Online Appendix

Link to Online Appendix I (Further Assessment of the Imputation)
Link to Online Appendix II (Alternative Household-Level Difference-in-differences)
Link to Online Appendix III (Parallel Trends and Synthetic Control)
Link to Online Appendix IV (Details of the Synthetic Control Procedures)
Link to Online Appendix V (Details on the Treasury Legal Clarification)

Online Appendix I. Further Assessment of Imputation

To assess the possibility that the procedure used to impute undocumented status introduces a bias towards lower homeownership among those classified as undocumented immigrants, I run a similar imputation procedure on the sample of U.S. citizens. If it is the procedure, itself, that drives the correlation between undocumented status and homeownership, then we should expect to see the same correlation arise among U.S. citizens who fulfill the imputation's criteria to be considered "undocumented" if it weren't for their citizenship status. I provide evidence that little, if any, of the observed relationship between undocumented status and homeownership arises mechanically from the imputation procedure employed.

I first return to the imputation procedure described in section 3.1, but instead apply each of the logical edits to citizens where applicable. The only difference in the imputation procedure applied to citizens is that any logical edit that relies on when a person arrived in the U.S. is not excluded.[1]

After citizens have been assigned their "pseudo-status" (the status they would be assigned by the imputation if they hadn't already been observed to be citizens), I restrict the sample in the same way the choice sample of immigrants was restricted in section 3.1[2] and generate summary statistics akin to those in Table 2. As can be seen in Table I.1, the raw ownership gap between undocumented immigrants and legal residents is much larger than the equivalent gap between citizens who are categorized as undocumented and citizens categorized as legal residents by a similar procedure. In other words, if a homeownership gap of 3 percentage points is attributable to the imputation procedure (because that is approximately

[1]This means that the edits to account for likely student visa holders, individuals who likely achieved legal status through IRCA 1982, and those who are likely in the U.S. on H-1B visas are not applied. Additionally, if a citizen's spouse is a citizen, they are not assigned legal resident status. However, if an individual's spouse has been assigned legal resident status by another logical edit, that individual *is* considered to be a legal resident by the last edit of the imputation procedure.

[2]The exception is that the sample is not restricted to those with a years in the U.S. term of 0 or greater than 37 because years in the U.S. is not meaningful for the majority of the sample of citizens.

the observed difference between "pseudo-undocumented and pseudo-legal residents"), then an unexplained gap of roughly 5 percentage points (as opposed to 8) between immigrants of different statuses still remains. Alternatively, if the imputation procedure mechanically drives those who are legal residents to be 3.8 (the percent change from 0.7066 to 0.7333) percent more likely to be homeowners, then legal residents are still nearly 18 percent more likely to be homeowners than undocumented immigrants (as opposed to roughly 21.5% more likely). In short, Table I.1 illustrates that very little of the raw homeownership gap between undocumented immigrants and legal residents can be attributed to any mechanical correlation that could arise from the imputation procedure used to assign immigrant status.

To further buttress the argument that the imputation procedure only negligibly influences the association between undocumented status and lower homeownership rates (if at all), I rerun descriptive regressions like those in section 3.2. Table I.2 presents results from the various descriptive regression specifications run on the sample of citizens who have been assigned their "pseudo-status" (i.e. the sample includes only citizens, and "undocumented" is now 1 if the citizen was categorized as "undocumented" by the modified imputation procedure and 0 otherwise). Columns 1-3 are identical to columns 4-6 in Table 3 and are provided for reference. Note that once controls are included, citizens classified as undocumented by the procedure are actually more likely to be homeowners, suggesting the imputation procedure applied to immigrants in the text may even yield estimates that are *lower* in magnitude than the true effect (i.e. the effect absent any mechanical bias from the imputation procedure). Even the negative coefficient estimates observed in the specifications that lack controls (columns 4 and 5) are of much smaller magnitudes than those observed for the sample of immigrants in columns 1 and 2. Altogether, there appears to be little evidence to suggest that the magnitude of the homeownership gap between undocumented immigrants and legal residents is inflated mechanically by the imputation procedure employed to assign immigrant status.

	Legal Resident	Undocumented	Citizen	Pseudo-Legal Resident	Pseudo-Undocumented
owned	0.4217	0.3469	0.7219	0.7333	0.7066
age	45.92	40.8	54.17	62.88	45.02
male	0.5564	0.6106	0.5155	0.4999	0.5372
married	0.7089	0.5318	0.5281	0.5071	0.5343
years in us	17.61	14.7	NA	NA	NA
monthly income (2010 dollars)	4489	4206	6075	4537	6571
people in household	3.631	3.594	2.408	2.178	2.677
workers in household	1.49	1.66	1.135	0.7446	1.581
children in household	1.249	1.267	0.5196	0.3525	0.7066

Table I.1: Summary statistics for the household-level microdata sample by immigrant status. Columns 1-3 are equivalent to Table 2. Columns 4 and 5 are derived from the sample of citizen households after undergoing the imputation procedure used to assign undocumented status as described in this section.

	owned	owned	owned	owned	owned	owned
(Intercept)	0.3735^{***}			0.6759^{***}		
	(0.0099)			(0.0036)		
undoc	-0.0710^{***}	-0.0928^{***}	-0.0206^{***}	-0.0292^{***}	-0.0263^{***}	0.0765^{***}
	(0.0054)	(0.0043)	(0.0029)	(0.0036)	(0.0030)	(0.0012)
years in U.S.			0.0141^{***}			
			(0.0007)			
years in U.S.^{2}			-0.0002^{***}			
			(0.0000)			
age			0.0163^{***}			0.0261^{***}
			(0.0005)			(0.0002)
age^2			-0.0001^{***}			-0.0002^{***}
			(0.0000)			(0.0000)
log(income)			0.0413^{***}			0.0576^{***}
			(0.0013)			(0.0007)
never married			-0.1407^{***}			-0.2411^{***}
			(0.0034)			(0.0021)
female			0.0196^{***}			-0.0157^{***}
			(0.0021)			(0.0008)
number workers			-0.0040^{**}			0.0127^{***}
			(0.0016)			(0.0007)
number people			0.0203^{***}			0.0002
			(0.0022)			(0.0009)
number kids			-0.0192^{***}			0.0039^{***}
			(0.0019)			(0.0009)
Fixed Effects	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes
Num. obs.	468960	468960	468960	10511358	10511358	10511358
Adj. R^2	0.0054	0.0945	0.2152	0.0009	0.0564	0.2690
N Clusters	1077	1077	1077	1078	1078	1078

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table I.2: Linear probability models for housing tenure (owned = 1) where columns 1-3 are run on the sample of immigrant households (equivalent to columns 4-6 of Table 3) and columns 4-6 are results from similar regressions run on citizen households that have been classified as undocumented or legal resident by the modified imputation procedure. Column 1 (2) is specified identically to column 4 (5). Column 3 includes controls for years in the U.S. and its square, whereas column 6 does not as years in the U.S. is not meaningful for most citizens (and would be almost perfectly collinear with age). Robust standard errors clustered at the CPUMA level (the most precise geographic variable available). All regressions use household weights provided by the ACS.

Online Appendix II. Alternative Household-Level Difference-in-differences

The specification chosen in Section 4.1 may be altered to focus on households where DACA is most likely to have an effect. In this section, I assign each household head an indicator that takes value 1 if anyone in the household meets the eligibility criteria for DACA. Specifically, any household in which any individual is born after 1980, has been in the U.S. since at least 2007, and arrived in the U.S. when they were no older than 16 is assigned a value $daca\ in\ hh = 1$. If the sample is restricted to undocumented households only, then the following specification could verify that the change in share of households residing in owner-occupied housing is driven by households in which at least one member was plausibly eligible for the program.

$$owned_{ipt} = \beta_1 daca\ in\ hh_i + \beta_2(daca\ in\ hh_i \times post_t) + X_i\theta + \alpha_p + \gamma_t + \varepsilon_{ipt} \tag{II.1}$$

This specification (or a triple differences specification) is not the choice specification for this paper for two reasons. First, this formulation does not account for any cases where a DACA recipient purchases a home in their name but does not live in that home. DACA recipients, who are primarily young adults with family members (of various statuses) living in the U.S., may use their DACA status as an avenue to procure a home loan for family members (e.g. parents) who would otherwise be restricted to mortgages offered to individuals without social security numbers, which are more limited in their prevalence and may be prohibitively costly in their terms. As an example, a DACA recipient may leave her parents' rental housing at 18 to move into her own apartment. Her parents have incomes (and willingness to pay) sufficient to afford the terms of a home loan for which she is eligible. She takes out the mortgage but remains in her apartment. Her parents (and perhaps siblings) move into the home and reimburse her for the mortgage payments. If the home the young DACA recipient can afford is small, it may be especially likely that she ends up living elsewhere to avoid crowding.

Second, as shown in Figures II.1 - II.3, it is less clear that the parallel trends assumption holds in these specifications, making it difficult to claim that the effect size is not biased due to pre-trends. If the trends are not believed to be parallel, then the estimated effects should be treated as upper bounds, and it is impossible to determine whether their statistical significance would remain absent the trends.

Nonetheless, if the trends are assumed to be parallel, the interpretation of the estimated effects is similar to the interpretation of the effects found in Section 4.1. The primary difference is that these estimates, while still "intent-to-treat" effects, are closer to the effect of "treatment on the treated."[3] The results are included in the table below. The first 3 columns replicate the results from Section 4.1 for comparison.

	owned	owned	owned	owned	owned	owned
undoc	-0.1154^{***}	-0.0242^{***}	-0.0293^{***}			
	(0.0050)	(0.0033)	(0.0034)			
undoc × post	0.0413^{***}	0.0065^{*}	0.0090^{***}			
	(0.0039)	(0.0035)	(0.0035)			
daca in hh				-0.1177^{***}	-0.0770^{***}	-0.0791^{***}
				(0.0075)	(0.0069)	(0.0068)
daca in hh × post				0.0326^{***}	0.0455^{***}	0.0490^{***}
				(0.0091)	(0.0087)	(0.0086)
years in U.S.		0.0141^{***}	0.0144^{***}		0.0130^{***}	0.0136^{***}
		(0.0007)	(0.0007)		(0.0007)	(0.0007)
years in U.S.^{2}		-0.0002^{***}	-0.0002^{***}		-0.0001^{***}	-0.0001^{***}
		(0.0000)	(0.0000)		(0.0000)	(0.0000)
age		0.0163^{***}	0.0173^{***}		0.0084^{***}	0.0110^{***}
		(0.0005)	(0.0006)		(0.0008)	(0.0009)
age^2		-0.0001^{***}	-0.0001^{***}		-0.0000^{*}	-0.0001^{***}
		(0.0000)	(0.0000)		(0.0000)	(0.0000)
log(income)		0.0413^{***}			0.0365^{***}	
		(0.0013)			(0.0013)	
never married		-0.1407^{***}	-0.1572^{***}		-0.1212^{***}	-0.1336^{***}
		(0.0034)	(0.0036)		(0.0034)	(0.0036)
female		0.0196^{***}	0.0101^{***}		0.0225^{***}	0.0126^{***}
		(0.0021)	(0.0020)		(0.0029)	(0.0028)
number workers		-0.0040^{**}	0.0240^{***}		-0.0207^{***}	0.0049^{***}
		(0.0016)	(0.0016)		(0.0019)	(0.0018)
number people		0.0203^{***}	0.0167^{***}		0.0291^{***}	0.0250^{***}
		(0.0022)	(0.0022)		(0.0024)	(0.0025)
number kids		-0.0192^{***}	-0.0174^{***}		-0.0207^{***}	-0.0181^{***}
		(0.0019)	(0.0019)		(0.0021)	(0.0021)
Controls	No	Yes	Yes*	No	Yes	Yes*
Adj. R^2	0.0949	0.2152	0.2037	0.0783	0.1895	0.1800
Num. obs.	468960	468960	468960	273768	273768	273768
N Clusters	1077	1077	1077	1077	1077	1077
Outcome Mean	0.3780	0.3780	0.3780	0.3469	0.3469	0.3469

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table II.1: Difference-in-differences regression results for the *owned* indicator. Columns 1-3 are identical to Table 4 and are provided for reference. Columns 4-6 are based on equation (II.1). In these regressions, the sample is restricted to undocumented households. As with the first 3 columns, columns 4-6 differ from each other only in their sets of controls. Column 4 includes no controls beyond CPUMA and year fixed effects. Column 5 includes the full set of controls as listed in Section 3. Column 6 includes the same controls except that log(income) is omitted as income is likely a bad control. Robust standard errors clustered at the CPUMA level.

[3] Only a small fraction of the undocumented population (the treated group in Section 4.1) received DACA, but roughly half of the DACA-eligible population (the treated group here) did.

Regardless of choice of controls, all three specifications find significant positive effects of DACA for households in which at least one member is eligible. The estimated effects of a four percentage point increase in homeownership propensities is notably larger than the effects in choice specifications. One explanation for this is that the proportion of the sample affected by treatment is several times larger here, meaning the intent-to-treat to effects more closely approximate what the treatment-on-treated effects would be (if it were possible to determine which individuals in the sample actually took up DACA). In other words, the treated group in these specifications is less contaminated by untreated households, which would bias estimates towards zero. However, given the event studies presented in Figures II.1 - II.3, it may be that effects are (artificially) larger due to a positive bias that could arise as a result of the failure of the parallel trends assumption. So, while the unbiasedness of the estimates in the final three columns of Table II.1 is subject to one's interpretation of the event studies below, the fact that estimates are, at least, in line with expectations is somewhat reassuring (the bias would have to be exceptionally large to yield significant and negative effects that would contradict the findings from choice specifications).

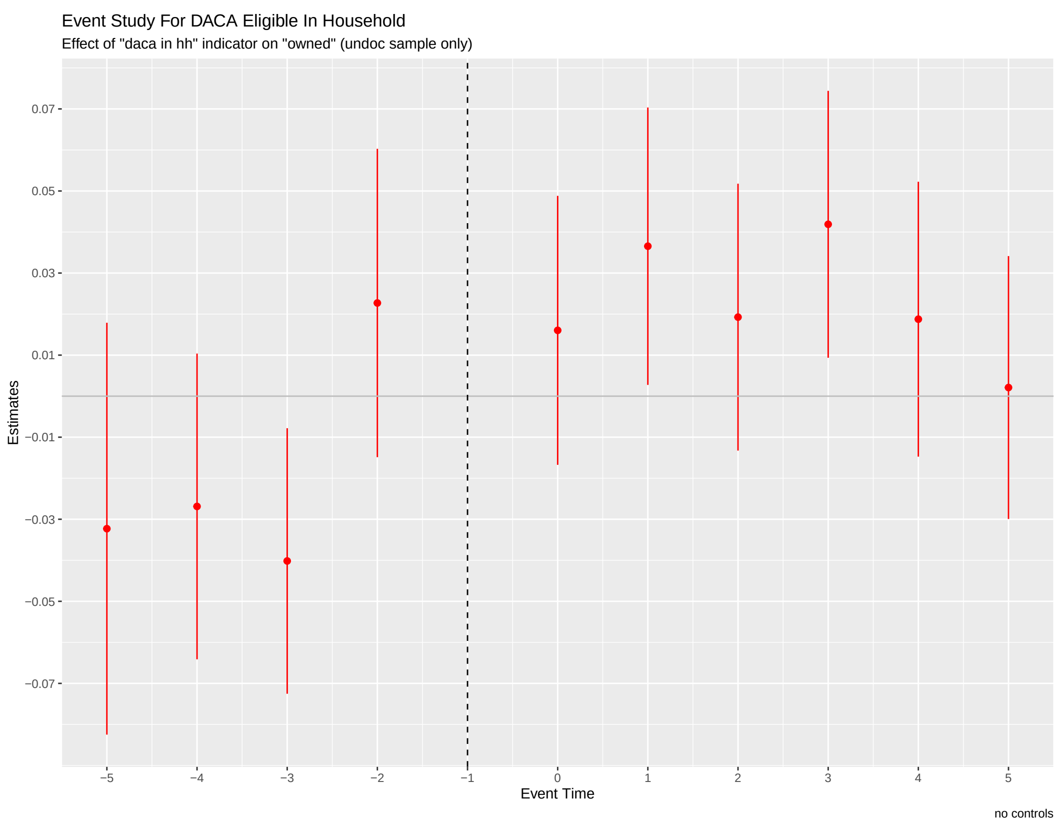

Figure II.1: Event study for the effect of having a DACA-eligible person living in the household (corresponds to the difference-in-differences results presented in column 4 of Table II.1)

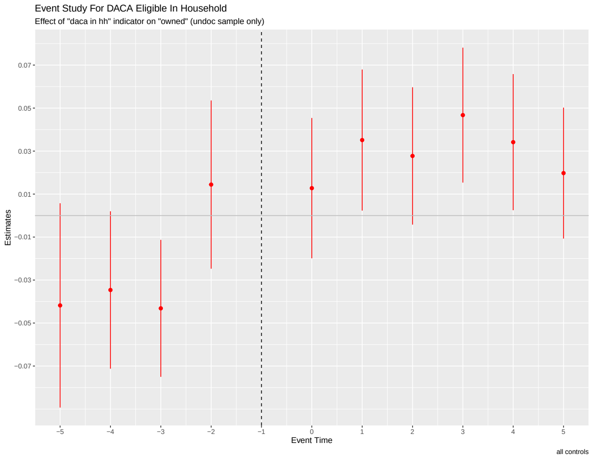

Figure II.2: Event study for the effect of having a DACA-eligible person living in the household (corresponds to the difference-in-differences results presented in column 5 of Table II.1)

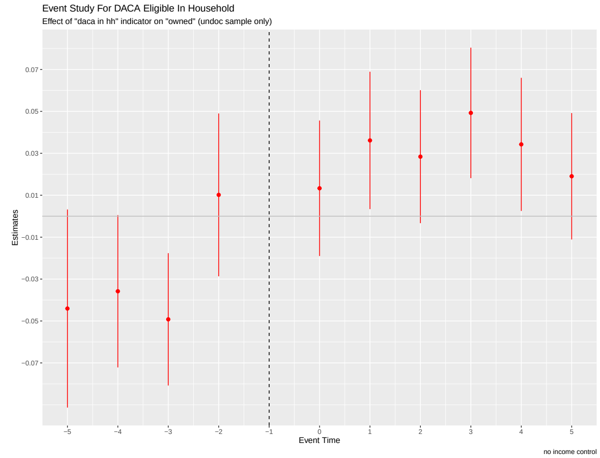

Figure II.3: Event study for the effect of having a DACA-eligible person living in the household (corresponds to the difference-in-differences results presented in column 6 of Table II.1)

Online Appendix III. Parallel Trends and Synthetic Control

Each difference-in-differences specification relies on the assumption of parallel trends. This section will assess each of the (in-text) county-level specifications in turn.[4] For each outcome, I present event study plots for transparency and to illustrate that, in most cases, there is little to no evidence of pre-trends that would bias the difference-in-differences estimates presented in the text. Then, because all of the analysis conducted at the county-level (Sections 4 and 5) relies on a panel of the same counties observed over time (as opposed to the household-level analysis, which is cross-sectional where individuals are observed only once), it is possible to produce estimates based on a synthetic control design. In the cases where the parallel trends assumption is unlikely to hold, estimates from synthetic control may be interpreted as more credible. In most cases, where there is little evidence of pre-trends, synthetic control estimates should closely resemble the difference-in-differences estimates and are therefore, presented for completeness and as tests of robustness to an alternative empirical strategy.[5]

I present four different p-values for the estimates throughout this section. They are defined in Galiani and Quistorff (2016). Where presented, one-sided p-values are computed as defined in Galiani and Quistorff (2016) and Abadie (2021). For further details on the procedures used, see Online Appendix IV.

Online Appendix III.1. Applications Outcome (DACA)

Refer to the estimates in Table 5. Figure 6 shows no evidence of pre-trends that may bias results, so synthetic control should produce estimates comparable to the difference-in-differences strategy. The synthetic control plots are presented below.[6]

[4]The trends assumptions for the analysis conducted at the household level have been assessed in other sections.

[5]Note that when there are two treatment categories (as in all of the county-level DACA analysis), synthetic control is run for the sample that excludes units in the "medium" category (i.e. synthetic control compares high DACA take-up units with the excluded category - low DACA take-up units).

[6]Note that one unit in the placebo group that is exceedingly difficult to match (due to its large baseline values of the outcome) is dropped from the placebo set of counties before the following plots and tables are generated. The 8 periods in which the unit is observed hold the top 8 spots in terms of magnitude of error. Therefore, it is matched poorly in both the pre-period and post-period and adds little meaningful information. If this unit is included, the synthetic placebo trend does not match the observed placebo trend as well in either period. However, even when included, p-values (and, of course, effect sizes) are practically identical.

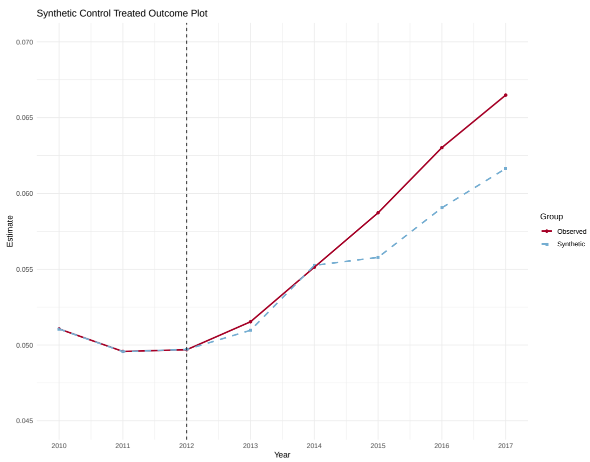

Figure III.1.1: Treated units

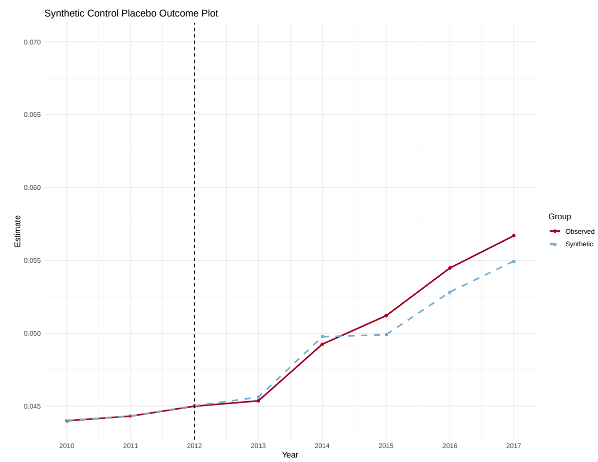

Figure III.1.2: Placebo units

Effect sizes and p-values are presented in the tables below.

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled
2013	0.0006	0.3233	0.0829		
2014	-0.0001	0.8909	0.4884		
2015	0.0029	0.0043	0.0325	0.8271	0.0002
2016	0.0040	0.0003	0.0061		
2017	0.0048	0.0000	0.0339		

Table III.1.1: unrestricted (pre-proportion = 1)

Note the surprisingly large p-value calculated using the post-period RMSPE. As noted by Galiani and Quistorff (2016), this might occur when some placebo units cannot be matched well (i.e. their pre-period RMSPE and post-period RMSPE are both large). Thus, when only considering the post-period RMSPE, these units would appear to be highly affected (even though, in reality, their deviations from their synthetic counterpart in the post-period are not much different from their deviations from the synthetic counterpart in the pre-period). Galiani and Quistorff (2016) recommend scaling p-values by the pre-period RMSPE (e.g. columns 4 and 6) as a solution.[7] An indicator of poor fit is a statistic that is, effectively, a p-value for the pre-period (i.e. it is computed identically to how "p-value joint post" is computed except that, instead of comparing observed values to synthetic values in the post-period, observed values are compared to synthetic values in the pre-period over which the data is trained). I will refer to this as the "pre-proportion" (as it is the proportion

[7]In other words, columns 4 and 6 are measurements of the size of deviations in the post-period(s) relative to the size of deviations in the pre-period. Columns 3 and 5 simply measure the size of deviations in the post-period(s), which is an adequate measure when the synthetic control procedure is able to produce trends that fit similarly well for both treated units and placebo units.

of random placebo samples that generate a pre-period RMSPE larger than the treated average pre-period RMSPE). An extreme value (i.e. close to 0 or close to 1) is an indicator that the synthetic control procedure performed much better for one group (treated when close to 1, placebo when close to 0) than the other. Therefore, another remedy to this problem of poor fit in the placebo group, as suggested by Galiani and Quistorff (2016) and Abadie, Diamond and Hainmueller (2010), is to restrict the placebo set of units to those which have a pre-period RMSPE no more than m times the average treated pre-period RMSPE. If the large "p-value joint post" is merely an artifact of including placebo units that are generally matched poorly by the synthetic control procedure, then imposing such a restriction will reduce the p-value.[8] Therefore, in addition to tables where p-values are constructed absent any sample restrictions on the quality of pre-period fit, I will include a few tables where p-values are re-computed under different restrictions (different values of m).[9]

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled
2013	0.0006	0.2260	0.0958		
2014	-0.0001	0.9449	0.5054		
2015	0.0029	0.0003	0.0484	0.1110	0.0012
2016	0.0040	0.0000	0.0115		
2017	0.0048	0.0000	0.0523		

Table III.1.2: $m = 100$ restriction (pre-proportion = 0.96)

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled
2013	0.0006	0.2456	0.0996		
2014	-0.0001	0.9599	0.5094		
2015	0.0029	0.0002	0.0531	0.0740	0.0016
2016	0.0040	0.0000	0.0136		
2017	0.0048	0.0000	0.0582		

Table III.1.3: $m = 75$ restriction (pre-proportion = 0.35)

Consistent with findings from event studies and difference-in-differences estimates, synthetic control detects a positive effect that is greater (in magnitude and significance) after two years have passed (i.e. the "adjustment period"). Weighting each post-period year equally, the joint post-period estimated effect is a 0.244 percentage point increase in the relative

[8]This is based on the assumption that the units driving the large p-value vary largely in the post-period for the same reason they vary largely in the pre-period (poor fit). If the units driving the large p-value only match poorly in the post-period, this may be indicative of an actual "effect" or unaccounted for trend. Because the restriction applies only to units with poor pre-period fit, such units would (appropriately) remain in the sample even under this restriction.

[9]Arguably, the comparison is most "fair" when the pre-proportion is close to 0.5.

number of Hispanic home loan applications, which is close to the unweighted difference-in-differences estimate of 0.33 percentage points.

Online Appendix III.2. Approvals Outcome (DACA)

Refer to the estimates in Table 9. Presented below is the event study corresponding to column 4.

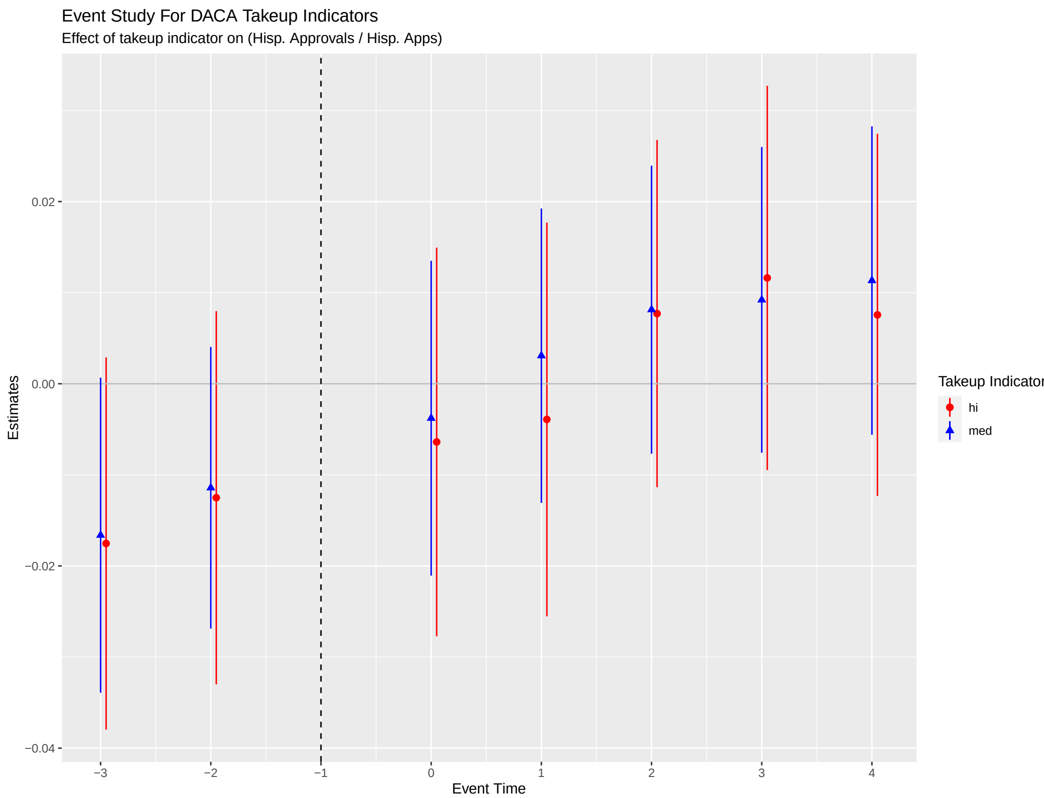

Figure III.2.1: Event study for Hispanic approval rate

Though pre-period estimates are not significantly below zero and the first two post-period estimates remain below zero, the points do appear to exhibit an upward trend, which would bias difference-in-differences estimates away from zero. If there is a meaningful pre-trend, one might find estimates from synthetic control to be more credible. While the resulting changes are not substantial, for the purpose of match accuracy, I impose that all counties must have at least 10 Hispanic home loan applications (the denominator of the outcome) in every year to be included in the sample. Prior to running synthetic control, any county with fewer than 10 Hispanic home loan applications in any year is dropped. The synthetic and observed trends are presented below.

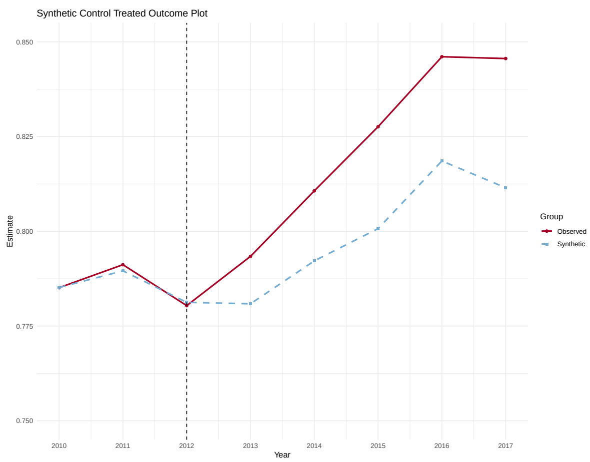

Figure III.2.2: Treated units

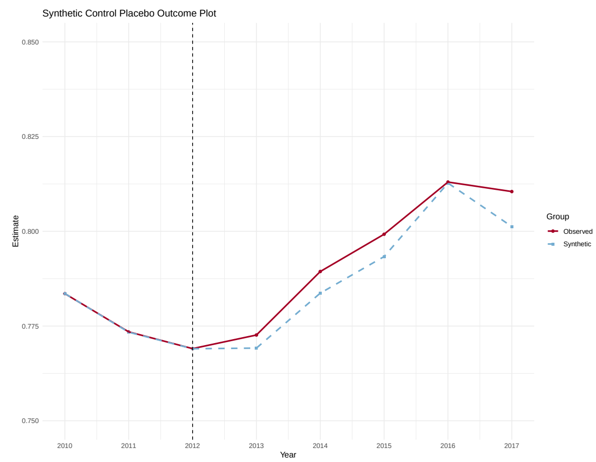

Figure III.2.3: Placebo units

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled
2013	0.0125	0.0206	0.0005		
2014	0.0185	0.0006	0.1645		
2015	0.0269	0.0000	0.0000	1.0000	0.8336
2016	0.0275	0.0000	0.0000		
2017	0.0341	0.0000	0.0000		

Table III.2.1: unrestricted (pre-proportion = 0.88)

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled
2013	0.0125	0.0219	0.0006		
2014	0.0185	0.0003	0.1698		
2015	0.0269	0.0000	0.0000	1.0000	0.8509
2016	0.0275	0.0000	0.0000		
2017	0.0341	0.0000	0.0000		

Table III.2.2: $m = 50$ restriction (pre-proportion = 0.54)

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled
2013	0.0125	0.0212	0.0007		
2014	0.0185	0.0005	0.1841		
2015	0.0269	0.0000	0.0000	1.0000	0.8977
2016	0.0275	0.0000	0.0000		
2017	0.0341	0.0000	0.0000		

Table III.2.3: $m = 25$ restriction (pre-proportion = 0.005)

Even when restrictions are imposed to reduce the pre-proportion, the p-values based on post-period RMSPE are exceedingly large even though p-values for individual periods are more reasonable and even indicate significance in most cases. This is likely the result of over-fitting. The first two sets of p-values (columns 3 and 4) are derived from comparing average "effects" (observed value - synthetic value) in the treated group with average, randomly sampled placebo effects. Over-fitting would result in synthetic values very close to average observed values in the placebo group. However, the deviations of any single unit from the synthetic prediction may be wild (e.g. placebo unit A's estimate is far below the synthetic, but placebo unit B's estimate is far above the synthetic to compensate). Then, the calculated average in a given period will likely be close to the synthetic prediction, but because RMSPE is calculated using a sum of *squared* deviations, it may still be large in the case of over-fitting. In this case, one-sided inference may prove more informative. The two-sided testing so far has tested against the null that (placebo) values (mean differences between observed and synthetic values or post-period RMSPE) are at least as extreme as the average of the values in the treated group. In other words, in two-sided inference, a comparison is made between the absolute value of mean differences (or between post-period RMSPE values, which, by construction, are non-negative). Galiani and Quistorff (2016) provide a method for one-sided inference for the p-values presented in columns 3 and 4. Abadie (2021) provide a method for one-sided inference for the p-values in columns 5 and 6.[10] The following tables present the p-values from one-sided inference.

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled	one-sided
2013	0.0125	0.0211	0.0005			
2014	0.0185	0.0006	0.1643			
2015	0.0269	0.0000	0.0000	0.1101	0.0159	positive
2016	0.0275	0.0000	0.0000			
2017	0.0341	0.0000	0.0000			

Table III.2.4: unrestricted (pre-proportion = 0.88)

[10] Adapting the two-sided testing to one-sided is not difficult. For the p-values in columns 3 and 4, rather than comparing absolute values of mean differences, the direction of the difference is taken into account. For the others, when computing the sum of squared deviations for calculating the post-period RMSPE, accept only positive (or only negative) deviations, treating all other deviations as zero. In other words, disallowing the possibility of a negative effect, any observed difference that takes a value less than zero must be evidence of a zero effect.

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled	one-sided
2013	0.0125	0.0218	0.0005			
2014	0.0185	0.0003	0.1697			
2015	0.0269	0.0000	0.0000	0.0756	0.0183	positive
2016	0.0275	0.0000	0.0000			
2017	0.0341	0.0000	0.0000			

Table III.2.5: $m = 50$ restriction (pre-proportion = 0.54)

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled	one-sided
2013	0.0125	0.0209	0.0007			
2014	0.0185	0.0005	0.1844			
2015	0.0269	0.0000	0.0000	0.0533	0.0283	positive
2016	0.0275	0.0000	0.0000			
2017	0.0341	0.0000	0.0000			

Table III.2.6: $m = 25$ restriction (pre-proportion = 0.004)

Estimates from the synthetic control empirical strategy support the in-text, difference-in-differences results. If anything, estimated effects are larger under synthetic control. The first two types of p-values both indicate statistical significance at (at least) the 95% confidence level in all periods but one. The p-values based on post-period RMSPE calculations are extremely high (indicating insignificance) under two-sided inference, but one-sided inference yields values that indicate statistical significance at the 95% confidence level when accounting for pre-period RMSPE (column 6) and at the 90% or 85% significance level (depending on the restriction imposed) when pre-period RMSPE is not taken into account (column 5).

Online Appendix III.3. Loan Amount (Applications) Outcome (DACA)

Refer to the estimates in Table 10. Presented below is the event study corresponding to column 4.

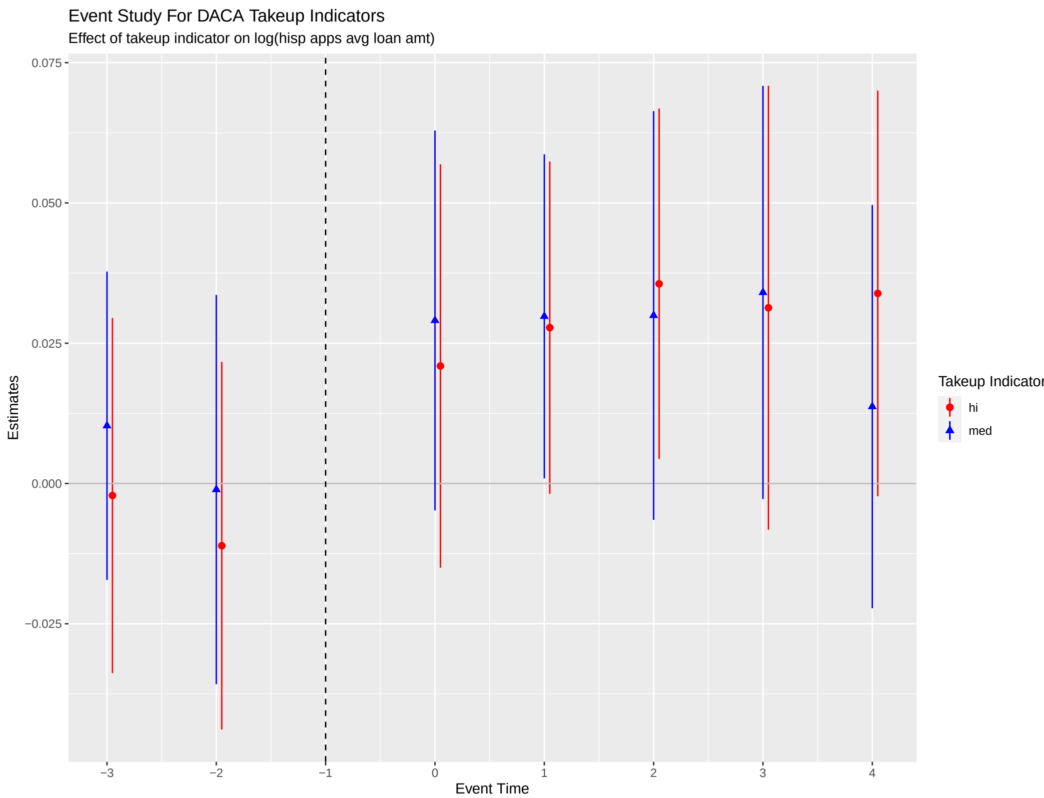

Figure III.3.1: Event study for log loan amount of Hispanic home loan applications

There is no noticeable pre-trend, so synthetic control estimates should be close to the difference-in-differences estimates in-text. Plots, estimated effects, and p-values are presented below.

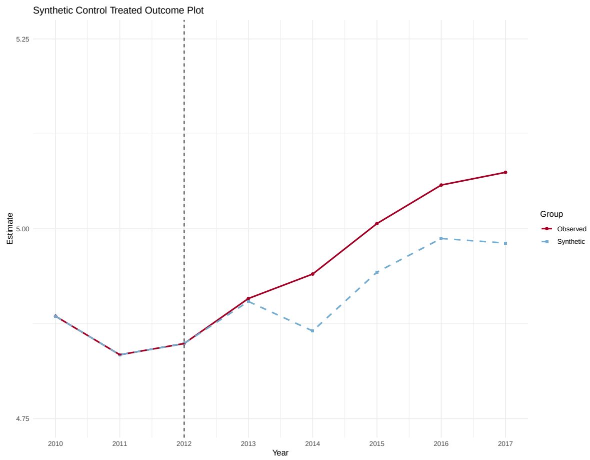

Figure III.3.2: Treated units

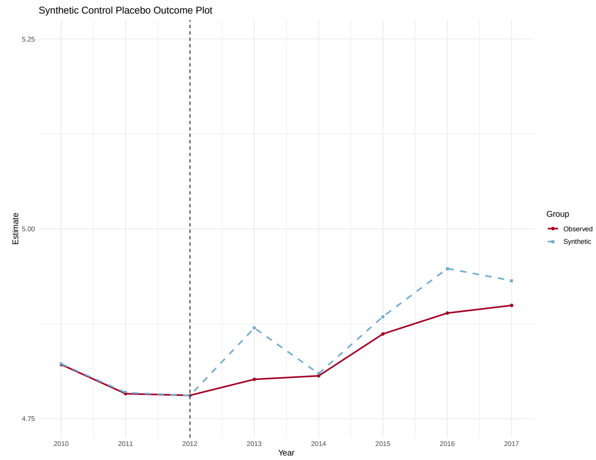

Figure III.3.3: Placebo units

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled
2013	0.0039	1.0000	1.0000		
2014	0.0749	0.0000	0.0004		
2015	0.0640	0.0000	0.0117	1.0000	0.9999
2016	0.0703	0.1121	0.2585		
2017	0.0932	0.0000	0.0000		

Table III.3.1: unrestricted (pre-proportion = 1)

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled
2013	0.0039	1.0000	1.0000		
2014	0.0749	0.0000	0.0006		
2015	0.0640	0.0000	0.0172	1.0000	1.0000
2016	0.0703	0.0889	0.3314		
2017	0.0932	0.0000	0.0000		

Table III.3.2: $m = 50$ restriction (pre-proportion = 0.49)

As in the previous section, the synthetic control appears to suffer from an issue of over-fitting. As before, tables with p-values for one-sided inference are produced.

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled	one-sided
2013	0.0039	0.0000	0.0000			
2014	0.0749	0.0000	0.0004			
2015	0.0640	0.0000	0.0000	0.0000	0.0056	positive
2016	0.0703	0.0000	0.0000			
2017	0.0932	0.0000	0.0000			

Table III.3.3: unrestricted (pre-proportion = 1)

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled	one-sided
2013	0.0039	0.0000	0.0000			
2014	0.0749	0.0000	0.0006			
2015	0.0640	0.0000	0.0000	0.0000	0.0162	positive
2016	0.0703	0.0000	0.0000			
2017	0.0932	0.0000	0.0000			

Table III.3.4: $m = 50$ restriction (pre-proportion = 0.5)

Results are, again, in line with the results from the difference-in-differences specifications. If anything, estimated effects are larger under synthetic control. Under two-sided inference, the first two versions of p-values indicate statistical significance at the 99% confidence level in 2014 and 2017 and at the 95% confidence level in 2015. Under one-sided inference, all p-values under all restrictions except one indicate significance at the 99% confidence level (the exception indicates significance at the 95% level).

Online Appendix III.4. Loan Amount (Approvals) Outcome (DACA)

Refer to the estimates in Table 11. Presented below is the event study corresponding to column 4.

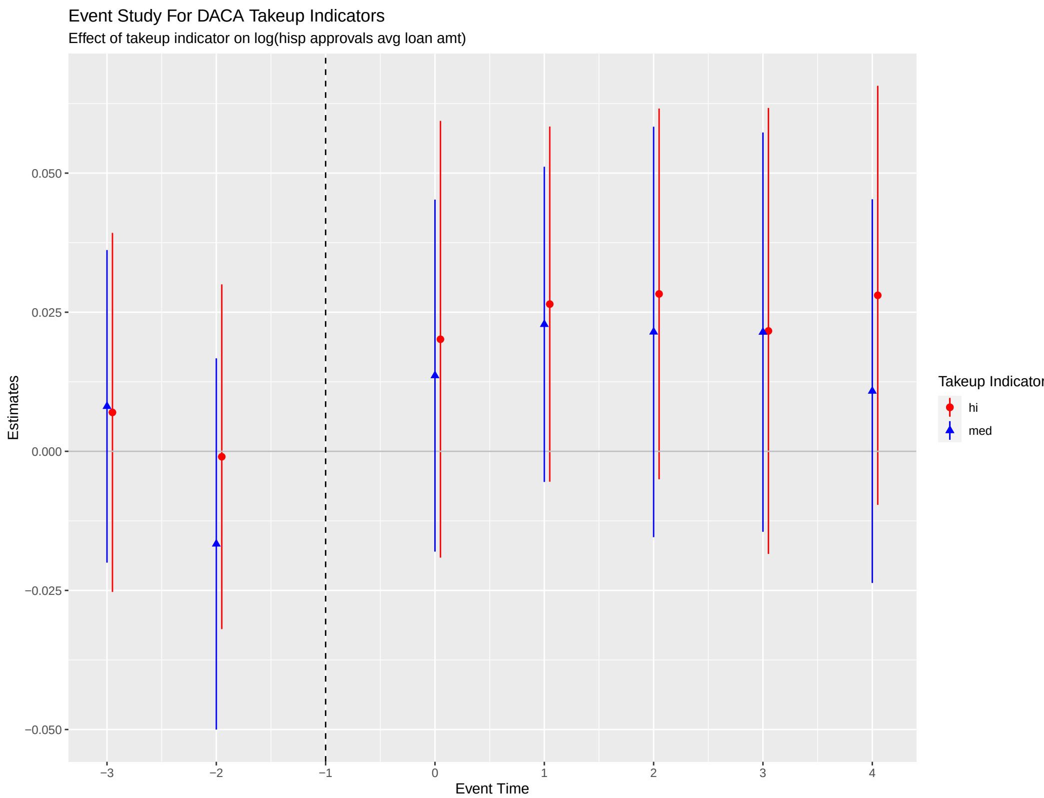

Figure III.4.1: Event study for log loan amount of Hispanic home loan applications that were approved

There is no noticeable pre-trend, so synthetic control estimates should be close to the difference-in-differences estimates in-text. Plots, estimated effects, and p-values are presented below.

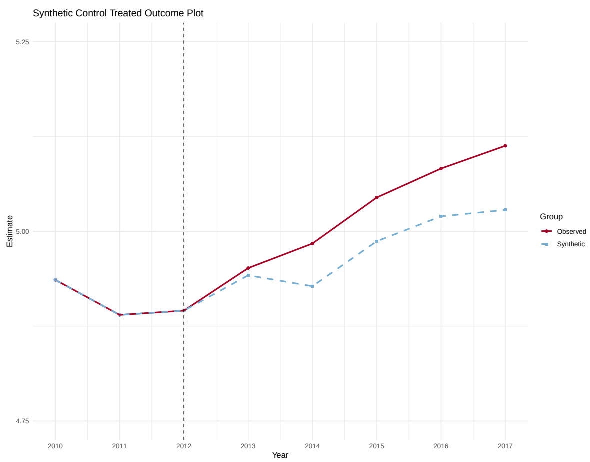

Figure III.4.2: Treated units

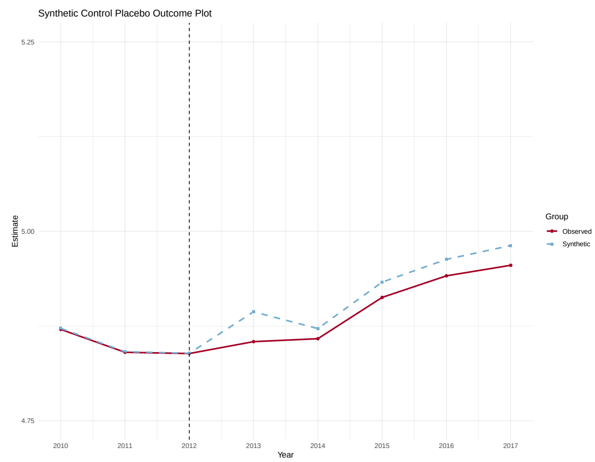

Figure III.4.3: Placebo units

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled
2013	0.0095	0.9968	0.9989		
2014	0.0563	0.0000	0.0145		
2015	0.0577	0.0000	0.0011	1.0000	0.9999
2016	0.0630	0.0000	0.0000		
2017	0.0845	0.0000	0.0000		

Table III.4.1: unrestricted (pre-proportion = 1)

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled
2013	0.0095	0.9981	0.9990		
2014	0.0563	0.0000	0.0212		
2015	0.0577	0.0000	0.0022	1.0000	1.0000
2016	0.0630	0.0000	0.0001		
2017	0.0845	0.0000	0.0000		

Table III.4.2: $m = 50$ restriction (pre-proportion = 0.06)

As in the previous two sections, the synthetic control appears to suffer from an issue of over-fitting. As before, tables with p-values for one-sided inference are produced.

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled	one-sided
2013	0.0095	0.0000	0.0010			
2014	0.0563	0.0000	0.0000			
2015	0.0577	0.0000	0.0000	0.0001	0.0001	positive
2016	0.0630	0.0000	0.0000			
2017	0.0845	0.0000	0.0000			

Table III.4.3: unrestricted (pre-proportion = 1)

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled	one-sided
2013	0.0095	0.0000	0.0010			
2014	0.0563	0.0000	0.0000			
2015	0.0577	0.0000	0.0000	0.0000	0.0011	positive
2016	0.0630	0.0000	0.0000			
2017	0.0845	0.0000	0.0000			

Table III.4.4: $m = 50$ restriction (pre-proportion = 0.06)

Results are, again, in line with the results from the difference-in-differences specifications. If anything, estimated effects are larger under synthetic control. Under two-sided inference, the first two versions of p-values indicate statistical significance at the 99% confidence level in 2015, 2016, and 2017 and at the 95% confidence level in 2014. Under one-sided inference, all p-values under all restrictions indicate significance at the 99% confidence level.

Online Appendix III.5. Applications Outcome (Treasury)

Refer to the estimates in Table 12. The event study corresponding to column 1 is presented in Figure 8. The pre-trends suggest that difference-in-differences estimates are likely to be positively biased. Therefore, an effective synthetic control strategy that does not suffer such bias would be expected to yield smaller estimated effects. Plots, estimated effects, and p-values are presented below.

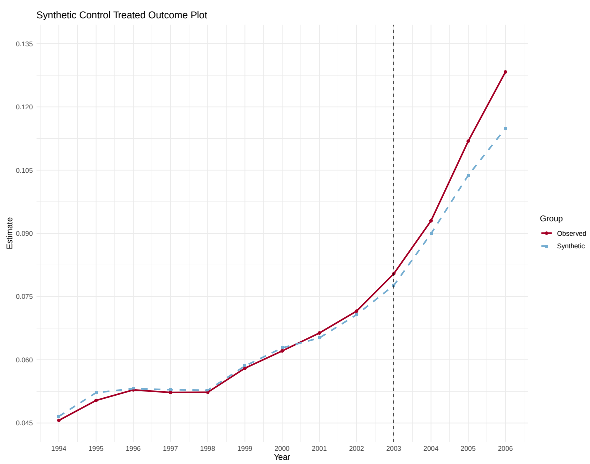

Figure III.5.1: Treated units

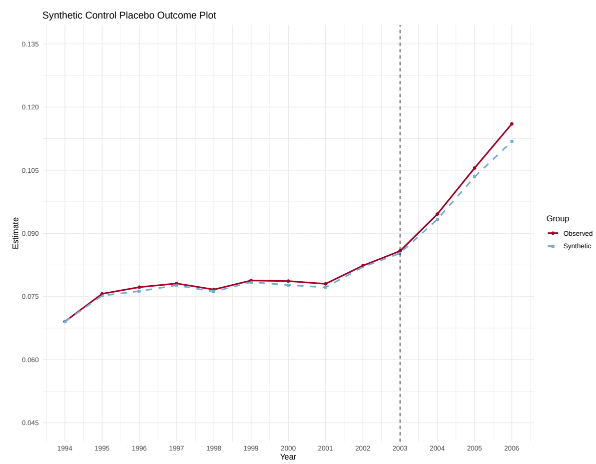

Figure III.5.2: Placebo units

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled
2004	0.0031	0.0204	0.0004		
2005	0.0081	0.0000	0.0000	0.0000	0.0017
2006	0.0134	0.0000	0.0000		

Table III.5.1: unrestricted (pre-proportion = 0.58)

The effects are consistent with expectations. All estimates are positive and significant at (at least) the 95% confidence level, and consistent with the idea that difference-in-differences estimates are upwards biased due to trends, the synthetic control estimates are smaller in magnitude. Thus, the synthetic control estimated effect of a 0.95 percentage point effect on the Hispanic home loan application rate should be considered more accurate than the 1.34 percentage point change indicated by the (biased) difference-in-differences results.

Online Appendix III.6. Approvals Outcome (Treasury)

Refer to the estimates in Table 13. The event study corresponding to column 4 is presented below.

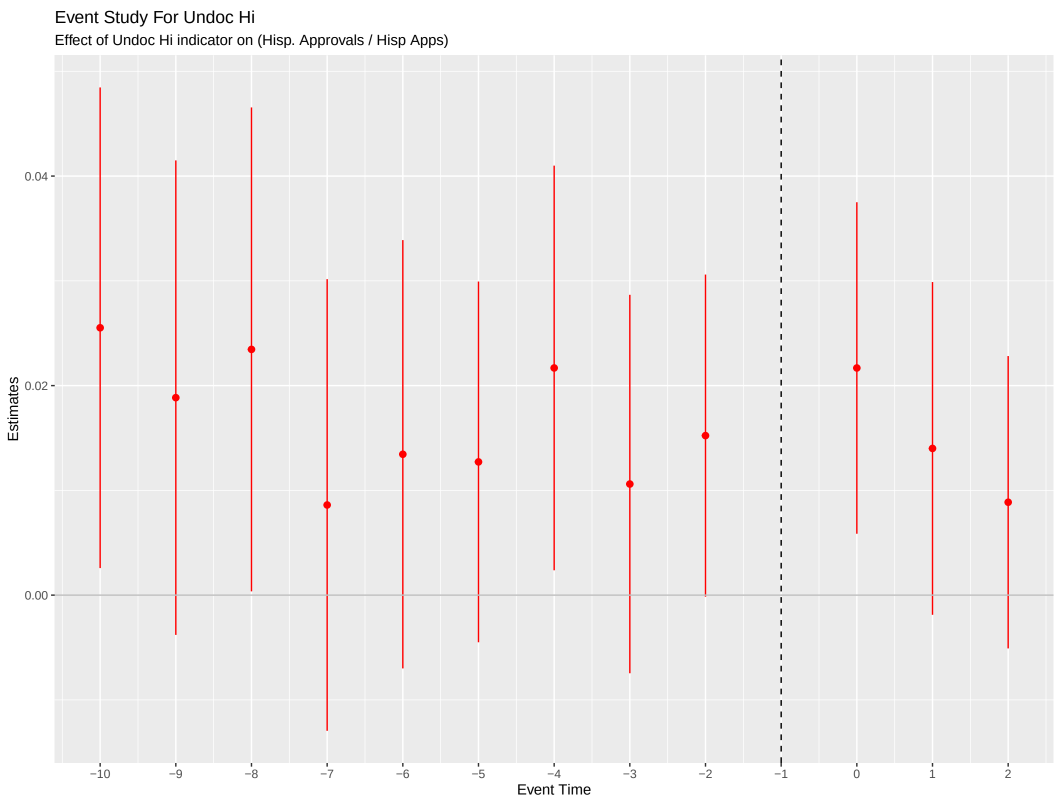

Figure III.6.1: Event study for log loan amount of Hispanic home loan applications

There are no apparent pre-trends, so we should expect synthetic control to yield estimates similar to those produced by the in-text difference-in-differences strategy. Plots, estimated effects, and p-values are presented below.

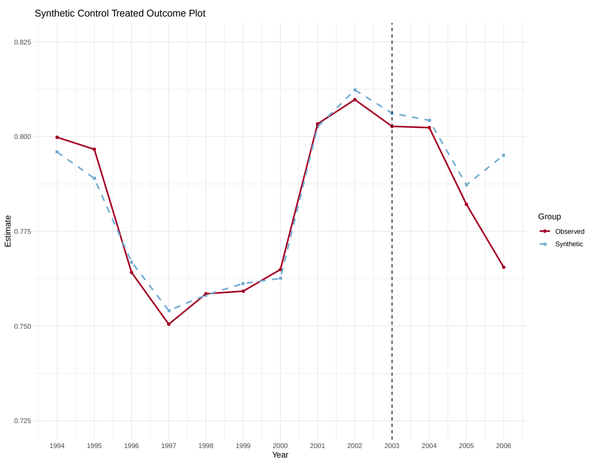

Figure III.6.2: Treated units

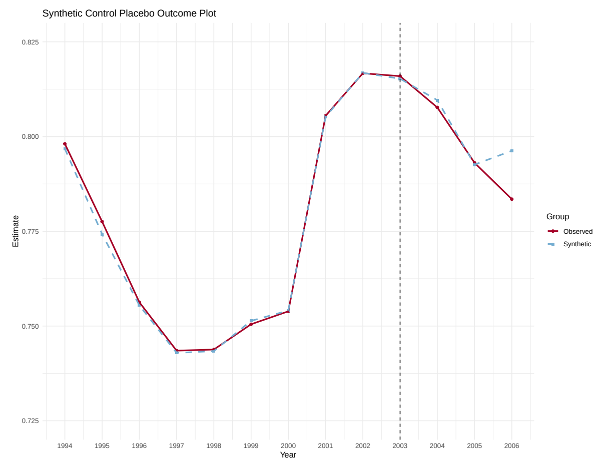

Figure III.6.3: Placebo units

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled
2004	-0.0019	0.6329	0.6797		
2005	-0.0052	0.2284	0.5428	0.5124	0.2997
2006	-0.0296	0.0001	0.1962		

Table III.6.1: unrestricted (pre-proportion = 0.99)

Because the pre-proportion value is near 1, I test to see if imposing a restriction on the placebo set meaningfully changes the p-values.

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled
2004	-0.0019	0.6207	0.7193		
2005	-0.0052	0.3096	0.6070	0.4858	0.4012
2006	-0.0296	0.0005	0.2755		

Table III.6.2: $m = 4$ restriction (pre-proportion = 0.47)

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled
2004	-0.0019	0.7202	0.8065		
2005	-0.0052	0.2595	0.7515	0.3044	0.6773
2006	-0.0296	0.0118	0.5260		

Table III.6.3: $m = 2$ restriction (pre-proportion = 0)

Results are similar regardless of pre-proportion value. Joint p-values are not extreme either,[11] suggesting over-fitting is not an issue in the way it was with the synthetic control for DACA's effect on some outcomes. Estimates are negative in all years and somewhat

[11] Additionally, p-values in columns 5 and 6 are consistent with those in columns 3 and 4.

larger in magnitude than the difference-in-differences estimates. The first p-value (column 3) indicates that the effect in 2006 is statistically significant at conventional levels. However, once pre-period fit is accounted for, the significance is lost. The estimates are insignificant in all other periods, and p-values for the joint effect across all post-period years (columns 5 and 6) indicate statistical insignificance, as well. Thus, the results are consistent with the results from the difference-in-differences specifications where point estimates were negative but statistically insignificant.

Online Appendix III.7. Loan Amount (Applications) Outcome (Treasury)

Refer to the estimates in Table 14. The event study corresponding to column 4 is presented below.

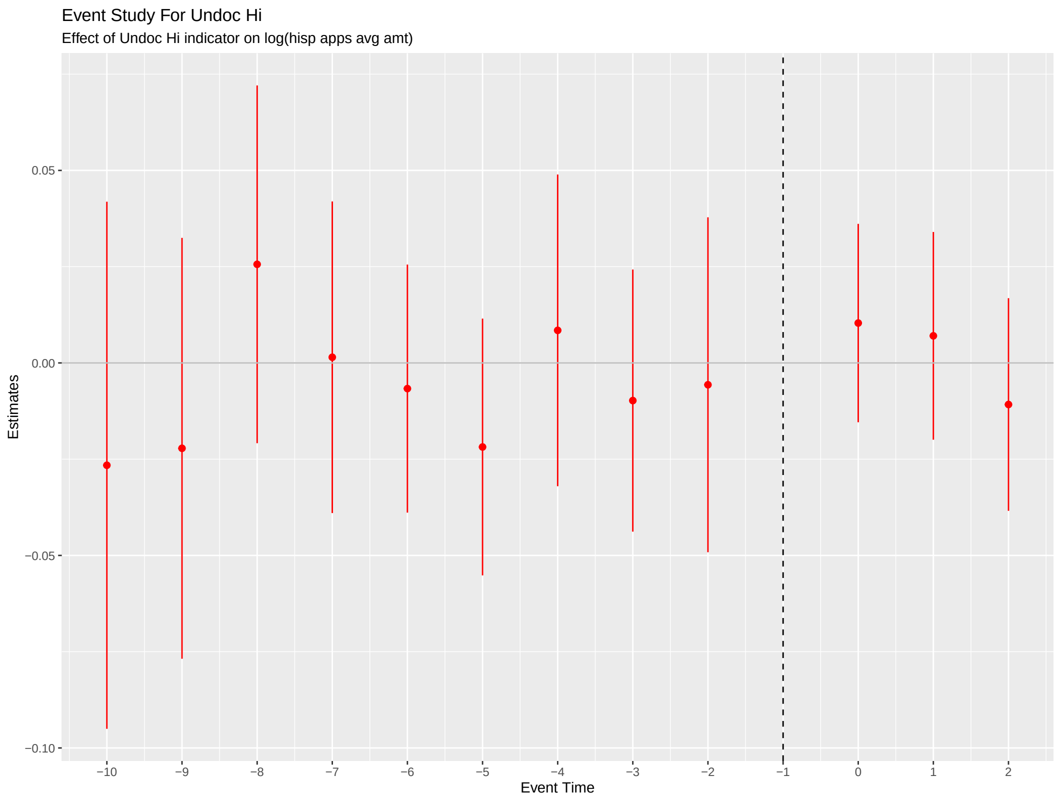

Figure III.7.1: Event study for log loan amount of Hispanic home loan applications

There are no apparent pre-trends, so we should expect synthetic control to yield estimates similar to those produced by the in-text difference-in-differences strategy. Plots, estimated effects, and p-values are presented below.

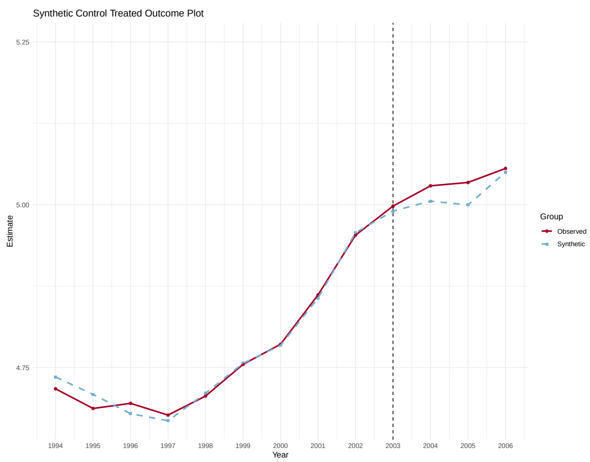

Figure III.7.2: Treated units

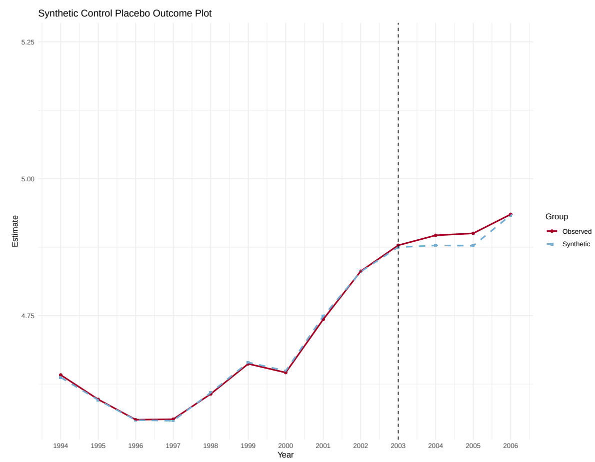

Figure III.7.3: Placebo units

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled
2004	0.0235	0.3076	0.8325		
2005	0.0345	0.1369	0.9380	0.4043	0.8638
2006	0.0058	0.6389	0.6634		

Table III.7.1: unrestricted (pre-proportion = 0.998)

Because the pre-proportion value is near 1, I test to see if imposing a restriction on the placebo set meaningfully changes the p-values.

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled
2004	0.0235	0.3163	0.8376		
2005	0.0345	0.2096	0.9403	0.1526	0.8982
2006	0.0058	0.6437	0.6746		

Table III.7.2: $m = 10$ restriction (pre-proportion = 0.82)

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled
2004	0.0235	0.4494	0.8471		
2005	0.0345	0.5798	0.9433	0.0623	0.9426
2006	0.0058	0.7224	0.6946		

Table III.7.3: $m = 5$ restriction (pre-proportion = 0.02)

Since the "p-value joint post" indicates significance under certain restrictions but the other p-values do not similarly indicate statistical significance, it is worth checking whether this significance remains under one-sided inference (which, if there are real positive effects, should produce p-values that are indicative of significance at even greater levels of confidence).

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled	one-sided
2004	0.0235	0.3077	0.8329			
2005	0.0345	0.1368	0.9051	0.2913	0.9880	positive
2006	0.0058	0.3631	0.9877			

Table III.7.4: unrestricted (pre-proportion = 0.998)

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled	one-sided
2004	0.0235	0.3172	0.8311			
2005	0.0345	0.2095	0.9074	0.1814	0.9920	positive
2006	0.0058	0.4068	0.9864			

Table III.7.5: $m = 10$ restriction (pre-proportion = 0.82)

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled	one-sided
2004	0.0235	0.4497	0.8264			
2005	0.0345	0.5805	0.9092	0.2967	0.9960	positive
2006	0.0058	0.6118	0.9847			

Table III.7.6: $m = 5$ restriction (pre-proportion = 0.02)

The absence of statistical significance when p-values for one-sided inference are considered is evidence that the marginal statistical significance detected under two-sided inference is the result of poor fit, not a true effect.[12]

The evidence from synthetic control is broadly consistent with the difference-in-differences results and the accompanying event study. Estimates are positive in direction, which is consistent with the comparable difference-in-differences specification (where California is included and weights are not applied), and nearly all p-values, including all p-values for one-sided inference, indicate that the estimated effects are statistically indistinguishable from zero.

[12]Contrast this with the one-sided testing for effects of DACA where switching to one-sided inference drastically increased statistical significance.

Online Appendix III.8. Loan Amount (Approvals) Outcome (Treasury)

Refer to the estimates in Table 15. The event study corresponding to column 4 is presented below.

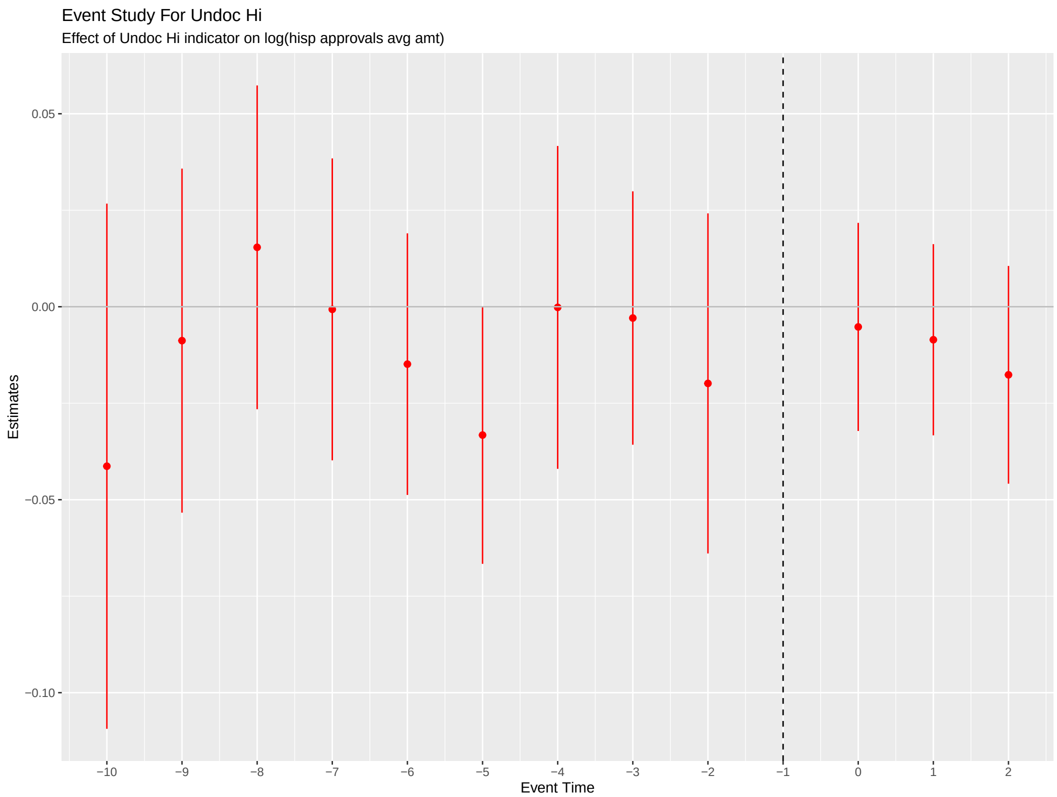

Figure III.8.1: Event study for log loan amount of Hispanic home loan approvals

There are no apparent pre-trends, so we should expect synthetic control to yield estimates similar to those produced by the in-text difference-in-differences strategy. Plots, estimated effects, and p-values are presented below.

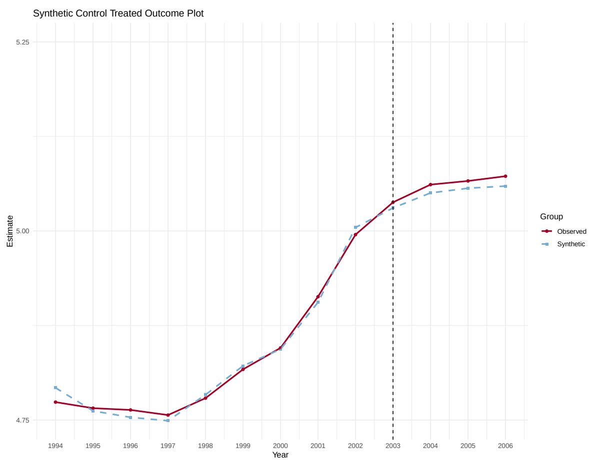

Figure III.8.2: Treated units

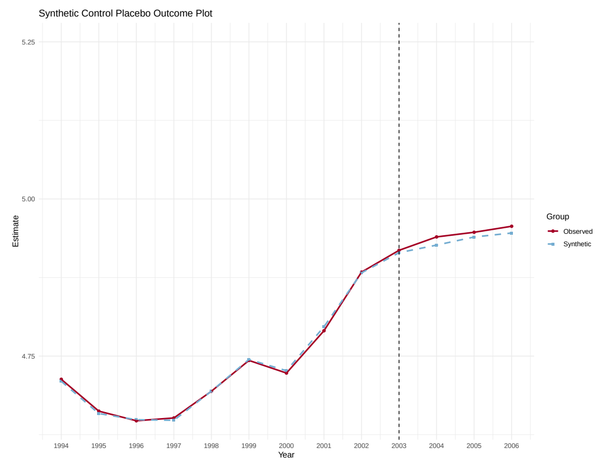

Figure III.8.3: Placebo units

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled
2004	0.0110	0.5913	0.9816		
2005	0.0096	0.4876	0.9256	0.5230	0.9546
2006	0.0131	0.4729	0.7001		

Table III.8.1: unrestricted (pre-proportion = 0.99)

Because the pre-proportion value is near 1, I test to see if imposing a restriction on the placebo set meaningfully changes the p-values.

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled
2004	0.0110	0.5721	0.9816		
2005	0.0096	0.5306	0.9279	0.3424	0.9686
2006	0.0131	0.4894	0.7094		

Table III.8.2: $m = 10$ restriction (pre-proportion = 0.71)

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled
2004	0.0110	0.5984	0.9831		
2005	0.0096	0.6148	0.9316	0.3047	0.9853
2006	0.0131	0.4784	0.7266		

Table III.8.3: $m = 5$ restriction (pre-proportion = 0.01)

Unlike the results for changes in average loan amounts among loan applications, changes in the size of approved loans are statistically insignificant even in the most restrictive case under two-sided inference. Thus, computing p-values under one-sided inference isn't as informative, but for completeness, I present them below, anyway.

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled	one-sided
2004	0.0110	0.5758	0.9782			
2005	0.0096	0.4253	0.8994	0.5618	0.9750	positive
2006	0.0131	0.4334	0.6900			

Table III.8.4: unrestricted (pre-proportion = 0.99)

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled	one-sided
2004	0.0110	0.5588	0.9781			
2005	0.0096	0.4808	0.9010	0.4835	0.9813	positive
2006	0.0131	0.4548	0.6991			

Table III.8.5: $m = 10$ restriction (pre-proportion = 0.71)

Year	Effect	p-value	p-value scaled	p-value joint post	p-value joint post scaled	one-sided
2004	0.0110	0.5864	0.9777			
2005	0.0096	0.5855	0.9039	0.4855	0.9897	positive
2006	0.0131	0.4366	0.7147			

Table III.8.6: $m = 5$ restriction (pre-proportion = 0.01)

Estimates from synthetic control are again, somewhat larger in magnitude, but like the estimates from the difference-in-differences specifications, they are statistically indistinguishable from zero.

Thus, all synthetic control estimates are consistent with their corresponding difference-in-differences estimates when the parallel trends assumption appears to hold.

Online Appendix IV. Details of Synthetic Control Procedures

Synthetic control is carried out using the synth package[13] in R (a subset of results were validated using the Stata version of the synth package). In all cases, predictors are the pre-period observations of the dependent variable. Choice of predictor weights is data-driven. Weights are chosen by an optimization algorithm that minimizes mean squared prediction error (MSPE) over all pre-treatment periods (the optimization algorithm used is the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm, which, in general, produced better synthetic trends than alternatives such as the Nelder-Mead, albeit at the cost of computation speed). Additional details about the optimization procedure are available upon request.

P-values are generated as suggested by Galiani and Quistorff (2016) in the case of multiple treated units. In all cases, the size of the full set of placebo averages exceeds 10 to the hundredth power. For this reason, as suggested by Galiani and Quistorff (2016), random samples of 1,000,000 are selected in the computations of all p-values.

[13] Abadie, Diamond and Hainmueller (2011)

Online Appendix V. Details on the Treasury Legal Clarification

A handful of media articles[14] have published that the rules implemented by the Treasury Department in 2003 allowed customers to set up bank accounts using ITIN's in place of Social Security numbers. Other sources[15] claim that (in or around) 2003 was when banks and credit unions first began offering mortgages to undocumented immigrants. These claims are *close* to the truth. In this section, I will elaborate on some of the relevant details that led to the massive spike in ITIN loans circa 2003.

In 2003, rules proposed by the PATRIOT Act's section on "Customer Identification Programs for banks, savings associations, credit unions, and certain non-federally regulated banks" were implemented. These new rules mark the first instance that Treasury Department policy formally listed the ITIN as an acceptable form of identification for the purpose of establishing bank accounts. Prior to the new rules, identifying information to be collected was regulated by the Bank Secrecy Act (BSA), which had not been updated since prior to 1996, when ITIN's were created. The BSA listed, more broadly, that institutions needed to secure a tax identification number as defined by IRS code 6109 of 1954. This IRS code states that it "shall determine what constitutes a taxpayer ID number..." However, the code is vague and only explicitly mentions Social Security numbers, employer identification numbers, or "an alternative identification number for purposes of identifying themselves." The BSA also stated that, for non-resident aliens, institutions also needed to retain a passport number or "a description of some other government document used to verify his identity."

This seems to leave room for institutions to justify offering ITIN loans if they are confident in their interpretation of existing Treasury rules. However, in 2002, the Treasury Department issued a statement that said, in part, "... because ITINs are issued without rigorous verification, financial institutions must avoid relying on the ITIN to verify the identity of a foreign national." Thus, at best, the rules on establishing bank accounts using an ITIN were ambiguous. At worst, they barred the use of ITIN's as acceptable identification for the establishment of bank accounts.

The ambiguity of the rules made the issuance of ITIN loans rare prior to 2003, though there is record of some smaller institutions reportedly offering such loans as early as the late 90's. In 2003, the Treasury Rules in the PATRIOT Act rendered parts of the Bank Secrecy Act obsolete and explicitly listed ITIN's as acceptable forms of ID for establishing bank accounts. Beginning in 2004, there are reports of organizations and financial entities beginning to engage with ITIN loans on a large scale.[16] In 2004, Suspicious Activity Reports

[14] See, for instance, Khimm (2014) and Roosevelt (2017).

[15] See, for example, Jordan (2008).

[16] For example, banks associated with the New Alliance Task Force, which is argued to have "pioneered"

for borrowers with ITIN's spiked, which may be a result of institutions reacting to the new stringency of the PATRIOT Act rules, but an alternative explanation would be that there simply were not many borrowers using ITIN's prior to 2004, following the Treasury's legal clarification.

A publication by the Chicago Fed's Consumer and Community Affairs Division in 2005 (Gallagher, 2005) reported that, as of September of 2004, there were 18 banks and 1 credit union accepting ITIN's for mortgage underwriting, including TCF Bank and Fifth Third Bank. It is also reported that "[t]he regulatory community cites language in Section 326 of the PATRIOT Act in explaining" that an ITIN is an acceptable form of ID. Finally, and perhaps most importantly, in 2004, Citibank (one of "the big 4") started issuing ITIN mortgages.[17]

In summary, there was some ITIN mortgage activity prior to 2003, but it appears to have been rare and legally ambiguous, at best. In 2003, through changes brought on by the PATRIOT Act, the Treasury Department amended the Bank Secrecy Act's rules to explicitly allow for the use of ITIN's as acceptable identification for the opening of bank accounts. An "explosion" of ITIN usage in banking followed, including Citigroup's decision to offer ITIN mortgages the next year.

the creation of ITIN mortgage products for individuals lacking Social Security numbers in 2003, reportedly used alternative forms of ID to open more than 50,000 new accounts for Latin American Immigrants in 2004. In January of 2004, Mortgage Guaranty Insurance Corporation became the first company to insure ITIN loans. In April of 2004, the Wisconsin Housing and Economic Development Authority created the first governmental agency to promote the use of secondary markets for ITIN loans, but they would be shut down by the state government the following year.

[17]In late 2005, Wells Fargo also experimented with offering ITIN mortgages in LA and Orange counties in California.

References

Abadie, Alberto. 2021. "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects." *Journal of Economic Literature*, 59(2): 391–425.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association*, 105(490): 493–505.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2011. "Synth: An R Package for Synthetic Control Methods in Comparative Case Studies." *Journal of Statistical Software*, 42(13): 1–17.

Galiani, Sebastian, and Brian Quistorff. 2016. "The synth_runner Package: Utilities to Automate Synthetic Control Estimation Using." 16.

Gallagher, Mari. 2005. "Alternative IDs, ITIN Mortgages, and Emerging Latino Markets." *Profitwise News and Views*.

Jordan, Miriam. 2008. "Mortgage Prospects Dim for Illegal Immigrants." *The Wall Street Journal*.

Khimm, Suzy. 2014. "The American Dream, undocumented." *MSNBC*.

Roosevelt, Margot. 2017. "Can undocumented workers get a mortgage?" *The Orange County Register*.