, and John Poe, "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature," *Journal of Econometrics*, 2023, *235* (2), 2218–2244.

Sun, Liyang and Sarah Abraham, "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, 2021, *225* (2), 175–199.

Appendix

A. Additional Summary Statistics

Table A1 summarizes the data sources used in this study. Due to limitations in the NOMIS API's geographic coverage at the LA level, self-employment rates from the Annual Population Survey could not be reliably merged into the panel for the full sample, limiting the direct self-employment analysis discussed in the main text.

Table A1: Data Sources and Coverage

Source	Variable	Coverage	Frequency
Companies House	Firm registrations	All GB LAs	Daily
ONS NSPL	Postcode-to-LA mapping	2.52M postcodes	Snapshot
ONS Mid-Year	Population estimates	All GB LAs	Annual
DWP Rollout	UC full service dates	336 LAs (332 in panel)	One-time
NOMIS APS	Self-employment rate	Partial	Annual

B. Robustness: Extended Results

Table A2: Extended Robustness: Coefficient Comparison Across Estimators

Estimator	Coefficient	SE	p-value	N
Callaway-Sant'Anna (DR)	0.005	(0.019)	0.793	9,296
TWFE	0.020	(0.028)	0.462	27,888
Sun-Abraham (IW)	0.037	(0.032)	0.256	27,888
TWFE excl. pilots	0.029	(0.037)	0.430	26,040

Notes: Callaway-Sant'Anna uses the quarterly-aggregated panel (332 LAs $\times$ 28 quarters); TWFE and Sun-Abraham use the monthly panel (332 LAs $\times$ 84 months = 27,888); "excl. pilots" drops the 22 LAs treated before January 2017 (310 LAs $\times$ 84 months = 26,040). Sun-Abraham estimates computed via `fixest::sunab()` with aggregate ATT averaging over cohort-specific effects. Standard errors clustered at the LA level.

Acknowledgements

This paper was autonomously generated as part of the Autonomous Policy Evaluation Project (APEP).

Contributors: @olafdrw

First Contributor: https://github.com/olafdrw

Project Repository: https://github.com/SocialCatalystLab/ape-papers