

Bringing College to Town: Local University Access and Women's Educational Attainment in Iran

Reza Tavakoli Jaghargh
Department of Economics, Virginia Tech
rtavakoli@vt.edu

Abstract

Between 1982 and 2013 the Islamic Azad University, a tuition-charging private system, opened branches in 285 of Iran's 397 counties, bringing college within reach of provincial towns while leaving tuition in place. This paper uses the staggered timing of these openings, matched to the 2 percent microdata of the 2016 and 2011 censuses, to ask who gained when college came to town. Comparing birth cohorts within counties with a Callaway and Sant'Anna estimator on a county-by-cohort panel, I find that a local branch raised tertiary completion by about 4 to 5 percentage points among cohorts who grew up with it, on a base of 12 to 20 percent. The gain is concentrated among women; men show no effect. The effect is larger for residents of the county town than for rural residents, although the difference is imprecisely estimated. The same branches raised the share of young women who remain unmarried, and the share living with their parents, by 7 to 9 points. Children of more educated parents and of wealthier households show larger gains, and educational persistence across generations is not lower in treated counties. The results suggest that distance was a binding constraint on college attendance for women in particular.

JEL: I23, I24, J16, O15. Keywords: higher education expansion, private universities, college proximity, gender, Iran, staggered difference-in-differences.

1. Introduction

Two constraints keep able young people out of college in developing countries: the college is far away, and it costs money. Public university expansions usually relax both at once, which makes it hard to learn which constraint binds for whom. Iran's Islamic Azad University offers a setting in which the two are, at least partly, separated. Founded in 1982 as a private, tuition-charging system operating alongside a free but exam-rationed public sector, Azad opened branches in hundreds of provincial towns over three decades. Each branch cut the distance to college for local youth; it lowered, but did not remove, the monetary cost, because tuition remained and only the cost of living away from home fell. Who enrolled is informative about which constraint was binding.

This paper answers that question with two independent 2 percent census microdata samples (2016 and 2011) linked to a new dataset of the opening year of 363 Azad branches, geocoded to 397 counties on fixed 2011 borders. The research design follows Duflo (2001): within a county, birth cohorts too old to use the branch when it opened serve as controls for cohorts who were 18 or younger at the opening; county fixed effects absorb permanent differences across places and cohort fixed effects absorb national trends. Because openings are staggered over 30 years, the dynamic effects are estimated with the Callaway and Sant'Anna (2021) estimator on a county-by-cohort panel, treating the cohorts aged 19 to 22 at the opening as an anticipation window so that the base period is a cohort that could not have used the branch, and show that it agrees with Gardner's (2022) two-stage estimator and with two-way fixed effects on the same panel.

Three results follow. First, a local branch raised college completion. Averaged over all exposed cohorts, the tertiary share is 4.5 points higher (s.e. 1.2) in 2016 and 2.4 points higher (s.e. 0.9) in 2011, on a base of 12 to 20 percent. The effect grows with exposure: about 1 point for the first exposed cohort, 2 points after five cohorts, 3 to 4 points after ten, and 4 points after twelve on a panel whose county composition is held fixed. Effects of 6 to 8 points at twenty cohorts are estimated only from the counties that received branches in the 1980s. The pre-period coefficients are within half a point of zero, and a placebo that assigns openings ten years early to people who were never exposed gives 0.5 points (s.e. 0.5). Second, the gain goes to women. Women born around the opening gain 12 points (s.e. 2.2); men gain 1.4 points (1.3); the difference is 11 points (1.4). The estimate is larger in the county town, 4.0 points (1.9), than in villages, 0.7 points (1.2), but the urban-rural difference of 3.3 points (1.9) is not precisely estimated. In the 2011 census, which asks place of birth, the urban effect is present for people born in the same city and absent for the village-born, whether they stayed in the village or moved to town. Third, the branches raised the share of young women who are unmarried by 7 to 9 points and the share living with their parents by the same amount, reproducing on census data the marriage-delay effect that Karimi and Taghvatalab (2020) found in household survey data for the 2000s.

The paper also asks whether the gain went to families with resources. Among children living with a parent, those whose parents have secondary schooling or more show gains of 10 to 17 points, and those whose parents have primary schooling or less show none. Households in the middle and upper terciles of a current wealth index show gains of 6 to 10 points, against 1 to 3 points in the poorest tercile. The intergenerational regression coefficient of schooling and the rank-rank slope are not lower in treated counties. Because parent education is observed only for co-resident children and wealth is measured at the census, these results are exploratory. They suggest that a tuition-financed expansion mainly served families who could pay, as Blanden and Machin (2004) found for the British expansion.

The paper relates to three literatures. The first studies college proximity. Since Card (1995) and Kane and Rouse (1995), distance to college has been used as a source of variation in schooling, and much of this work treats distance as a cost that falls most heavily on students from low-income families (Frenette 2006; Carneiro, Heckman and Vytlačil 2011). New universities raise participation among local youth (Frenette 2009), and college openings have been used to measure how a mother's education affects her children (Currie and Moretti 2003; Carneiro, Meghir and Parey 2013). In Iran the distance constraint was relaxed for everyone, but the response is concentrated among urban women, which points to a constraint that depends on gender and not only on income.

The second literature studies the expansion of higher education. Large expansions in Europe and China changed who attends college and what graduates earn (Blanden and Machin 2004; Oppedisano 2011; Che and Zhang 2018), and cross-country evidence links new universities to regional growth (Valero and Van Reenen 2019). Most of this work concerns public systems. The closest paper, Karimi and Taghvatalab (2020), uses Iranian county-level enrollment in the 2000s from both the Azad and Payame Noor systems with marriage as the outcome. This paper extends the period to the full 1982–2013 Azad expansion, uses census attainment as the outcome, and adds the sex, place and family-background dimensions; Section 2 discusses the overlap with Payame Noor openings. The results also fit a body of work showing that the cost of access weighs more on girls, from village schools in Afghanistan (Burde and Linden 2013) to bicycles for schoolgirls in India (Muralidharan and Prakash 2017), and that schooling and the timing of marriage move together (Field and Ambrus 2008; Jensen 2012). In Iran, Moeeni and Tanaka (2023) show that women's college attendance responds to limits on their hiring.

Finally, the paper adds to the small literature on inequality of opportunity and educational persistence in Iran and the Middle East (Salehi-Isfahani, Hassine and Assaad 2014; Assaad, Hendy and Salehi-Isfahani 2019) and on local school supply and mobility (Assaad and Saleh 2018), with direct estimates from co-resident children. Like Asher, Novosad and Rafkin (2024), it has to work with parent–child links that are observed only when parents and children share a household.

The design has limits, and they are collected here. First, adding county-trait-specific cohort trends over the full cohort range reduces the effect to about one point. Section 8 argues that this over-controls, because opening dates are themselves determined by the traits; trends estimated on unexposed cohorts only, and province-by-cohort-group fixed effects, leave the effect intact. Second, the effect survives a smoothness restriction on pre-trends but not a relative-magnitudes restriction of the Rambachan and Roth (2023) type, because the horizon is long. Third, a joint test of the pre-period bins in the person-level event study rejects at the 5 percent level, because the two oldest bins are 0.6 to 1.5 points below zero; this may be a floor effect in cohorts with almost no tertiary education, and the panel estimator shows no such pattern. Fourth, openings within 60 km of never-treated counties show a 3-point spillover in one specification and none in others; if real, it biases the main estimates toward zero. Finally, 113 of the 363 opening years do not separate founding from first intake, the 2011 education sub-codes were inferred from the data, and the treatment does not separate Azad from Payame Noor openings in the same towns. None of the data choices tested in Section 8 moves the main estimate by more than 2.5 points.

Section 2 gives the institutional background, Section 3 the data, Section 4 the empirical strategy, Section 5 the main results, Section 6 heterogeneity by sex and place, Section 7 the exploratory evidence on family background, Section 8 placement, robustness and sensitivity, and Section 9 concludes.

2. Background: higher education in Iran and the Azad expansion

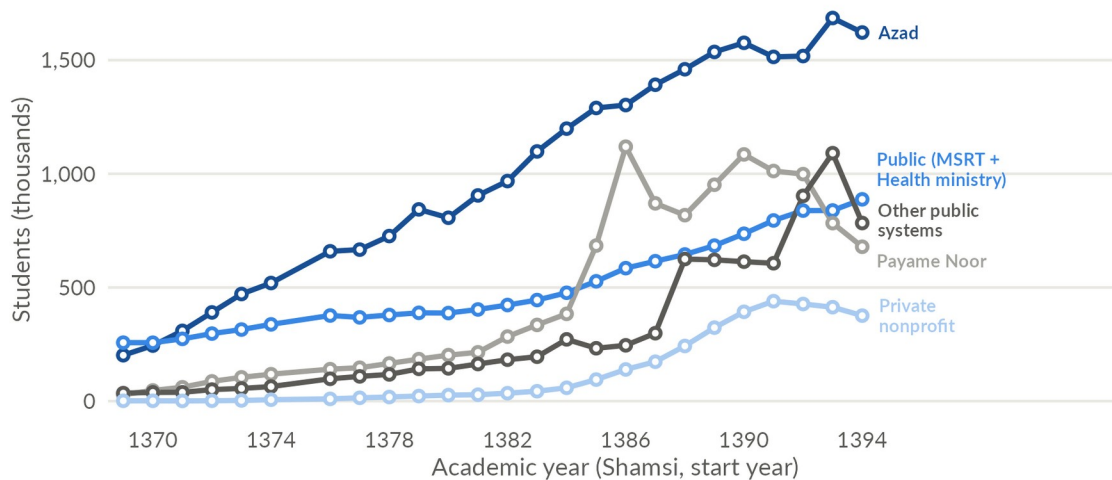
Iran's public universities are free and admit students through a national entrance examination (the *konkur*). Seats have always been far fewer than applicants, and the public system is concentrated in provincial capitals and large cities, so a public seat usually implies moving away from home. The Islamic Azad University was founded in 1361 (1982) as a non-governmental, self-financed institution. It charges tuition, admits through its own less selective examination, and expanded by opening branches (*vahed*) and smaller centers (*markaz*) in provincial towns.

Figure 1 documents the scale of the expansion with the national enrollment series of the Institute for Research and Planning in Higher Education (IRPHE). In 1369-70 (1990), eight years after its founding, Azad enrolled 202,000 students, 38 percent of all university students in Iran, against 258,000 in the public universities of the science and health ministries. Its enrollment then grew to 520,000 in 1374-75, 843,000 in 1379-80 and 1.2 million in 1384-85, and for the decade from the mid-1370s to the mid-1380s Azad accounted for about half of all students in the country, more than the entire public system. Enrollment was overwhelmingly at the bachelor's level (603,000 bachelor's, 192,000 associate and 34,000 master's students in 1379-80). After 1385 the Azad share fell to about 37 percent, not because Azad shrank (it reached 1.6 million students by 1394-95) but because the Payame Noor distance-learning university, which had fewer than 200,000 students

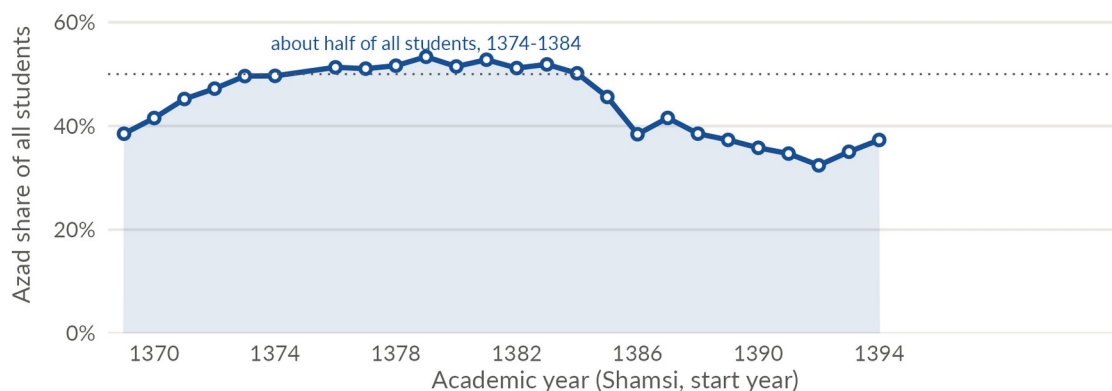
until 1383, tripled to 1.1 million within two years, and the applied-science and technical systems expanded. Payame Noor is a public system that also opened centers in many provincial towns; it was small during most of the period in which the cohorts studied here reached college age, but the two systems overlap in the towns and years of the late 1380s, and the treatment used in this paper does not separate them.

Figure 1. Students by university system, 1369-70 to 1394-95

A. Enrolled students by university system



B. Azad's share of national enrollment

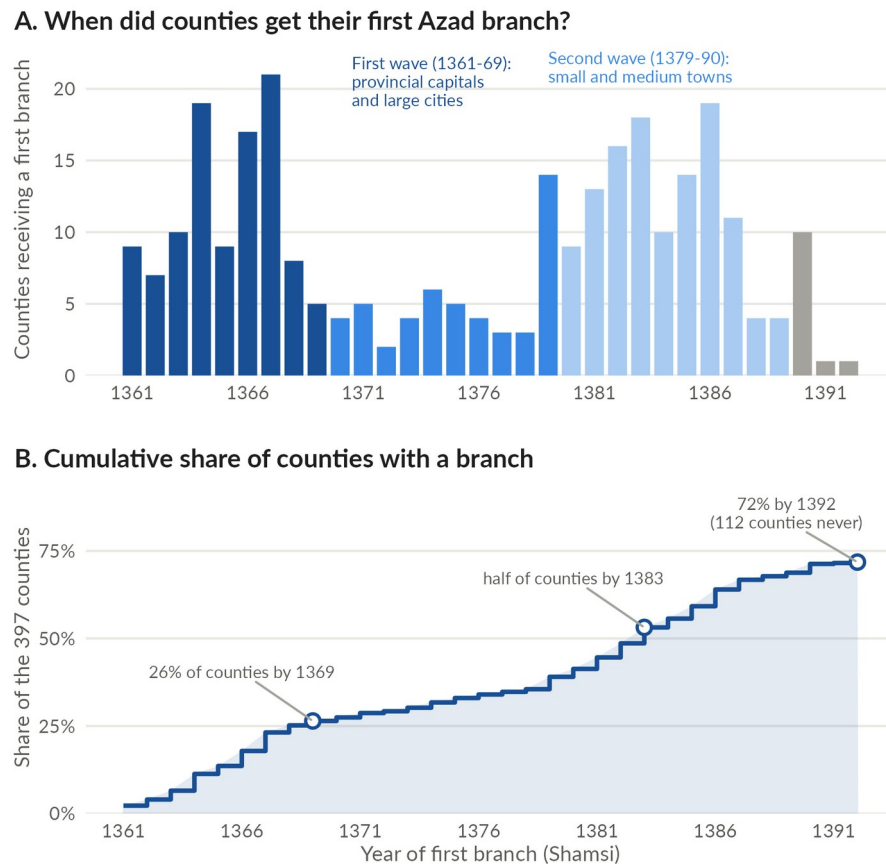


Notes: Source: IRPHE, students by affiliation. The Azad figure for 1375-76 is missing in the source and the year is omitted. “Public” combines the universities of the Ministry of Science and the Ministry of Health; “other public systems” combines the applied-science, technical, teacher-training and ministry institutions.

Two features of the expansion matter for identification and interpretation. First, the timing was staggered over three decades. Figure 2 shows that 9 counties received a branch in 1361 (1982), about 100 by 1367, 200 by 1382 and 285 by 1392, with 112 counties never receiving one. Second, placement was not random: larger and more literate counties received branches earlier (Section 8).

Figure 3 maps the pattern: the first wave in the 1360s went to provincial capitals and other large cities, the second wave in the 1380s filled in smaller towns, and the counties that never received a branch are mostly small and lie away from the main urban corridors. The research design therefore relies on the timing of openings within counties, not on the comparison of counties with and without branches.

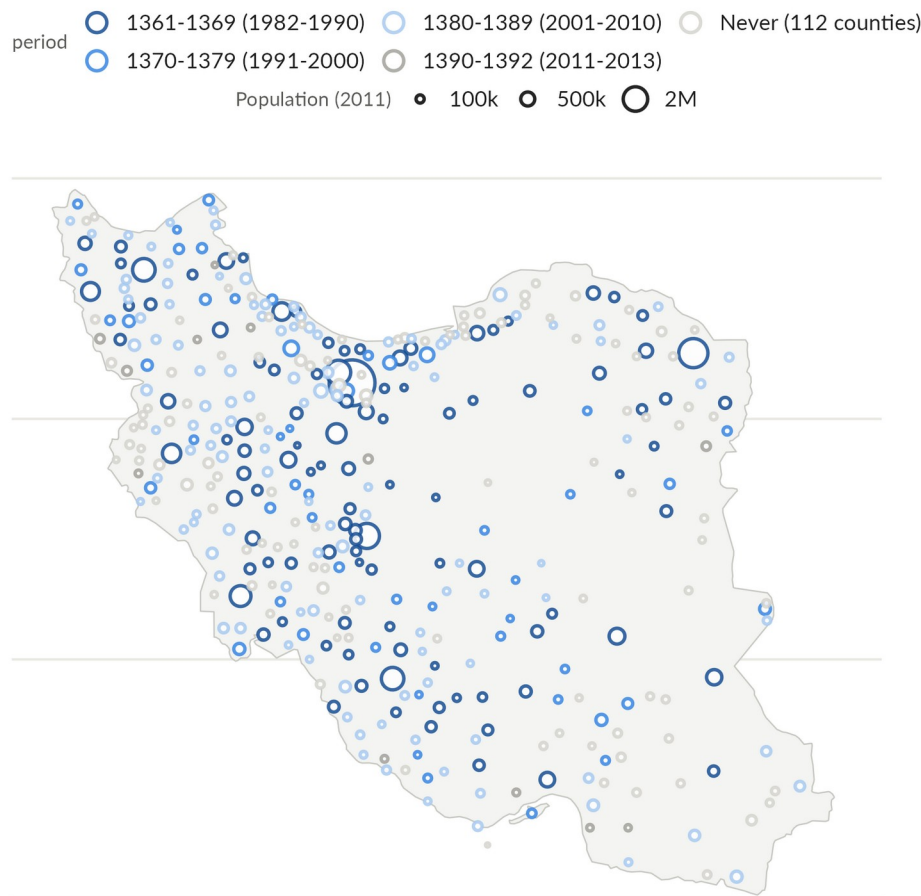
Figure 2. The Azad expansion across counties



Source: authors' branch dataset (363 branches), 1390 county borders. 285 counties ever received a branch; 112 never did.

Notes: Panel A counts counties by the year of their first branch; Panel B shows the cumulative share of the 397 counties (2011 borders) with at least one branch.

Figure 3. Where the branches went: first opening year by county



Notes: Each point is a county (2011 borders) at its center, colored by the year of its first Azad branch and sized by 2011 population. Sources: author's branch dataset; coordinates from English Wikipedia; population from the 2011 census sample.

3. Data

3.1 Branch openings

No official list of branch founding dates exists. One was built from the official history (*tarikhcheh*) page of each branch website, supplemented by the IAU knowledge dashboards, Persian Wikipedia, and the Azad news agency, for 363 branches; the official directory lists 406 units, of which about 40 are recent, merged, or duplicate entries. For each branch the year of first student intake, where stated, and the formal establishment year are recorded; where only one date is given it is used and the row is flagged. In 113 rows the source gives a founding year without separating the intake year. Branches were assigned to counties on 2016 borders by name (175 exact matches and 59 unique substring matches) and by hand (127 assignments with a confidence flag); two of the 363 could not be identified, leaving 361 assigned. The 33 counties created between 2011 and 2016 were mapped back to their 2011 parents. The treatment date for a county is the first opening year among

its branches. Robustness checks use the formal establishment year, drop the flagged rows, exclude *markaz* centers, and drop counties with low-confidence assignments; dropping branches never turns a treated county into a control, because the earliest surviving branch still defines treatment, but it can move the treatment date later.

3.2 Census microdata

The Statistical Center of Iran releases 2 percent household samples of the 1395 (2016) and 1390 (2011) censuses with county identifiers (1,579,435 and 1,481,586 persons). The 2016 census is the primary source: its exposed cohorts have completed university, its education variable is an official 1 to 8 scale, and it records residence five years earlier for movers. The 2011 census is used as an independent replication and for three things it alone provides: a place-of-birth question, a direct link from each person to his or her mother’s row in the household, and a five-year-earlier picture of co-residence. Its education variable is a three-digit code whose sub-codes are not documented in the released codebook; Appendix C explains how they were recovered from the data, and the tertiary indicator does not depend on the uncertain part. The analysis uses ordinary households and, for completed schooling, persons aged 25 to 64. Age is asked directly, so birth cohort is the census year minus age. Tertiary is an associate degree or higher.

Parent education is observed only for co-resident children: through the mother link and the head’s spouse in 2011, and through the “child of head” relationship in 2016. The household files provide dwelling ownership and area, and, in 2011, ownership of a car, a computer and central heating, from which a wealth index is built.

3.3 Samples and descriptive statistics

Table 1. Samples and descriptive statistics

	Census 1395 (2016)	Census 1390 (2011)
Persons in ordinary households (2% sample)	1,575,658	1,475,746
Persons aged 25–64 (main sample)	869,301	751,485
Weighted population represented (million)	78.9	74.0
Tertiary completion, ages 25+	0.214	0.138
Tertiary, men / women, ages 25–60, 18 or younger at opening	0.30 / 0.31	0.20 / 0.21
Tertiary, men / women, ages 25–60, 23 or older at opening	0.16 / 0.08	0.13 / 0.06
Share living with a parent, ages 18–24, men / women	0.86 / 0.53	0.84 / 0.57
Share living with a parent, ages 25–29, men / women	0.49 / 0.26	0.45 / 0.29
Counties (2011 borders)	397	397

	Census 1395 (2016)	Census 1390 (2011)
Counties ever treated / never treated	285 / 112	285 / 112
First opening year: min / median / max	1361 / 1379 / 1392	1361 / 1379 / 1392
Persons by age at opening: ≤ 18 / 19–22 / ≥ 23 / never treated	1,035,810 / 73,693 / 299,479 / 166,676	890,086 / 74,754 / 350,564 / 160,342

Notes: Census weights. Raw means mix cohort trends with treatment and are descriptive only.

4. Empirical strategy

Let y be an indicator for holding a tertiary degree, observed for person i of birth cohort c living in county j , and let T_j be the year the first branch opened in j . A person's age at the opening, $a = T_j - c$, determines exposure: people aged 18 or younger could enroll at the usual age, people aged 19 to 22 could enroll late, and people aged 23 or older are assumed not to have used the branch. The person-level event-study specification is

$$y_{ijc} = \sum_k \beta_k 1[a \in \text{bin } k] + \gamma \text{female}_i + \alpha_j + \delta_c + \varepsilon_{ijc},$$

with three-year bins of age at opening from -10 to 40 , the bin 23 – 25 as the omitted category, and never-treated counties contributing to the fixed effects only. The bins for ages 19 to 22 are estimated, not pooled with the reference. County fixed effects absorb permanent differences in schooling across places; cohort fixed effects absorb the national rise in schooling. Coefficients for bins with $a \geq 26$ are the pre-period test. Standard errors are clustered by county; observations are weighted by census weights. Each event study is summarized by two averages: over the bins with a from -10 to 4 , called “cohorts born around the opening” throughout the paper, and over the bins with a from 5 to 16 . In the 2011 census the youngest available bin is $a = -4$ to -2 , so the first average there runs from -4 to 4 . Standard errors of the averages are computed from the full cluster-robust covariance matrix of the coefficients. Differences between groups (women and men, urban and rural) are tested in a single model that interacts the bins with the group indicator and gives each group its own county fixed effects, cohort fixed effects and sex control, so that the interaction equals the difference between the two separate event studies and its standard error uses their joint covariance.

Because openings are staggered and effects grow with exposure, the data are also collapsed to a county-by-cohort panel (cell outcome = weighted tertiary rate; cell weight = sum of census weights; cells with fewer than 20 observations dropped) and Callaway and Sant'Anna (2021) group-time average treatment effects are estimated. The treatment cohort is $g_j = T_j - 18$, the first cohort aged 18 at the opening, and event time is $e = c - g_j$. The four cohorts before g_j (ages 19 to 22 at the opening) could enroll late, so they are treated as an anticipation window: the base period for every comparison is $e = -5$, the cohort aged 23 at the opening, and all pre-period effects are measured against that same base (a universal base period). With not-yet-treated controls, the effect for opening group g in cohort c is

$$ATT(g, c) = E[y_{jc} - y_{j,g-5} \mid G_j = g] - E[y_{jc} - y_{j,g-5} \mid G_j > c + 4],$$

where y_{jc} is the weighted tertiary rate in county j and cohort c , G_j is the county's treatment cohort, and the comparison counties are those whose first fully exposed cohort comes after $c + 4$, so that they are still outside the anticipation window (the alternative uses never-treated counties only). The group-time effects are aggregated by event time,

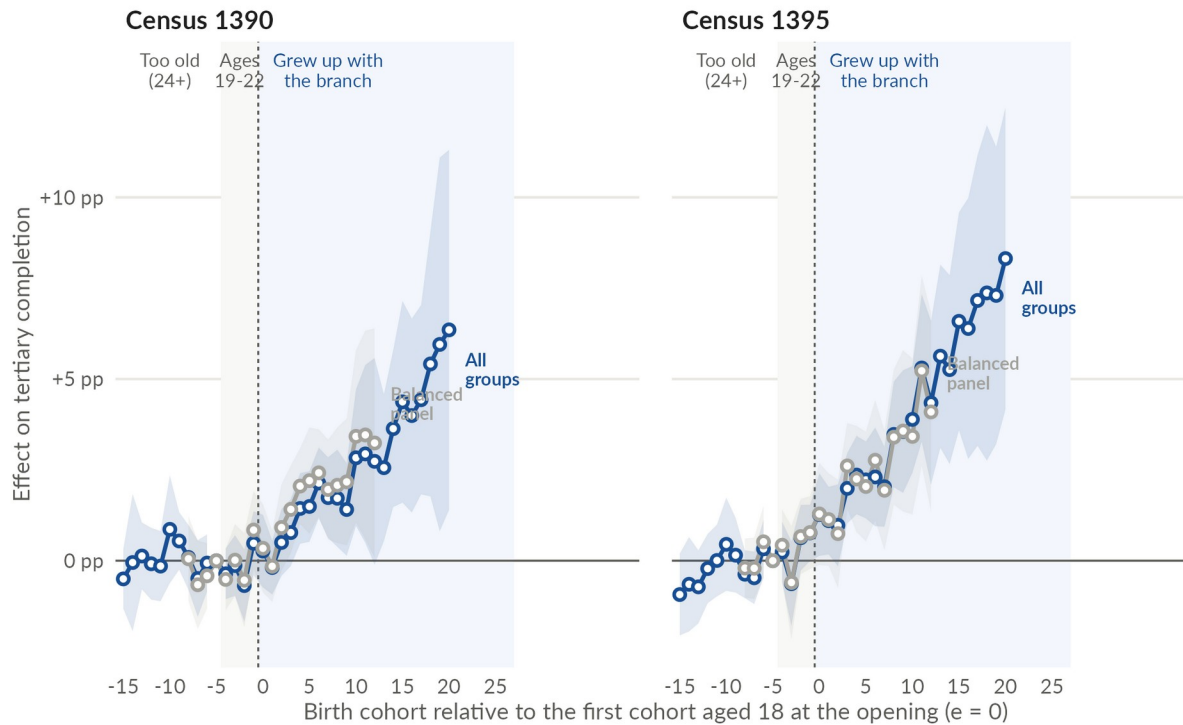
$$\theta(e) = \sum_g w_{g,e} ATT(g, g + e),$$

where $w_{g,e}$ is the weighted share of group g among the groups observed at event time e , and into a simple overall average over all treated group-cohort cells, with cluster bootstrap inference by county. Because the set of counties contributing shrinks as e grows, a balanced version is also reported restricted to opening groups observed at every event time from -8 to $+12$. Gardner's (2022) two-stage estimator and plain two-way fixed effects, both with reference period $e = -5$, are reported for comparison. Appendix Figure A1 shows how the pre-period changes under the base-period and anticipation choices: with no anticipation window and a varying base period, the partly exposed cohorts enter the comparison group, the effect is understated, and the pre-period looks flatter than it is.

The design has four threats. (i) Non-random placement: branches went first to larger, more literate counties. County fixed effects absorb levels; the concern is differential trends, which is addressed with pre-period estimates, sensitivity bounds, province-by-cohort-group fixed effects, and trait-specific cohort trends (Section 8). (ii) Migration: the census records current residence, and the target population is people currently living in the county. Results are shown dropping recent movers and, in 2011, by place of birth, with a statement of what each can and cannot establish. (iii) Spillovers: people in neighboring counties may use a branch. All counties are geocoded and effects are tested of openings within 30, 60 and 100 km on never-treated counties. (iv) Co-residence selection in the family-background analysis, which is tested directly and which is the reason that analysis is presented as exploratory.

5. Main results

Figure 4. Callaway–Sant’Anna event study: tertiary completion and cohort exposure to a local branch



Notes: County-by-cohort panel, persons aged 25–64. $e = 0$ is the first cohort aged 18 when the branch opened; the base period is $e = -5$ (age 23) and $e = -4$ to -1 are anticipation periods. Not-yet-treated controls. Bands are 95 percent cluster-bootstrap intervals by county (499 draws). The balanced panel keeps only opening groups observed at every event time from -8 to $+12$.

Figure 4 is the main result. In both censuses the pre-period coefficients (cohorts aged 24 or older at the opening) lie within about half a point of zero: their mean is -0.2 points in 2016 and 0.0 in 2011. The anticipation cohorts (ages 19 to 22) show a small rise, as late enrollment implies. From the first fully exposed cohort the effect grows steadily: 1.3 points at $e = 0$ (s.e. 0.6), 2.2 at $e = 5$, 3.9 at $e = 10$, 6.6 at $e = 15$ and 8.3 at $e = 20$ in 2016; 0.3, 1.5, 2.8, 4.4 and 6.4 points in 2011. The 2011 profile is lower because its most exposed cohorts had not finished university by 2011.

The growth of the effect with e has three sources that the aggregate profile does not separate: longer exposure of the individual, growth of the branch itself, and a changing set of counties. The third is important. Table 2 reports the number of opening groups and counties behind each event time: at $e = 0$, 28 groups and 224 counties contribute; at $e = 10$, 18 groups and 131 counties; at $e = 20$, only 8 groups (openings 1361–1368) and 96 counties. The estimates at long horizons therefore describe the counties that received branches in the 1980s, which are the larger and more literate ones. The balanced panel in Figure 4 (grey), which holds the composition fixed at 130 counties in 2016 and 112 in 2011, follows the full-sample profile closely where the two overlap and reaches 4.1 points (s.e. 1.4) at $e = 12$ in 2016 and 3.2 (1.6) in 2011. The paper takes 4 points over twelve

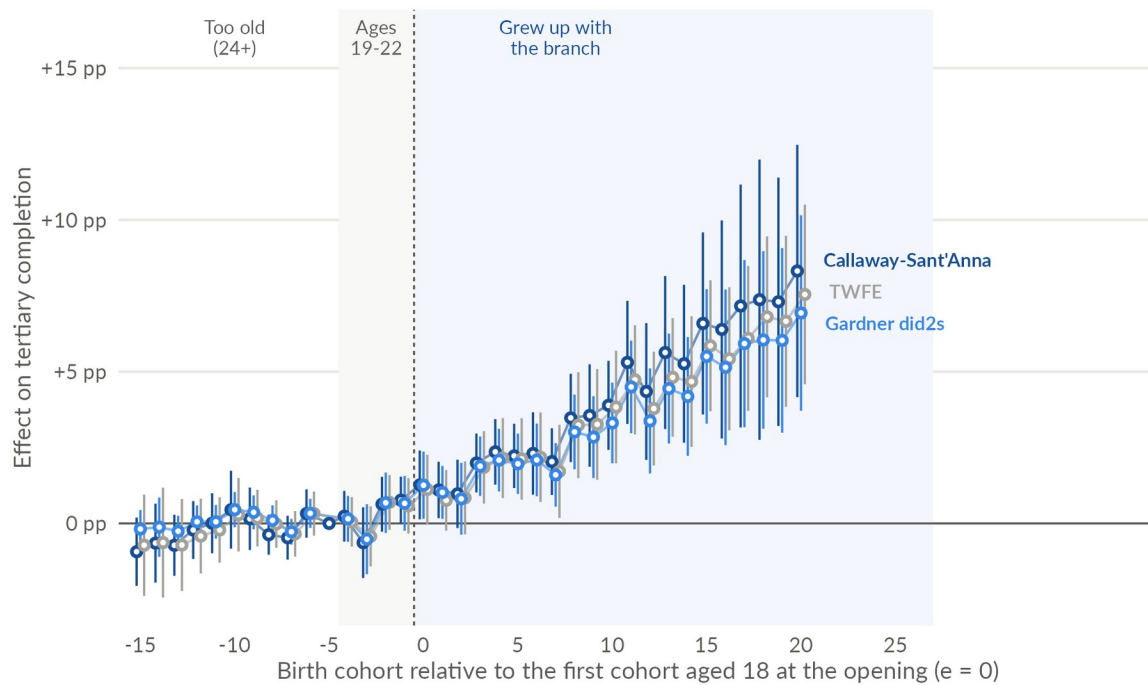
cohorts as the headline size and treats the 6 to 8 point estimates at twenty cohorts as effects for early-opening counties.

Table 2. Effect of a local Azad branch on tertiary completion

	Census 1395 (2016)	Census 1390 (2011)
A. Callaway–Sant’Anna, county x cohort panel (anticipation 4, universal base)		
Overall ATT, not-yet-treated controls	0.045 (0.012)	0.024 (0.009)
Overall ATT, never-treated controls	0.045 (0.013)	0.029 (0.012)
Overall ATT, balanced panel	0.047 (0.014)	0.029 (0.011)
Dynamic effect at $e = -10$	0.005 (0.007)	0.009 (0.008)
Dynamic effect at $e = 0$ (age 18 at opening)	0.013 (0.006)	0.003 (0.005)
Dynamic effect at $e = 5$	0.022 (0.005)	0.015 (0.005)
Dynamic effect at $e = 10$	0.039 (0.007)	0.028 (0.010)
Dynamic effect at $e = 12$, balanced panel	0.041 (0.014)	0.032 (0.016)
Dynamic effect at $e = 15$	0.066 (0.015)	0.044 (0.014)
Dynamic effect at $e = 20$	0.083 (0.021)	0.064 (0.025)
Mean of pre-period effects ($e \leq -6$)	-0.002	0.000
Groups / counties contributing at $e = 0$	28 / 224	—
Groups / counties contributing at $e = 10$	18 / 131	—
Groups / counties contributing at $e = 20$	8 / 96	—
Counties / cells	368 / 8,854	366 / 7,896
B. Person-level event study (ref. age 23–25 at opening)		
Cohorts born around the opening (age -10 to 4), average	0.069 (0.017)	0.044 (0.011)
Cohorts aged 5–16 at opening, average	0.029 (0.005)	0.018 (0.004)
Pre-period bins (age 26+), average	-0.005 (0.005)	-0.006 (0.003)
Joint test of pre-period bins, p-value	0.037	0.007
Observations	869,301	751,485

Notes: Panel A: cluster-bootstrap standard errors by county (499 draws). Panel B: county and cohort fixed effects, female dummy, census weights; averages over bins with delta-method standard errors from the full cluster-robust covariance; standard errors clustered by county. The balanced panel restricts to opening groups observed at every e from -8 to 12 (130 counties in 2016, 112 in 2011). The joint pre-period test in Panel B rejects because the bins for ages 35–40 at opening are 0.6 to 1.5 points below zero; see Appendix Figure A2.

Figure 5. Three estimators on the same panel, census 1395



Notes: Callaway–Sant’Anna (anticipation 4, universal base, not-yet-treated controls), Gardner (2022) did2s, and two-way fixed effects, all with reference period $e = -5$. 95 percent intervals clustered by county.

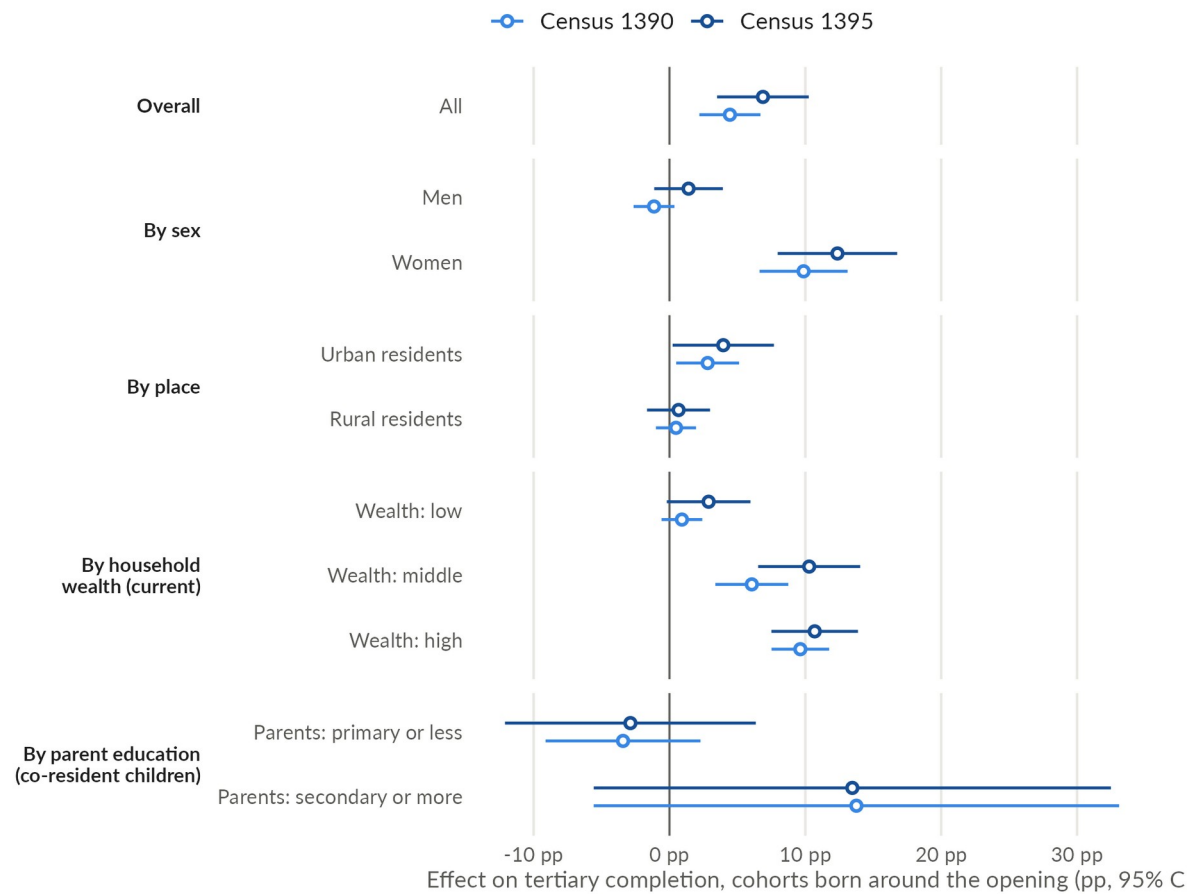
Figure 5 shows that the choice of estimator is immaterial once the reference period is the same: Callaway–Sant’Anna, did2s and two-way fixed effects give point estimates within half a point of each other at every event time. The reason is that 112 never-treated counties and many late-treated counties provide clean comparison units, so the forbidden comparisons that bias two-way fixed effects (de Chaisemartin and D’Haultfœuille 2020; Goodman-Bacon 2021; Sun and Abraham 2021) carry little weight here.

The overall ATT of 4.5 points and the long-horizon dynamic effects are different estimands. The first averages over every treated group-cohort cell, most of which are only a few cohorts past the first exposed one; the second describes cohorts with long exposure in early-opening counties.

6. Who gained? Sex and place

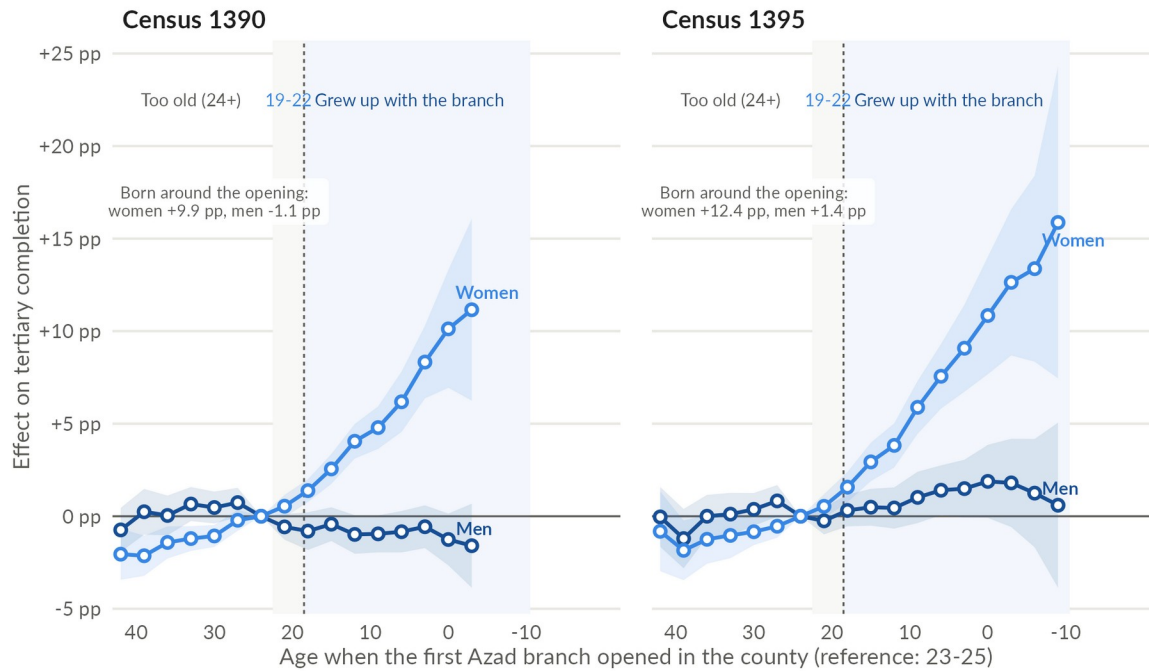
Figure 6 collects the effect for cohorts born around the opening for every group examined in this section and the next: by sex, by urban or rural residence, by current household wealth, and by parent education. The pattern is the same in both censuses. The gain is concentrated among women and among people in households with more resources, and the point estimate is larger for urban than for rural residents; the estimates for men, rural residents and the bottom wealth third are small and imprecise. The rest of this section takes the groups one at a time.

Figure 6. Who gained: effect for cohorts born around the opening, by group



Notes: Person-level event studies, ages 25–64 (parent-education rows: co-resident children aged 25–39 in 2011 and 25–34 in 2016). Effect is the mean coefficient over the bins with age –10 to 4 at opening (–4 to 4 in 2011), reference age 23–25. All rows except parent education use delta-method standard errors; the parent-education rows use root-mean-square standard errors and are exploratory. 95 percent intervals clustered by county.

Figure 7. Effect by sex



Notes: Person-level event study, ages 25–64, by sex. Reference bin: age 23–25 at opening. Shading marks the three exposure groups: too old (24 or older at opening), partly exposed (19–22) and fully exposed (18 or younger). The box gives the mean effect for cohorts born around the opening (age –10 to 4; –4 to 4 in 2011).

Table 3. Heterogeneity by sex and place

Sample	Born around opening (s.e.)	Age 5–16 at opening (s.e.)	CS overall ATT (s.e.)	N
<i>Census 1395 (2016)</i>				
All	0.069 (0.017)	0.029 (0.005)	0.045 (0.012)	869,301
Men	0.014 (0.013)	0.008 (0.006)	0.000 (0.010)	435,910
Women	0.124 (0.022)	0.051 (0.006)	0.071 (0.017)	433,391
Difference: women – men	0.110 (0.014)	0.042 (0.006)		
Urban residents	0.040 (0.019)	0.019 (0.008)	0.030 (0.012)	663,555
Rural residents	0.007 (0.012)	0.010 (0.005)	0.024 (0.008)	205,746
Difference: urban – rural	0.033 (0.019)	0.009 (0.008)		
<i>Census 1390 (2011)</i>				
All	0.044 (0.011)	0.018 (0.004)	0.024 (0.009)	751,485
Men	–0.011 (0.008)	–0.008 (0.004)	–0.015 (0.008)	373,478
Women	0.099 (0.017)	0.044 (0.005)	0.055 (0.016)	378,007
Difference: women – men	0.110 (0.013)	0.052 (0.006)		
Urban residents	0.028 (0.012)	0.010 (0.006)	0.010 (0.009)	557,689
Rural residents	0.005 (0.008)	0.004 (0.004)	0.014 (0.007)	193,796
Difference: urban – rural	0.023 (0.012)	0.006 (0.006)		

Notes: Person-level event-study averages with delta-method standard errors from the full cluster-robust covariance (clustered by county); group differences from a single model interacting the bins with the group indicator, with county fixed effects, cohort fixed effects and the sex control all specific to each group, so that the difference equals exactly the gap between the two separate event studies and its standard error comes from their joint covariance. CS overall ATT (anticipation 4, universal base, not-yet-treated controls, 199 cluster-bootstrap draws) from the corresponding county-by-cohort panel.

Figure 7 and Table 3 split the sample by sex. In 2016, women born around the opening are 12.4 points (s.e. 2.2) more likely to hold a degree; men are 1.4 points (1.3); the difference, estimated in one interacted model, is 11.0 points (1.4). In 2011 the numbers are 9.9 and -1.1 points, with a difference of 11.0 (1.3). The pattern holds in the panel estimator: the CS overall ATT is 0.071 (0.017) for women and 0.000 (0.010) for men in 2016, and 0.055 (0.016) and -0.015 (0.008) in 2011.

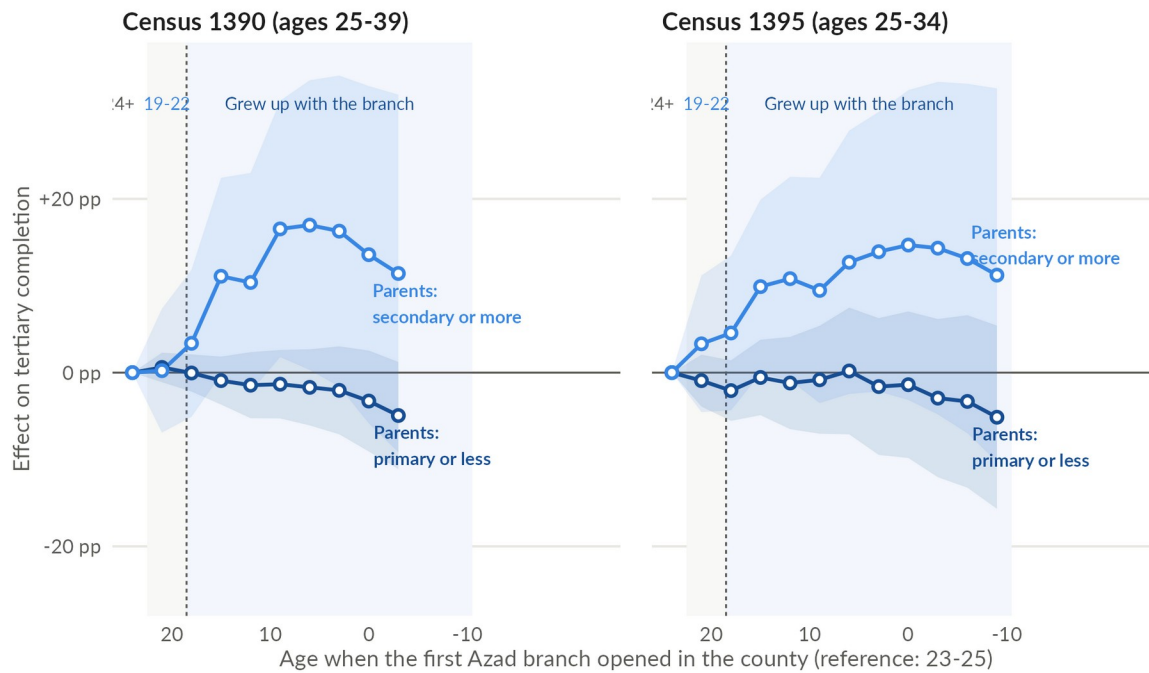
The estimate is larger in the county town, but the place result is weaker than the sex result. Urban residents show 4.0 points (1.9) in 2016 and 2.8 (1.2) in 2011; rural residents 0.7 (1.2) and 0.5 (0.8). The urban–rural difference is 3.3 points (s.e. 1.9, $p = 0.08$) and 2.3 (1.2, $p = 0.05$); the rural estimate is small, but the gap between the two groups is not precisely estimated. The panel estimator does not reproduce the gap: the CS overall ATT is 0.030 (0.012) for urban and 0.024 (0.008) for rural residents in 2016, and 0.010 (0.009) and 0.014 (0.007) in 2011. The rural event-study estimate for cohorts born around the opening is small, but rural cohorts with shorter exposure and the panel aggregate both show gains of one to two points, so the place contrast rests on the long-exposure cohorts only. Two checks address migration. Dropping people who moved into a city from a village in the five years before the 2016 census, or all movers, changes nothing (village-to-city movers are under one percent of urban residents), but this cannot speak to migration earlier in the treatment period. The 2011 census asks place of birth: the effect is 2.3 points (s.e. 1.1; CS ATT 0.025, s.e. 0.011) for urban residents born in the same city, 1.8 (1.5) for those born in another city, -1.8 (1.2) for those born in a village, and 0.5 (1.0) for rural residents who never left their village. These checks cannot rule out earlier migration, but the effect is present among urban residents born in their city and absent among the village-born, wherever they now live.

7. Family background and mobility: exploratory evidence

This section asks whether the gain went to families with resources. The evidence is suggestive, for three reasons. Parent education is observed only for children living with a parent, and Section 7.3 shows that a local branch changes which daughters live at home. Household wealth is measured at the census and is partly an outcome of the child’s own schooling, marriage and migration. And persistence measures compared across treated and never-treated counties are descriptive, because the counties differed before treatment.

7.1 Effects by parent education

Figure 8. Effect by parent education, children living with a parent



Notes: Person-level event study for co-resident children (ages 25–39 in 2011 using the mother link; 25–34 in 2016 using “child of head”). Parent education is the higher of the two parents. Reference bin: age 23–25 at opening.

Figure 8 splits co-resident children by the higher of their parents’ education. Children of parents with secondary schooling or more show gains of 10 to 17 points; children of parents with primary schooling or less show point estimates at or slightly below zero in both censuses. The reference and pre-period bins are wide because few co-resident children were 23 or older when their county’s branch opened (1,635 in 2016). Because co-residence responds to a branch for daughters but not for sons (Section 7.3), the split is also estimated for sons only: sons of parents with primary schooling or less show 0.1 points (2016) and –1.8 (2011); sons of parents with secondary schooling or more show 6.7 and 20.2 points, with standard errors of 12 to 15 points. The direction matches the pooled split, but these estimates are imprecise.

7.2 Persistence measures

Table 4. Intergenerational educational persistence, children living with a parent, ages 20–29 (descriptive)

	Census 1395 (2016)	Census 1390 (2011)
Intergenerational regression coefficient (child years on parent years)	0.333 (0.006)	0.284 (0.006)
Rank-rank slope	0.482 (0.021)	0.462 (0.024)
Regression coefficient: exposed (≤ 18 at opening) / never-treated counties	0.329 / 0.318	0.272 / 0.278
Rank-rank: exposed / never-treated	0.484 / 0.420	0.456 / 0.421

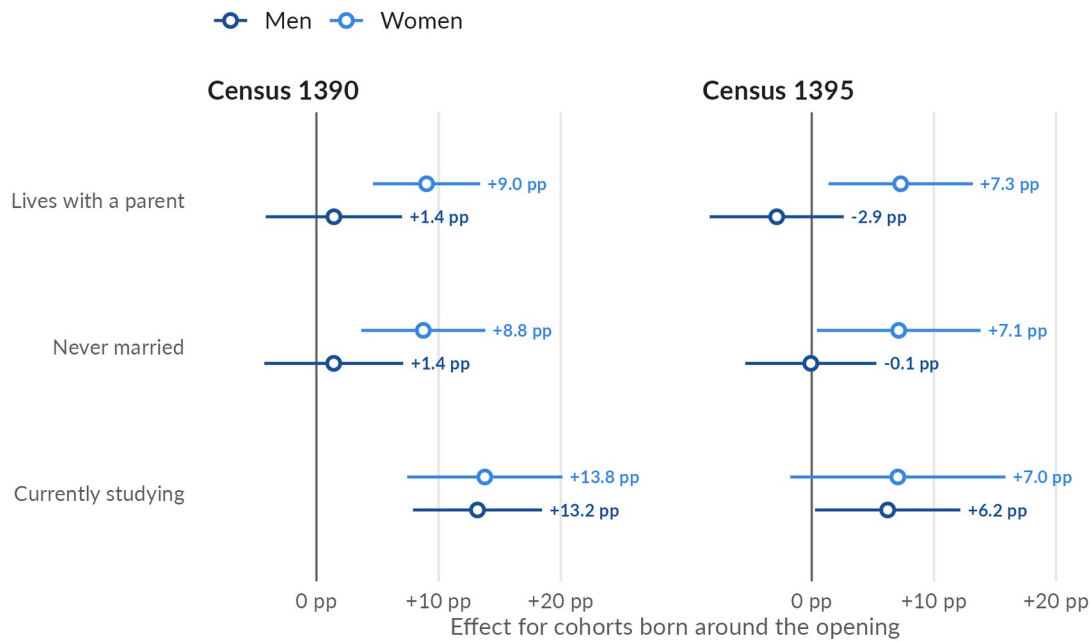
	Census 1395 (2016)	Census 1390 (2011)
counties		
P(child tertiary parents primary or less): exposed / never	0.32 / 0.27	0.27 / 0.21
P(child tertiary parents secondary, no diploma): exposed / never	0.50 / 0.44	0.44 / 0.35
P(child tertiary parents diploma): exposed / never	0.70 / 0.59	0.65 / 0.53
P(child tertiary parents tertiary): exposed / never	0.84 / 0.76	0.82 / 0.74
Observations	138,816	168,942

Notes: Weighted; cohort fixed effects and a female dummy in the regressions; standard errors clustered by county. Child outcome: years of schooling on the level scale (0, 5, 8, 11, 12, 14, 16, 18, 21), which is a regression coefficient in levels, not an elasticity; tertiary in 2011 includes current attendance, and schooling is not complete for all at ages 20–29. The exposed/never comparison is descriptive: never-treated counties are poorer and smaller.

Table 4 reports standard persistence measures on co-resident children aged 20 to 29. Persistence is high: a regression coefficient of 0.28 to 0.33 years per parental year and a rank-rank slope of 0.46 to 0.48, in line with the low mobility that Assaad, Hendy and Salehi-Isfahani (2019) document for Iran. The gradient is not lower in counties whose young people grew up with a branch, and the rank-rank slope is steeper there. Transition rates are higher for every parent group in treated counties, by 4 to 5 points for children of parents with primary schooling or less and by 11 to 13 points for children of parents with a diploma. These are cross-county comparisons and do not identify an effect of the expansion on mobility.

7.3 Co-residence, marriage and studying

Figure 9. Effects on living with a parent, never having married, and studying, ages 18–34



Notes: Person-level event study; effect for cohorts born around the opening (average of bins with age -10 to 4; -4 to 4 in 2011), by sex and census; the three outcomes are estimated separately.

Figure 9 tests whether a branch changes who lives at home. For men the estimated effect on living with a parent is -2.9 points (s.e. 2.8) in 2016 and 1.4 (2.9) in 2011. For women it is 7.3 points (3.0) in 2016 and 9.0 (2.2) in 2011; the effect on never having married is 7.1 and 8.8 points; the CS overall ATT for women's co-residence in 2011 is 0.031 (0.012). The share currently studying rises for both sexes. These are the marriage-delay effects of Karimi and Taghvatalab (2020) reproduced on census data, and they are consistent with the sex asymmetry in the main result: a local branch lets a daughter study without leaving home and postpones her marriage. They also mean that the daughters' side of Figure 8 is affected by selection into co-residence.

7.4 Household wealth

Table 5. Effect by current household wealth tercile, ages 25–64 (exploratory)

Wealth tercile	1395: born around opening (s.e.)	1395 CS ATT (s.e.)	1390: born around opening (s.e.)	1390 CS ATT (s.e.)
Bottom	0.029 (0.016)	0.018 (0.009)	0.009 (0.008)	0.004 (0.003)
Middle	0.103 (0.019)	0.063 (0.019)	0.061 (0.014)	0.030 (0.008)
Top	0.107 (0.016)	0.068 (0.013)	0.096 (0.011)	0.055 (0.019)

Notes: Wealth index: 2016 = home ownership, dwelling area per member, apartment; 2011 = car, computer, central heating, ownership, rooms and area per member; terciles within urban and rural. Event-study standard errors are delta-method values; CS ATTs use anticipation 4 and the universal base period. Current wealth is partly an outcome of own education, so the gradient is biased in a direction that is not known.

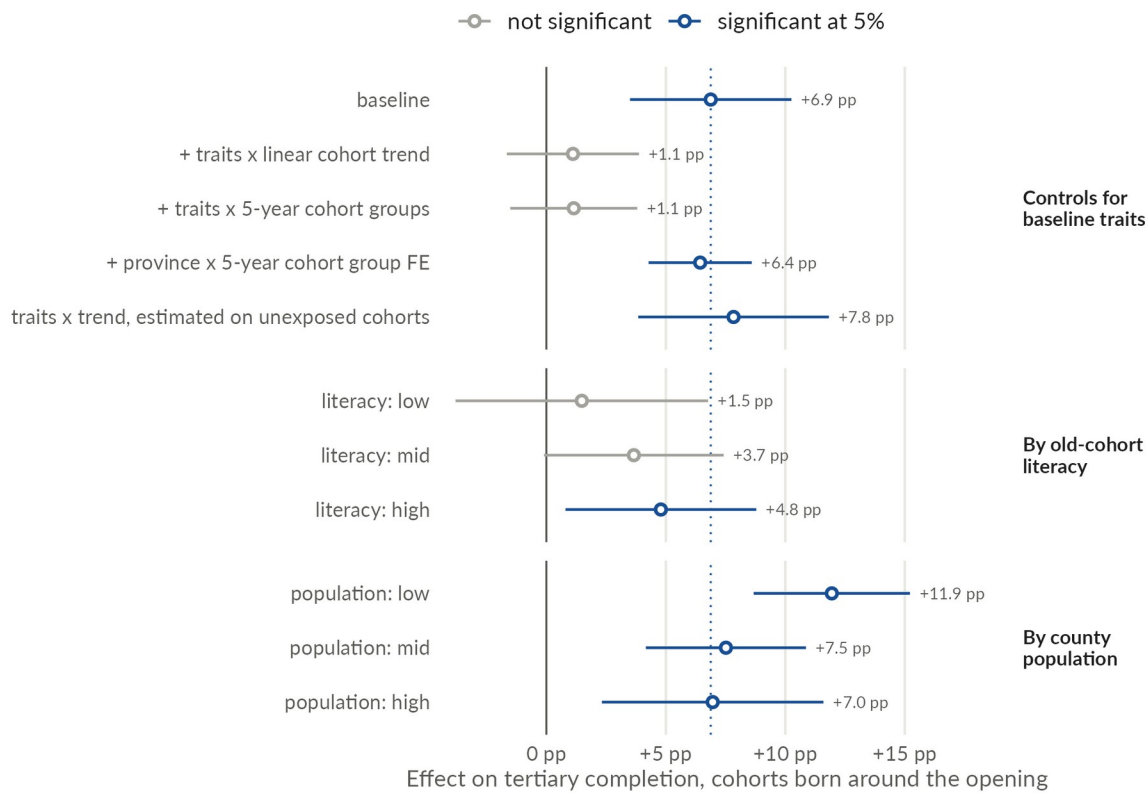
The wealth gradient is steep in both censuses: 1 to 3 points in the bottom tercile, imprecisely estimated, and 6 to 10 points in the middle and top. Because current wealth is partly an outcome, this cannot be read as the effect by pre-treatment resources. The pre-determined version, parents' wealth for co-resident children, has standard errors of 3 to 8 points and unstable pre-periods and is not informative. Together with the parent-education split, this pattern suggests that the expansion mainly served families who could pay.

8. Placement, robustness, and sensitivity

8.1 Non-random placement

Larger and more literate counties received branches first. Regressing the first opening year on standardized county traits measured on adults born before 1340 (1961), before any branch existed, one standard deviation higher adult literacy predicts an opening 3.8 years earlier (s.e. 0.7) and one standard deviation larger population 3.5 years earlier (0.5); urban share and distance to the provincial capital do not matter conditional on these. Literacy and population also predict ever receiving a branch (8 and 12 points per standard deviation). Identification rests on cohort timing within counties, and the question is whether counties that received branches early were also on steeper cohort trends in schooling.

Figure 10. Effect for cohorts born around the opening under alternative controls for county traits



Notes: Census 1395, county-by-cohort-by-sex cells, ages 25–64. Traits: old-cohort literacy, old-cohort urban share, log population, distance to the provincial capital, standardized. Delta-method standard errors, 95 percent intervals clustered by county.

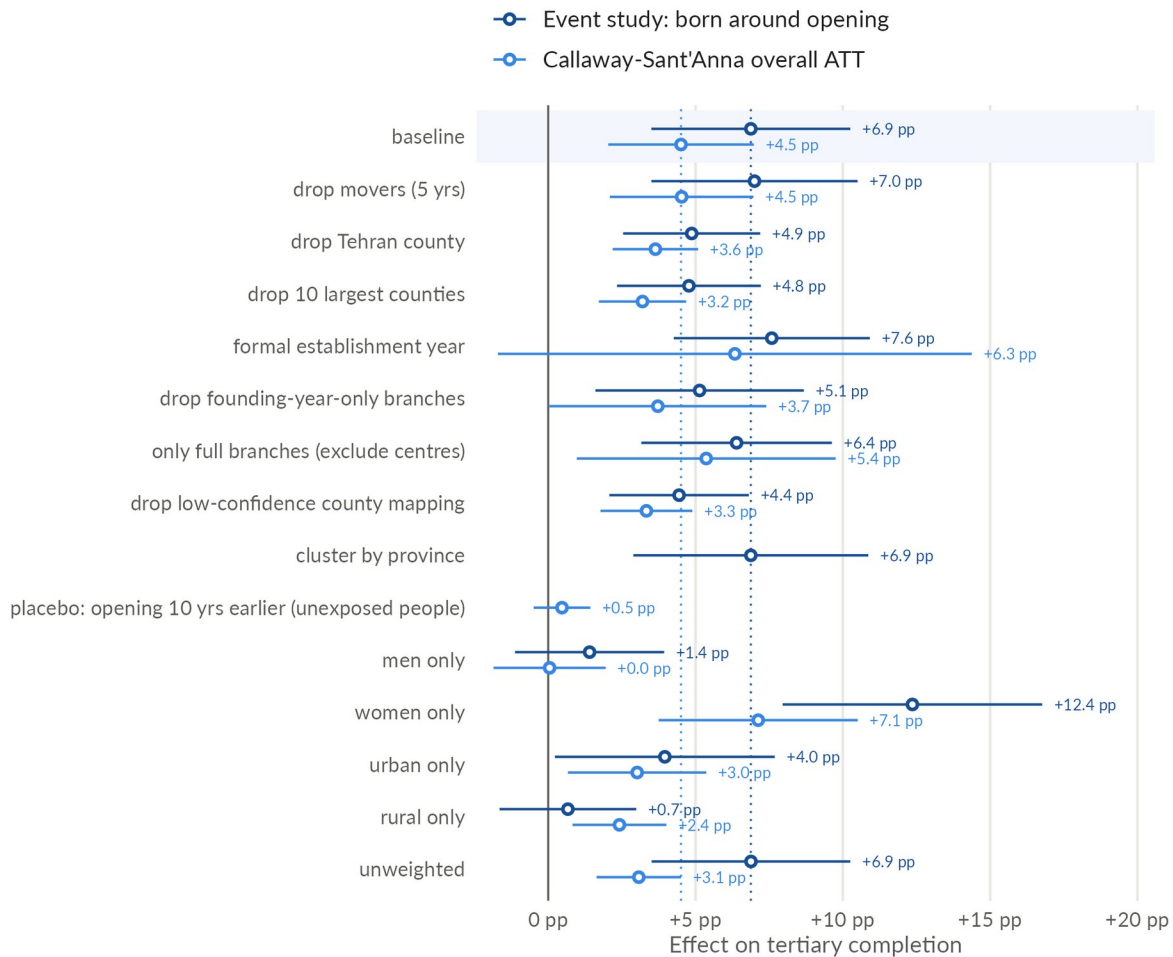
Figure 10 reports the effect for cohorts born around the opening under alternative controls. Adding the four traits interacted with a linear cohort trend, or with five-year cohort-group dummies, reduces the effect from 6.9 points to about 1.1 (s.e. 1.4). This is read as over-control, for two reasons. Opening timing is a function of literacy and size, so “literacy $\times$ cohort” is close to collinear with treatment, and an effect that grows across cohorts is absorbed by any trait-specific trend estimated over the treated cohorts (Wolfers 2006; Goodman-Bacon 2021). And the two specifications that avoid this problem leave the effect intact: estimating the trait-by-cohort trends on unexposed cohorts only (never-treated counties and cohorts aged 23 or older at opening) and removing them gives 7.8 points (s.e. 2.0) with a flat pre-period, and province-by-five-year-cohort-group fixed effects give 6.4 points (1.1). The effect is also present within every population tercile (11.9, 7.5 and 7.0 points in small, medium and large counties) and within the middle and top literacy terciles (3.7 and 4.8 points); it is 1.5 points (s.e. 2.7) in the least literate third. The extrapolation in the unexposed-cohort version assumes the pre-period trend continues unchanged; its own uncertainty is not included in the standard error. With full trait-specific trends, the effect cannot be distinguished from zero.

8.2 Pre-trend sensitivity

Two checks on the Callaway–Sant’Anna estimates (2016). First, a linear trend fitted to the pre-period coefficients has a slope of +0.08 points per cohort; extrapolating it through the post-period and subtracting leaves 0.9 points at $e = 0$, 2.6 at $e = 10$ and 6.2 at $e = 20$, so under a smoothness restriction on the trend violation the effect survives. Second, a relative-magnitudes restriction of the kind proposed by Rambachan and Roth (2023) bounds the post-period violation by a multiple M of the largest consecutive change observed in the pre-period, accumulated over the horizon. The largest consecutive pre-period change is 0.8 points; over horizons of 5 to 25 cohorts a violation of that size per cohort is as large as the estimated effect, so with $M = 1$ the bound does not exclude zero at any event time, and the breakdown value of M is about 0.2 (near zero in 2011). Over a horizon of 20 cohorts, a per-cohort violation one-fifth the size of the largest pre-period movement accumulates to the size of the effect; the identification argument therefore rests on the pre-period being close to flat, not on its being small relative to the effect. This calculation approximates the formal Rambachan and Roth bounds.

8.3 Robustness

Figure 11. Robustness of the main effect, census 1395



Notes: Each row re-estimates the person-level event study (effect for cohorts born around the opening) and the Callaway–Sant’Anna overall ATT. Event-study standard errors are delta-method values from the full cluster-robust covariance; CS ATTs use anticipation 4, the universal base period and not-yet-treated controls. 95 percent intervals clustered by county.

Table 6. Robustness, census 1395

Specification	Pre-period mean	Born around opening (s.e.)	CS overall ATT (s.e.)	N
Baseline	−0.005	0.069 (0.017)	0.045 (0.013)	869,301
Drop movers (other county or abroad, 5 yrs)	−0.005	0.070 (0.018)	0.045 (0.012)	833,336
Drop Tehran county	−0.008	0.049 (0.012)	0.036 (0.007)	760,843
Drop 10 largest counties	−0.008	0.048 (0.012)	0.032 (0.008)	575,831
Formal establishment year as opening	0.006	0.076 (0.017)	0.063 (0.041)	869,301
Drop branches with founding year only	0.010	0.051 (0.018)	0.037 (0.019)	869,301
Full branches only (exclude centers)	−0.004	0.064 (0.016)	0.054 (0.022)	869,301
Drop counties with low-confidence assignment	−0.008	0.044 (0.012)	0.033 (0.008)	698,148
Cluster by province	−0.005	0.069 (0.020)	—	869,301

Specification	Pre-period mean	Born around opening (s.e.)	CS overall ATT (s.e.)	N
Unweighted	-0.005	0.069 (0.017)	0.031 (0.007)	869,301
Placebo: opening 10 years earlier, unexposed people only	-0.006	n/a (0.005 at ages 5–16)	0.005 (0.005)	295,726

Notes: Pre-period mean: average coefficient over the bins for ages 26 or older at opening. Person-level standard errors are delta-method values from the full cluster-robust covariance, clustered by county (by province in that row). CS ATTs use anticipation 4, the universal base period and not-yet-treated controls, with 199 cluster-bootstrap draws. The placebo assigns each county a fake opening ten years before the true one and uses only people who were 23 or older at the true opening or live in never-treated counties.

The effect survives every change in sample and definition (Table 6, Figure 11). Dropping Tehran or the ten largest counties lowers the effect for cohorts born around the opening from about 7 to about 5 points and sharpens the panel ATT: the largest cities contribute large but noisy effects. Using the formal establishment year, dropping the 113 branches whose sources give only a founding year, excluding *markaz* centers, and dropping counties with uncertain assignment leave the effect between 4.4 and 7.6 points. The placebo estimate is 0.5 points (s.e. 0.5). Clustering by province raises the standard error from 1.7 to 2.0 points.

8.4 Spillovers

On the 112 never-treated counties (85,000 persons aged 25–64 in 2016), treating a county as exposed when a branch opens within 30 km gives an effect of -0.9 points (s.e. 2.9) for cohorts born around the nearby opening and a CS overall ATT of 0.030 (0.023); within 60 km, 0.7 (2.8) and 0.031 (0.011); within 100 km the control group shrinks to five counties and the estimates are uninformative. Among own-branch counties, cohorts too old for their own branch show no response to a nearby opening (0.9 points, s.e. 1.9). The 60 km panel estimate is statistically significant at conventional levels; the corresponding person-level estimate is not; the two agree on sign. If spillovers of about 3 points within 60 km are real, the not-yet-treated and never-treated controls in the main design are partly treated, and the main estimates are biased toward zero by that amount. The control counties in the balanced panel are, by construction, later-treated counties that are not more distant from earlier branches than average, so this bias is not specific to any one specification.

9. Conclusion

When a tuition-charging university opened a branch in an Iranian town, college completion rose by about 4 points over twelve cohorts among people who grew up with it, and by more in the counties that received branches earliest. The gain went to women; men did not gain. The estimate is larger for residents of the town than for rural residents, though the difference between the two is not precise. The same branches delayed young women's marriage and kept them in their parents' homes longer. The descriptive evidence on family background points to families with more

educated parents and more assets, and shows no fall in the persistence of schooling across generations.

The results point to a gender difference in the geographic constraint: for daughters of families that would not send them to another city, a branch in town removed a barrier that a free but distant public seat could not. Whether the money constraint also bound, and for whom, is the next question. Answering it requires tuition levels relative to household income by period, county enrollment data that separate new enrollment from substitution away from public universities, and a measure of family resources that is not itself an outcome. Labor-market outcomes for the women who gained are a natural extension.

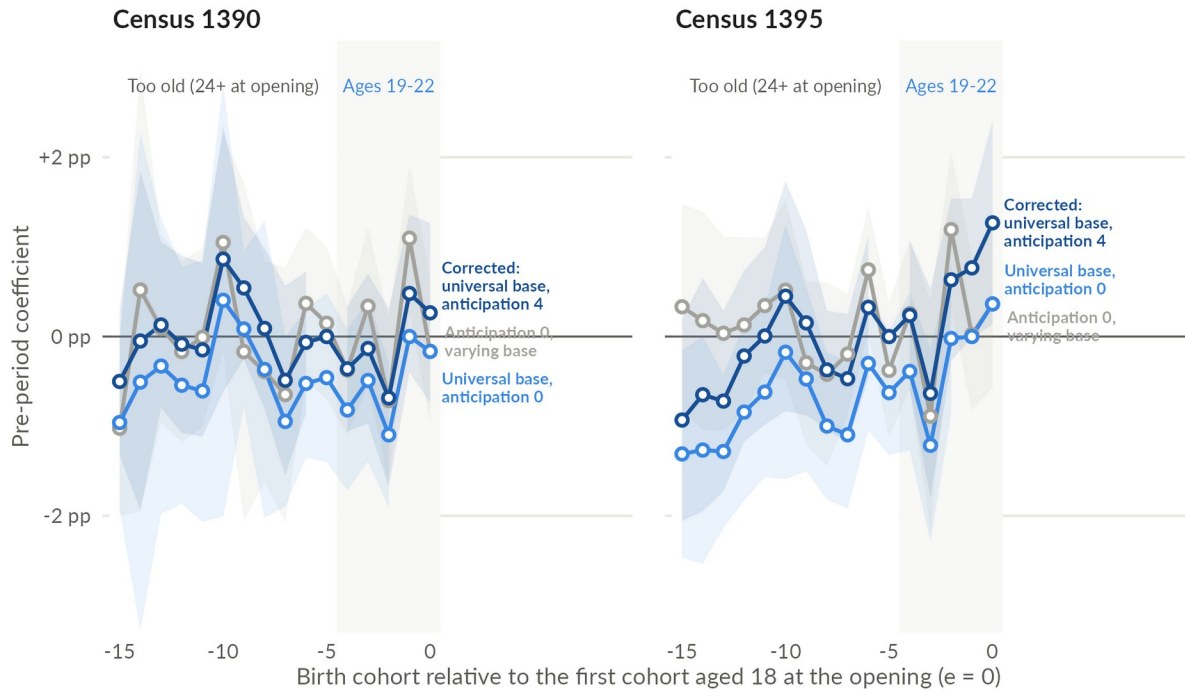
References

- Asher, S., P. Novosad and C. Rafkin (2024). Intergenerational mobility in India: new measures and estimates across time and social groups. *American Economic Journal: Applied Economics* 16(2): 66–98.
- Assaad, R., C. Hendy and D. Salehi-Isfahani (2019). Inequality of opportunity in educational attainment in the Middle East and North Africa: evidence from household surveys. *International Journal of Educational Development* 66: 24–43.
- Assaad, R. and M. Saleh (2018). Does improved local supply of schooling enhance intergenerational mobility in education? Evidence from Jordan. *World Bank Economic Review* 32(3): 633–655.
- Blanden, J. and S. Machin (2004). Educational inequality and the expansion of UK higher education. *Scottish Journal of Political Economy* 51(2): 230–249.
- Burde, D. and L. L. Linden (2013). Bringing education to Afghan girls: a randomized controlled trial of village-based schools. *American Economic Journal: Applied Economics* 5(3): 27–40.
- Callaway, B. and P. H. C. Sant’Anna (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics* 225(2): 200–230.
- Card, D. (1995). Using geographic variation in college proximity to estimate the return to schooling. In L. Christofides, E. K. Grant and R. Swidinsky (eds.), *Aspects of Labour Market Behaviour*. University of Toronto Press.
- Carneiro, P., J. J. Heckman and E. J. Vytlačil (2011). Estimating marginal returns to education. *American Economic Review* 101(6): 2754–2781.
- Carneiro, P., C. Meghir and M. Parey (2013). Maternal education, home environments, and the development of children and adolescents. *Journal of the European Economic Association* 11(S1): 123–160.
- Che, Y. and L. Zhang (2018). Human capital, technology adoption and firm performance: impacts of China’s higher education expansion in the late 1990s. *Economic Journal* 128(614): 2282–2320.
- Currie, J. and E. Moretti (2003). Mother’s education and the intergenerational transmission of human capital: evidence from college openings. *Quarterly Journal of Economics* 118(4): 1495–1532.
- de Chaisemartin, C. and X. D’Haultfœuille (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review* 110(9): 2964–2996.
- Duflo, E. (2001). Schooling and labor market consequences of school construction in Indonesia: evidence from an unusual policy experiment. *American Economic Review* 91(4): 795–813.

- Field, E. and A. Ambrus (2008). Early marriage, age of menarche, and female schooling attainment in Bangladesh. *Journal of Political Economy* 116(5): 881–930.
- Frenette, M. (2006). Too far to go on? Distance to school and university participation. *Education Economics* 14(1): 31–58.
- Frenette, M. (2009). Do universities benefit local youth? Evidence from the creation of new universities. *Economics of Education Review* 28(3): 318–328.
- Gardner, J. (2022). Two-stage differences in differences. arXiv:2207.05943.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics* 225(2): 254–277.
- Institute for Research and Planning in Higher Education (IRPHE). Higher education statistics of Iran: students by affiliation, 1369–70 to 1394–95. Tehran: IRPHE.
- Islamic Azad University. Directory of university units. <https://iau.ir/fa/center> (accessed 2026).
- Jensen, R. (2012). Do labor market opportunities affect young women’s work and marriage decisions? Experimental evidence from India. *Quarterly Journal of Economics* 127(2): 753–792.
- Kane, T. J. and C. E. Rouse (1995). Labor-market returns to two- and four-year college. *American Economic Review* 85(3): 600–614.
- Karimi, S. M. and G. Taghvatalab (2020). Access to higher education and the likelihood of being married: evidence from Iran. *Quarterly Review of Economics and Finance* 78: 22–33.
- Moeeni, S. and A. Tanaka (2023). The effects of a female hiring ceiling on college attendance: evidence from Iran. *Journal of Public Economics* 224: 104896.
- Muralidharan, K. and N. Prakash (2017). Cycling to school: increasing secondary school enrollment for girls in India. *American Economic Journal: Applied Economics* 9(3): 321–350.
- Oppedisano, V. (2011). The (adverse) effects of expanding higher education: evidence from Italy. *Economics of Education Review* 30(5): 997–1008.
- Rambachan, A. and J. Roth (2023). A more credible approach to parallel trends. *Review of Economic Studies* 90(5): 2555–2591.
- Salehi-Isfahani, D., N. B. Hassine and R. Assaad (2014). Equality of opportunity in educational achievement in the Middle East and North Africa. *Journal of Economic Inequality* 12(4): 489–515.
- Statistical Center of Iran (2011). National Population and Housing Census 1390: 2 percent public-use sample. Tehran.
- Statistical Center of Iran (2016). National Population and Housing Census 1395: 2 percent public-use sample. Tehran.
- Sun, L. and S. Abraham (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics* 225(2): 175–199.
- Valero, A. and J. Van Reenen (2019). The economic impact of universities: evidence from across the globe. *Economics of Education Review* 68: 53–67.
- Wolfers, J. (2006). Did unilateral divorce laws raise divorce rates? A reconciliation and new results. *American Economic Review* 96(5): 1802–1820.

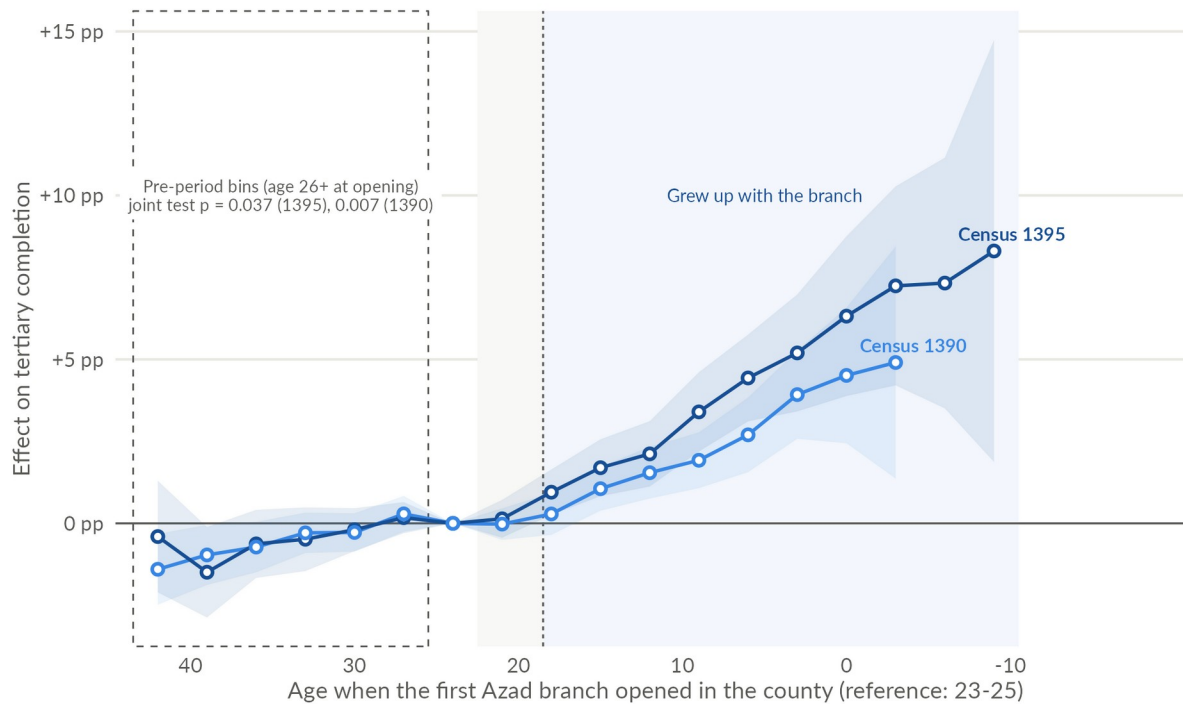
Appendix A. Pre-period diagnostics

Figure A1. Pre-period coefficients under three estimator setups



Notes: Callaway–Sant’Anna, county-by-cohort panel, not-yet-treated controls, pre-period only. Grey: no anticipation window, varying base period. Light blue: no anticipation, universal base at $e = -1$. Dark blue: anticipation window of four cohorts, universal base at $e = -5$ (the setup used in the paper). The varying base hides a level difference; with the base at $e = -1$, a partly treated cohort, earlier cohorts appear 0.5 to 1 point lower.

Figure A2. Person-level event study with the pre-period bins outlined



Notes: Ages 25–64, county and cohort fixed effects, reference bin age 23–25. The joint test of the outlined bins rejects ($p = 0.037$ in 2016, 0.007 in 2011) because of the bins for ages 35–40.

Appendix B. The branch dataset

The official IAU directory (iau.ir/fa/center) lists 406 units with name, province and address but no dates. For each unit the branch website and its history page were located, and the following were recorded: (i) the year of first student intake, (ii) the formal establishment or license year, (iii) the source type and (iv) a note on how the date was interpreted. 363 branches have a main opening year; in 250 rows the intake year is stated or coincides with establishment; in 113 rows only a founding year is given. Three slugs refer to the same Sonqor branch and two pairs (Amol, Chalus, Mianeh) are duplicates; they do not affect the first-opening year. The dataset, the county crosswalk and the county coordinates are in the replication files.

Appendix C. Education coding in the 2011 census

The 2011 file codes education with three digits: the first digit is the level (1 primary, 2 lower secondary, 3 upper secondary, 5 university, 6 seminary, 9 literacy movement and other), the last digit the grade or, for completed levels, a completion code. The official code sheet is not distributed with the sample. The meaning of each code was inferred from the joint distribution of age, current enrollment, literacy and code frequency (for example, code 511 is held by people with median age 22, all enrolled; code 512 by people with median age 32, none enrolled; the weighted counts, 2.5

and 3.8 million, match published enrollment and attainment figures for bachelor's programs). Tertiary completion is defined as codes 512, 514, 517 and 522, and "ever enrolled" as any 5xx code. The distinction between associate (52x) and bachelor (51x) programs affects the years-of-schooling scale but not the tertiary indicator.