WPS7605

Policy Research Working Paper 7605

Revisiting the Impact of the Brazilian SIMPLES Program on Firms' Formalization Rates

Caio Piza

WORLD BANK GROUP
Development Research Group
Impact Evaluation Team
March 2016

Abstract

A recent survey of rigorous impact evaluations of programs to help small and medium-size firms to formalize indicates that the programs do not seem to work for most informal firms. One of the few exceptions finds large effects of a tax simplification program in Brazil called SIMPLES on firms' formalization rates and performance indicators. Using the same data set but a different identification strategy, another study concludes that the program had limited effect on formalization rates. The aim of this paper is twofold. First, it revisits the two studies to reconcile their conflicting conclusions. Second, it investigates the validity of the identification strategy of both studies. The findings suggest that the conflicting results between the two studies are caused by the dates each used to identify when the program was put into effect. A robustness check indicates that data heaping and seasonality around November cast doubts on the identification strategy used in both studies to estimate the effect of this particular program.

This paper is a product of the Impact Evaluation Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at caiopiza@worldbank.org.

Revisiting the Impact of the Brazilian SIMPLES Program on Firms' Formalization Rates[1]

Caio Piza
Development Impact Evaluation Unit (DECIE)
The World Bank Development Research Group
Contact email: caiopiza@worldbank.org

Keywords: Business Environment, Business Taxation, SME Development, SIMPLES Program, and regression discontinuity design.

JEL: C26, K34, L20.

[1] I would like to thank Joana Monteiro and Gabriel Montes-Rojas for sharing their do-files. I would like to thank Guadalupe Bedoya, Astrid Zwager, Miriam Bruhn, David McKenzie, Aart Kraay, Bill Maloney, Gabriel Montes-Rojas, Florence Kondylis, Aidan Coville, Arndt Reichert, Arianna Legovini, Tulio Cravo, Prashant Bharadwaj, Owen Ozier, Eeshani Kandpal, Clement Joubert, Pedro Olinto, and the participants of the World Bank DECRG's half-baked seminar for comments and suggestions. All conclusions are mine.

1. Introduction

Access to credit is a major barrier preventing informal firms from growing. Some argue that reducing the costs of formalization (red tape costs) would help firms to access credit in commercial banks and subsidized credit lines from governmental agencies to support scale and growth. The relationship between access to finance and growth is suggested by anecdotal evidence (de Soto, 1989 and 2000) and, more formally, by theoretical models (see Paula and Scheinkman, 2010 for a reference).

Recently, applied researchers have used experimental and quasi-experimental impact evaluation methods to test the effect of tax simplification policies on formalization rates and performance outcomes of micro, small, and medium-sized enterprises (MSMEs). In a recent survey, Bruhn and McKenzie (2014) summarize the evidence of rigorous impact evaluations; they argue that most informal firms do not seem to formalize and therefore do not benefit from such policies. Fajnzylber et al. (2011) can then be seen as an exception.

Fajnzylber et al. (2011) found large effects of the Brazilian tax simplification program SIMPLES (*Sistema Integrado de Pagamento de Impostos e Contribuições das Microempresas e Empresas de Pequeno Porte*) on formalization rates and performance indicators of firms with up to five employees. They used the change in legislation at the end of 1996 as an event study and used regression discontinuity (RD) design and difference-in-difference (DID) methods to evaluate the impact of the program.

Their main findings show an increase in the formalization rates of about 11 percentage points (or 50 percent) and local average treatment effects on revenues and profits of about 361 and 326 percent, respectively (see table 6 in their paper). They also found an impact on fixed assets and employment. Interestingly, they found no effect on access to credit. This could suggest that formal firms created just after the change in the law were not credit constrained and may have decided to start up as formal for other reasons.

Curiously, the same journal published another evaluation of the same tax simplification program that arrived at different conclusions. Unlike Fajnzylber et al. (2011), Monteiro and Assunção (2012) found a positive and significant impact on formalization rates but only among firms in the retail sector. The authors report no results on firm performance indicators and are, in fact, very skeptical about the impact of SIMPLES on formalization rates and firm performance.

Looking carefully at both studies, a few differences stand out. For instance, Monteiro and Assunção (2012) use firms that are 0-20 months old, while Fajnzylber et al. use firms that are 3-20 months old, with a kernel function with one month bandwidth to assign higher weights to observations close to the threshold (the date the law was passed). Another difference may have to do with the sample composition. Fajnzylber et al. (2011) considered entrepreneurs aged 20-65 without college degrees while Monteiro and Assunção (2012) apparently used the whole sample.

Monteiro and Assunção (2012) used ineligible firms that started business after December 1996 as a counterfactual. In the RD framework, the counterfactual is less clear, as the data do not allow one to distinguish between genuine new entrants and informal firms that decided to formalize with the program. A key difference between these two studies is the date each used for the implementation of the program. Monteiro and Assunção (2012) used December, while Fajnzylber et al. (2011) used November 1996 as

the date the law was passed. As will be shown below, this explains in great part the disagreement between the studies.

This paper revisits the identification strategy used in both studies and performs a thorough check of the RD estimates of the impact of the SIMPLES program on firms' formalization rates. I focus on the RD estimates because the identification strategy of both RD and DID relies on a jump in formalization rates after the program started. Notice that an invalidation of the first-stage results would call into question any treatment effect estimate of the program on outcomes related to firms' performance obtained through an IV method.

In my analysis I use the same data set as in both studies, the Brazilian Survey of the Urban Informal Sector (ECINF) collected in October 1997 by the Brazilian Bureau of Statistics (IBGE), but in a format that is closest to Monteiro and Assunção (2012), that is, I do not restrict the sample as in Fajnzylber et al. (2011), and I look at the sample of all micro firms without distinguishing between micro-entrepreneurs and small firms.[2] It is important to emphasize that this paper should not be seen as a replication exercise. The objectives of this paper are (1) to identify reasons for the studies coming to different conclusions, even though they used the same data set, and (2) to test the robustness of the observed jump in formalization rates detected by Fajnzylber et al. (2012).

Even though ECINF has more than one variable that could be used to capture formalization rates (see Fajnzylber et al. 2011), I will focus on the results obtained with the variable, 'license'. This variable tells us whether a firm had a municipal or state license to operate. This variable was used in both studies to measure the impact of SIMPLES on formalization rates and to estimate the impact of formalization on firms' performance through two-stage least squares by Fajnzylber et al. (2011).[3]

Fajnzylber et al.'s (2011) main contribution in terms of identification strategy is that they explore more deeply the information on firms around the threshold. It is unfortunate, however, that they do not perform RD validity tests. When revisiting their RD estimates, I will perform a couple of validity tests that are now standard among RD design practitioners. These include the McCrary density test (2008) for the manipulation of the assignment variable and a balance test around the cutoff point to check whether the covariates are similar for firms created just after and just before the law passed (see Lee and Lemieux, 2010).

I will also investigate whether the potential measurement error in the assignment variable, as discussed in Monteiro and Assunção (2012), may bias RD estimates to some extent and, if so, how (see Barreca et al. 2015 for this point). Finally, I will run a placebo test using different cutoffs to check whether the discontinuity in formalization rates around November/December 1996 is exclusively likely due to the law or influenced by some other factors, such as seasonality issues that could be correlated to firms' creation.

Beyond this brief introduction, this paper is organized as follows. The next section briefly describes the program, while section 3 revisits the first-stage estimates using an RD design. Section 4 discusses some robustness checks of the RD design. The main findings are summarized in the conclusion.

[2] Fajnzylber et al. (2011) estimates the effect of the program for all micro-firms and then separately for microenterprises and small firms with at least one employee.

[3] Fajnzylber et al. (2011) show first stage estimates for five alternative measures of formalization, but license showed larger point estimates for the sample of all micro-firms and micro-entrepreneurs. For firms with at least one employee, the variable 'legal entity' showed the largest point estimate.

2. The SIMPLES Program[4]

The SIMPLES program was envisaged with the objective to simplify the tax system for SMEs. To be able to participate in the new program, microenterprises (self-employed) should have an annual revenue up to Brazilian real 120,000, while the annual revenue of small firms should be under Brazilian real 1.2 million.[5] The system combines six different federal taxes and social contributions into one single monthly-based rate: IRPJ (corporate income tax), PIS/PASEP (contribution to employee savings programs), CSLL (contribution on net profit), COFINS (contribution for financing the social security system), IPI (industrialized products tax), and employer social security contributions. Firms still had to pay other federal, state, and municipal taxes.

The reforms reduced the tax burden considerably. Rather than paying from 5 to 11 percent of gross revenues on taxes, under SIMPLES, micro firms would pay from 3 to 5 percent and small firms between 5.4 and 8.6 percent. In the beginning, only some sectors were covered by the new law. Eligible firms had to belong to one of the following sectors: retail trade, manufacturing, transportation, civil construction, and other services that do not require professionals with regulated occupations such as auditors, architects, dentists, engineers and others.

3. Reconciling the Fajnzylber el al. and Monteiro and Assunção Results

A key difference between the two studies is the date considered for implementation of the program. According to Monteiro and Assunção (2012, p. 108):

"*Although we have only cross-sectional survey of firms in October 1997 [the Ecinf 1997], we introduce a time dimension in our analysis by considering firms created before and after the new legislation, which was December 1996.*"

Fajnzylber et al. considered November 1996 as cutoff instead. According to them (2011, p. 263):

"*In November, the Brazilian Government implemented a new simplified tax system for micro and small firms, the SIMPLES.*"

In other words, Monteiro and Assunção (2012) used December 1996, the month the law was approved in the Brazilian Congress, as cutoff. Fajnzylber et al. (2011) used November 1996 instead, the month the federal government drew up the provisional measure creating the new program.[6]

Since ECINF 1997 indicates the number of months a firm is in business, Monteiro and Assunção (2012) specify their DID regression to obtain estimates comparing outcomes of eligible and ineligibles firms created before and after December 1996. Their estimates include firms from 0 to 20 months old. That is, firms that started business between February 1996 and October 1997.

On the other hand, Fajnzylber et al. (2011) use an RD design to compare eligible firms created after November 1996 with eligible firms that were in business already. Like Monteiro and Assunção, Fajnzylber et al. (2011) also employed a DID approach, but they used a different sample of firms and gave more weight to observations close to the cutoff point.

---

[4] This section is brief summary of section 2 in Monteiro and Assunção (2012).
[5] The exchange rate at that time was 1.04 Brazilian Real per 1 U.S. dollar. See www.bacen.gov.br
[6] Although the provisional measure gave firms the right to formalize under the simplified tax program, Congress would still have to approve it to turn it into a law.

Both studies test whether firms born after the law change were more likely to start becoming formal and their conclusions are different to some extent. Fajnzylber et al. (2011) found a 50 percent increase in formalization rates among all firms. Monteiro and Assunção (2012) found a similar effect, but only among firms in the retail sector. For the pooled sample of firms, they found an effect not statistically different from zero. As shown in the section below, the main source of disagreement has to do with the dates used as thresholds.

3.1. First-Stage Estimates: The Cutoff Matters

Fajnzylber et al. (2011) use RD (before-and-after) and DID methods to estimate the impact of the tax simplification program on firms' formalization rates. Although the counterfactual of each method is different, and hence the interpretation of the treatment effect, both methods rely on some jump (discontinuity) in formalization rates among eligible firms just after November 1996.

With the RD design, they ask what the formalization rates of eligible firms would have been in the absence of the intervention. Unfortunately, the data do not allow one to know whether the formal firms created just after November 1996 are new entrants, informal firms that decided to formalize, or a mix of the two.

If the program mostly created new formal firms, one could ask why informal active firms did not formalize. It could be that it was faster to create a new formal firm than formalize an active informal one. If that was the case among those that remained informal, there would be a subgroup that would have formalized had the formalization cost been as low as for the new entrants. Otherwise, the informal ones would not be a valid counterfactual to the new entrants. To use Angrist, Imbens, and Rubin's (1996) terminology, one would be comparing 'never takers' with a mix of 'always takers' and 'compliers'.

On the other hand, if the program mostly incentivized a group of informal firms to formalize, firms that decided to remain informal would not be a valid counterfactual, given that the tax simplification applies equally for all eligible firms. Thus, informal firms that opted for formalizing would be different from those that did not.

The DID method improves the identification strategy to some extent. First, it uses the ineligible group of firms as the counterfactual. One can question whether ineligible firms are a valid counterfactual, but with the parallel trend assumption it would require that, in the absence of the program, the formalization rates would have evolved equally between eligible and ineligible firms. In other words, with the DID approach, the outcome variable of eligible and ineligible groups is allowed to be different in level. Second, it would control for time (seasonality) effects; that could explain to some extent the creation of new formal firms in the end of a calendar year. In a sense, the DID would clean up estimates by reducing the influence of always-takers in the composition of new formal firms.

That said, the estimation of the effect of the program on formalization rates depends on an observed jump in formalization rates among the eligible group of firms and a smooth continuous function among ineligible firms by the time the program came into effect.

The charts below basically replicate figures 1A and 1B in Fajnzylber et al. (2011). The main difference is that the authors fitted a cubic polynomial, while I fit a local linear regression. Notice, as in Fajnzylber et al., young firms (0-3 months of age) are dropped from the analysis.[7]

[7] Using the code of Monteiro and Assunção (2012,) I get a similar figure as shown in the appendix.

Figure 1A – Proportion of Formal Eligible Firms Before and After SIMPLES Threshold: November 1996

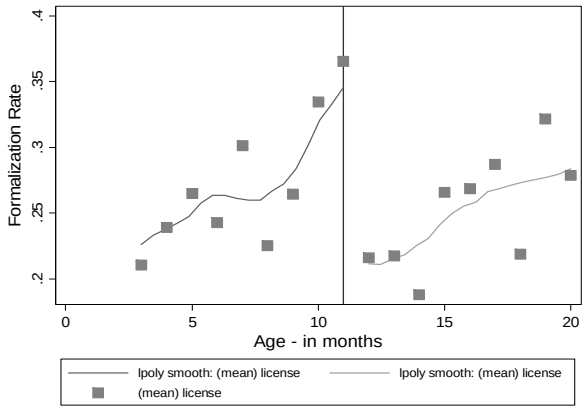
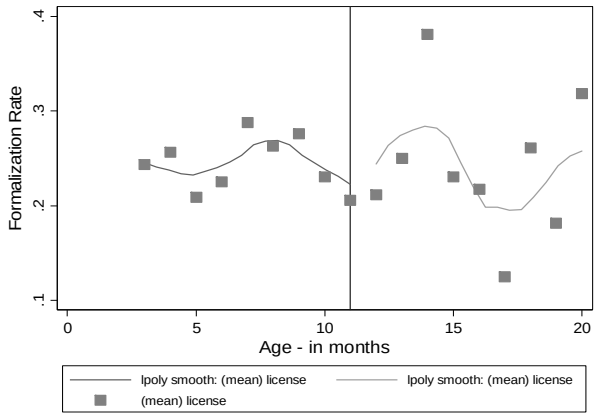


Figure 1B – Proportion of Formal Ineligible Firms Before and After SIMPLES Threshold: November 1996



The figures show a pattern that one would expect: a higher proportion of formal firms after the program and no apparent change in formalization rates of ineligible firms. These two figures could suggest that RD and DID would be valid methods to estimate the impact of the program.

It is interesting to note that in figure 1A, the cluster of firms created in November and December 1996 seem key to the overall impact of the program on formalization rates, particularly if one uses a kernel function that gives more weight to observations close to the threshold. In fact, as pointed out by Fajnzylber et al. (2011, p. 265), "a striking feature of the data is that SIMPLES does not appear to have a permanent effect on the share of firms registered."

Indeed, the impact of SIMPLES begins to vanish as one moves away from the threshold. In other words, the effect seems to be very local. Therefore, the results in both Fajnzylber et al. (2011) and Monteiro and Assunção (2012) seem mostly driven by the cluster of firms created at the end of 1996.

Based on the observed jump, Fajnzylber et al. (2011) estimate reduced form and two-stage least square RD regressions using firms with time in business ranging from 3 to 20 months and a kernel function with one-month bandwidth. Table 1 below reproduces their first stage results.

Table 1 – Fajnzylber et al. (2011) First-Stage Estimates for All Micro-Firms

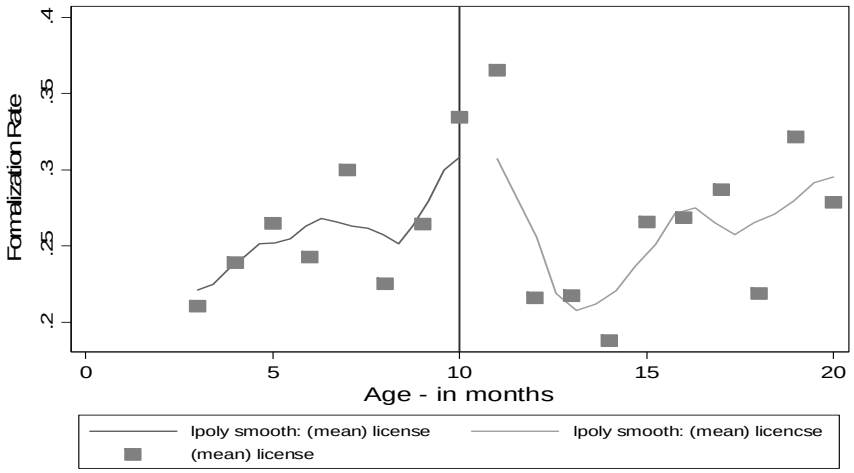
Formality Measures	Eligible firms before-after (All micro-firms)	
License to operate	0.116***	(0.034)
Legal entity	0.075***	(0.022)
Micro-firm registration	0.063***	(0.020)
Registered with tax authorities	0.072***	(0.022)
Paid taxes	0.031**	(0.016)
Paid social security	0.043***	(0.015)

Source: Fajnzylber et al. (2011, p. 273), Table 5.
Note: ***, ** Statistically significant at 1 and 5 percent respectively. Standard errors in parentheses. The estimates include as covariates, age, gender, and education of the firm's owner, household size, dummies for the reasons of why the entrepreneurs decided to start business, and dummies for state and industry variables. The sample considers firms 3-20 months old and entrepreneurs aged 20-65 without college degrees.

The results presented in table 1 suggest that the program had a large impact on formalization rates. Given the percentage of 'control' firms with license to operate of 22.4 percent, 11 percentage points is equivalent to an effect of about 50 percent.[8]

As discussed below, the impact on formalization rates found in Fajnzylber et al. (2011) seems to be mostly driven by firms created in November 1996, that is, the date the provisional measure was drawn up. Figure 2 illustrates the discontinuity with the threshold defined in December 1996 by Monteiro and Assunção (2012). The discontinuity practically disappears. This explains why Monteiro and Assunção (2012) found limited impact of the program on firms' formalization rates.

Figure 2 - Proportion of Formal Firms Before and After SIMPLES Threshold: December 1996



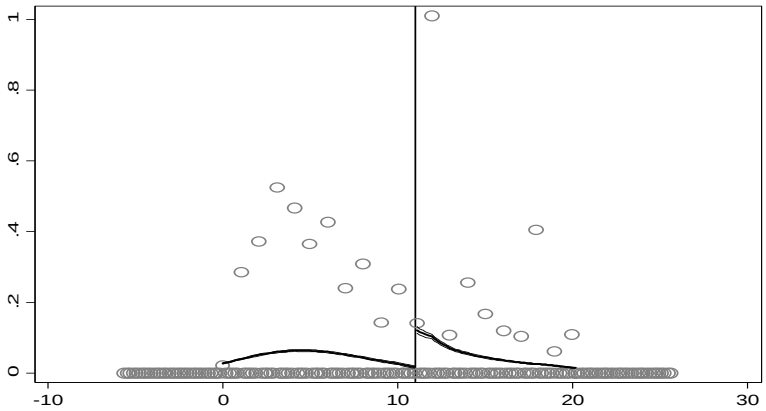
[8] The authors then use the law as an instrumental variable and estimate the impact of formalization rates on firms' performances through the weighted two-stage least squares.

According to figure 2, firms created in November 1996 seem to be driving the impact on formalization rates. In addition, this spike in formalization rates in November 1996 could also result from some sorting, that is, some manipulation of the assignment variable by some firms. It could be that firms provide their 'age' to the surveyors based on the time they operated formally, not including the time they were in business informally.

Because the results seem to be driven by firms created by the time the provisional measure was drawn up, it is crucial to check whether there are indications of sorting in the data. To check whether there was manipulation of the assignment variable, I ran the McCrary density test. This is a non-parametric test that compares the distribution of the assignment variable on each side of the cutoff point. Under the null hypothesis of no-sorting, the distribution of the assignment variable should be continuous (smooth) around the cutoff point. With the rejection of the null, one could then question the validity of the RD estimates because the rejection usually suggests that firms on each side of the cutoff point are systematically different (see Lee and Lemieux 2010, p. 348).

Figure 3 shows that the null hypothesis is strongly rejected. There is an indication that the assignment variable is not smooth around the threshold. However, the test might be driven by the spike in the number of firms created in October 1996, 12 months before the collection of the survey.[9]

Figure 3 – McCrary Density Test for the Manipulation of the Assignment Variable Threshold: November 1996



In fact, both studies make the important point about the risk of manipulation of the assignment variable. Fajnzylber et al. (2011, p. 268) argue, "*The way SIMPLES was introduced may have affected the type of firms that registered before and after. For example, to the degree that the reforms were discussed in the Congress prior to implementation, they may have been anticipated by some entrepreneurs who may have delayed starting their business until SIMPLES was in place. If firms that engaged in this kind of thinking are non-randomly distributed, our identification strategy is weakened.*"

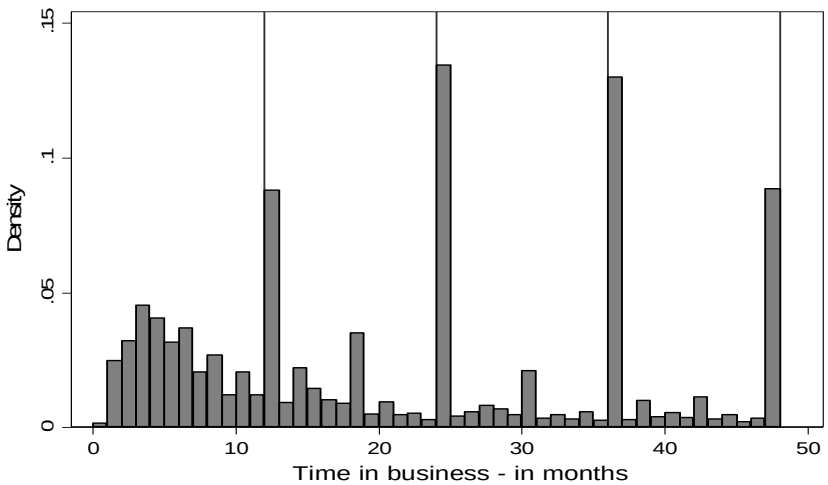
Monteiro and Assunção (2012) dedicate a whole section in their paper to discuss the potential measurement error in the assignment variable. They show a histogram with spikes in the proportion of firms with reported time in business of 12 months (created in October 1996) and interpret this result as an

[9] The appendix shows that the test is also rejected if December is used as threshold.

indication of measurement error. This is very likely to be the case, but it would have direct consequences on the RD estimates. As discussed in Barreca et al. (2015), it would not necessarily harm the results in case the data heap does not predict the outcome, but, if it does, the RD estimates would be biased.

Figure 4 basically reproduces figure 1 in Monteiro and Assunção (2012), but it covers firms 0 to 50 months old to check whether the spike observed in October 1996 might have to do with the way firms reported their age. The solid vertical lines are for months 12, 24, 36, and 48. The spikes suggest that some firms seem to round off their ages when reporting their time in business.[10]

Figure 4 – Evidence of Heaping Data in the Assignment variable

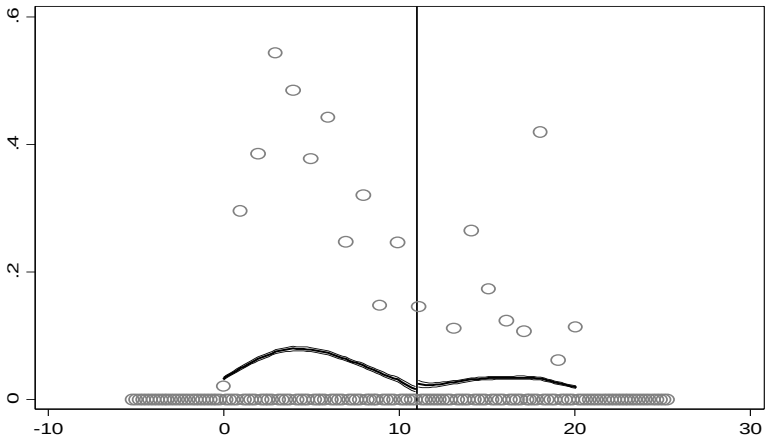


According to Barreca et al. (2015), heaping is very common in self-reported data. The problem is that RD estimates with heap data might be biased if the heap is observed close to the cutoff point. Notice that in this case, even a very narrow bandwidth, say one month, would not circumvent the problem. On the contrary, in the present case, the result could be upward biased due to the heap.

One way of dealing with this issue is through a 'donut-RD approach,' that is, dropping data at the vicinity of the cutoff point. Figure 5 shows the McCrary density test with the heap data excluded. Dropping all firms that supposedly entered business in October 1996, one year before the survey was collected, almost eliminates the problem of sorting. Notice that the null hypothesis of non-manipulation is still rejected. The test suggests that the increase in the proportion of younger firms might influence the overall result.

[10] I would like to thank Miriam Bruhn for raising this point.

Figure 5 – McCrary Density Test for the Manipulation of the Assignment Variable Threshold: November 1996



The next step consists of checking the balance around the cutoff point with and without the data heap. Table 2 below shows the coefficient for a t-test of difference in means comparing firms created after November 1996 with firms created before. The table shows estimates with different bandwidth sizes. Overall, firms born before and after November 1996 do seem comparable even for relatively large bandwidths. However, a joint-test for difference in means rejects the null of equal means, even for a three-month bandwidth. Table 3 replicates the exercise, but this time dropping firms created in October 1996.

Interestingly, except for a three-month bandwidth, dropping firms created in October 1996 worsen the balance condition, according to the joint test shown in the last column of table 3. One could argue that those firms created in October 1996 do not seem to be different, on average, from the firms created after the law. For that to be the case, one should observe no big difference in first stage estimates once the heap is controlled for, which does not seem to be the case, as discussed below.

Table 2 – Balance Check with Multiple Bandwidth Sizes

Bandwidth size in months	Obs	Primary school	High school	College	Gender	Skin color	Household size	Age	Start working age	Firms' location (=1 if outside home)	Firms' location (=1 if exclusively outside home)	Borrowed money from relatives to start	Used owned resources	Asset (=1 for owned dwellings)	*Joint test*
3	1476	-0.0070	-0.0079	0.024	-0.035	0.079	-0.31	-0.41	-0.66	0.026	-0.072**	0.034	0.0064	-0.061	
		(-0.14)	(-0.17)	(1.08)	(-0.69)	(1.52)	(-1.64)	(-0.36)	(-0.97)	(0.53)	(-2.46)	(0.99)	(0.12)	(-1.32)	[0.075]
4	1748	-0.0039	-0.017	0.016	-0.034	0.059	-0.22	0.0026	-0.70	0.041	-0.062**	0.034	-0.0097	-0.061	
		(-0.086)	(-0.43)	(0.88)	(-0.75)	(1.28)	(-1.34)	(0.0025)	(-1.14)	(0.94)	(-2.31)	(1.15)	(-0.21)	(-1.48)	[0.021]
5	2114	0.0018	-0.025	0.019	-0.034	0.071*	-0.16	0.18	-0.87	0.026	-0.048*	0.022	0.011	-0.058	
		(0.044)	(-0.68)	(1.16)	(-0.83)	(1.73)	(-1.06)	(0.19)	(-1.51)	(0.67)	(-1.95)	(0.83)	(0.27)	(-1.57)	[0.019]
6	2427	-0.020	-0.019	0.024	-0.013	0.057	-0.20	0.30	-0.85	0.026	-0.040*	0.0086	0.029	-0.029	
		(-0.53)	(-0.56)	(1.53)	(-0.34)	(1.48)	(-1.45)	(0.36)	(-1.61)	(0.72)	(-1.74)	(0.35)	(0.76)	(-0.84)	[0.023]
7	3011	-0.035	-0.0066	0.037**	-0.018	0.046	-0.085	0.036	-0.61	0.0093	-0.034	0.0015	0.043	-0.00091	
		(-1.04)	(-0.21)	(2.53)	(-0.52)	(1.33)	(-0.68)	(0.047)	(-1.29)	(0.29)	(-1.64)	(0.069)	(1.24)	(-0.030)	[0.001]
8	-0.042	0.0040	0.037***	-0.031	0.061*	-0.076	0.32	-0.37	0.026	-0.013	-0.020	-0.0011	0.037	-0.0045	
	(-1.27)	(0.13)	(2.59)	(-0.94)	(1.83)	(-0.63)	(0.43)	(-0.81)	(0.53)	(-0.40)	(-1.00)	(-0.050)	(1.10)	(-0.15)	[0.000]

Note: ***, **, * statistically significant at 1, 5, and 10 percent. Robust t-statistics in parenthesis. The coefficients are from a linear regression of each covariate on a constant, a (treatment) dummy that is 1 for firms created after the threshold and 0 otherwise, and the assignment variable Z. Robust t-statistics in parentheses. P-value of the joint test in brackets.

Table 3 – Balance Check with Multiple Bandwidth Sizes (Excludes Firms Created in October 1996)

Bandwidth size in months	Obs	Primary school	High school	College	Gender	Skin color	Household size	Age	Start working age	Firms' location (=1 if outside home)	Firms' location (=1 if exclusively outside home)	Borrowed money from relatives to start	Used owned resources	Asset (=1 for owned dwellings)	*Joint test*
3	798	0.042	-0.046	0.042	-0.082	-0.22	0.060	-2.79	3.72*	0.27	0.27	-0.17	-0.24	0.10	
		(0.22)	(-0.26)	(0.52)	(-0.43)	(-1.10)	(0.088)	(-0.59)	(1.91)	(1.41)	(1.43)	(-1.36)	(-1.29)	(0.59)	[0.40]
4	1070	0.12	-0.098	-0.060	-0.023	-0.29**	0.29	1.02	0.062	0.17	0.18	-0.049	-0.19	-0.022	
		(0.97)	(-0.88)	(-1.31)	(-0.19)	(-2.34)	(0.66)	(0.35)	(0.047)	(1.37)	(1.55)	(-0.64)	(-1.57)	(-0.19)	[0.002]
5	1436	0.061	-0.064	-0.034	-0.050	-0.095	0.37	1.09	-0.30	0.026	0.061	-0.0011	-0.086	-0.068	
		(0.66)	(-0.75)	(-0.90)	(-0.54)	(-1.00)	(1.12)	(0.50)	(-0.27)	(0.29)	(0.68)	(-0.019)	(-0.94)	(-0.79)	[0.001]
6	1749	0.019	-0.073	-0.020	-0.028	-0.11	0.23	1.49	-0.64	0.046	0.041	0.0060	-0.040	0.033	
		(0.25)	(-1.06)	(-0.64)	(-0.37)	(-1.44)	(0.83)	(0.86)	(-0.73)	(0.62)	(0.57)	(0.13)	(-0.53)	(0.50)	[0.000]
7	2333	-0.031	-0.011	0.031	-0.029	-0.051	0.25	1.16	-0.077	-0.013	-0.0070	0.0050	0.0035	0.045	
		(-0.53)	(-0.21)	(1.21)	(-0.52)	(-0.87)	(1.20)	(0.88)	(-0.11)	(-0.24)	(-0.13)	(0.15)	(0.062)	(0.89)	[0.000]
8	2724	-0.043	0.0050	0.036	-0.053	-0.033	0.31	1.01	0.16	-0.041	-0.039	0.024	-0.0019	0.042	
		(-0.78)	(0.099)	(1.41)	(-0.98)	(-0.60)	(1.51)	(0.79)	(0.25)	(-0.75)	(-0.74)	(0.74)	(-0.034)	(0.85)	[0.000]

Note: ***, **, * statistically significant at 1, 5, and 10 percent. Robust t-statistics in parenthesis. The coefficients are from a linear regression of each covariate on a constant, a (treatment) dummy that is 1 for firms created after the threshold and 0 otherwise, and the assignment variable Z. Robust t-statistics in parentheses. P-value of the joint test in brackets.

4. Robustness Check and Placebo Tests

Two straightforward robustness checks of RD estimates consists of (1) running the RD regressions using December 1996 as threshold and (2) running the RD regressions with November 1996 as threshold, but excluding firms created in October 1996. The latter strategy is referred to as the 'donut-RD approach'. First stage estimates are provided with several bandwidth sizes. Throughout this paper, I use spline linear specification because the results are similar with quadratic polynomials and quadratic splines.

Table 4 – First Stage Regression – Effect of SIMPLES on Formalization Rates

	3-month	4-month	5-month	6-month	7-month	8-month
			Cutoff = November 1996			
D	0.13***	0.099**	**0.099****	**0.11*****	**0.10*****	**0.12*****
	(2.63)	(2.3)	(2.57)	(3.03)	(3.27)	(3.92)
N	1399	1664	2012	2315	2860	3236
			Cutoff = December 1996			
D	**-0.17****	-0.042	0.026	0.042	0.051	0.058*
	(-2.06)	(-0.84)	(0.61)	(1.11)	(1.5)	(1.92)
N	1404	1831	2178	2562	2971	3453
			Cutoff = November 1996 (excludes firms created in Oct 1996)			
D	0.059	0.13	0.11	0.079	0.076	0.095*
	(0.33)	(1.19)	(1.31)	(1.16)	(1.45)	(1.88)
N	744	1009	1357	1660	2205	2581
			Cutoff = November 1996 (dummy for firms created in Oct 1996)			
D	0.055	0.0069	0.035	0.065*	0.072**	0.096***
	(0.90)	(0.13)	(0.80)	(1.67)	(2.07)	(2.98)
N	1312	1577	1925	2228	2773	3149

Note: ***, **, * statistically significant at 1, 5, and 10 percent. These are spline linear regressions in which y is regressed on a constant, a dummy that is 1 for firms created after the threshold and 0 otherwise (D), the assignment variable defined in months (Z) and an interaction term between the 'after' dummy and the assignment variable (DZ) to allow for different trends in each side of the threshold. Robust t-statistics in parentheses. The average of eligible firms aged 12 to 20 months (control firms) that are formal is 22.6 per cent. Excluding firms 12 months old, the average is 23.5 per cent.

Table 4 has some interesting results. The first block of results uses November 1996 as cutoff. The coefficients are large and statistically significant at the 1 percent level in most cases. The effect size and the statistical significance are very similar, regardless of the window size.

The second block of results shows the first stage with the threshold defined in December 1996 instead. The coefficient is negative, large and statistically significant for a three-month bandwidth. It turns positive and statistically significant only for a relatively larger window. This result is further evidence of

why Monteiro and Assunção (2012) may not have found similar impact of the program on formalization rates as Fajnzylber et al. (2011).

The third block of results shows the first-stage estimates when firms that started their business in October 1996 are dropped from the sample. The coefficients are positive, a bit unstable, and statistically significant only for relatively larger bandwidth size. Since dropping firms reduces power, the fourth block shows estimates with dummy variables for firms created in October 1996 (see Barreca et al. 2015). In fact, the point estimates are similar to those in the previous block, but more precisely estimated.

One should bear in mind some indication of imbalance in observed characteristics of firms created before and after November 1996 when interpreting the positive and statistically significant impact of the program on the formalization rates shown in the fourth block of table 4. The more one moves away from the threshold, the more likely the estimates will be confounded by unobserved factors.

A final robustness check consists in running placebo tests setting the cutoff in different points in time. The tests involve checking whether the discontinuity observed around November 1996 exclusively captures the effect of the program or is some other factor responsible? A factor not investigated in this paper, but that could be correlated with firms' creation.

Table 5 – Placebo Tests

	3-month	4-month	5-month	6-month	7-month	8-month
			Cutoff = November 1993			
D	-0.024	0.057	0.091	0.069	0.059	0.079
	(-0.24)	(0.64)	(1.16)	(0.93)	(0.86)	(1.22)
N	724	787	876	920	1012	1039
Mean						
			Cutoff = November 1994			
D	0.15	0.17**	0.21***	0.20***	0.20***	0.19***
	(1.62)	(2.12)	(2.98)	(3.04)	(3.24)	(3.45)
N	1120	1166	1351	1416	1544	1622
Mean						
			Cutoff = November 1995			
D	0.018	0.0098	0.070	0.083	0.067	0.043
	(0.23)	(0.14)	(1.19)	(1.48)	(1.31)	(0.93)
N	1198	1287	1570	1670	1887	2020
Mean						

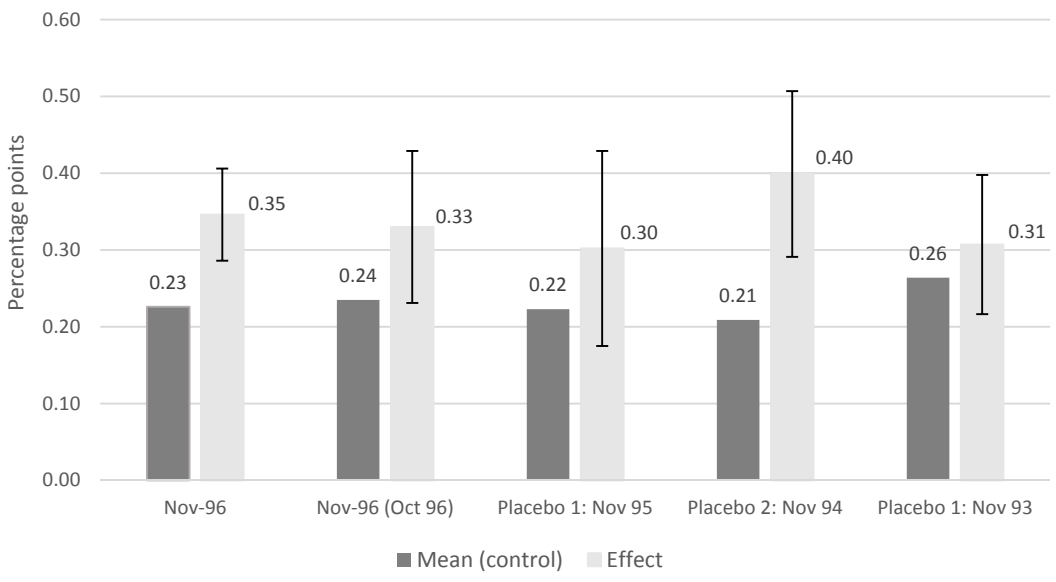
Note: ***, **, *statistically significant at 1, 5, and 10 percent. These are spline linear regressions in which y is regressed on a constant, a dummy that is 1 for firms created after the threshold and 0 otherwise (D), the assignment variable defined in months (Z) and an interaction term between the 'after' dummy and the assignment variable (DZ) to allow for different trends in each side of the threshold. Robust t-statistics in parentheses. In 1995, the proportion of eligible firms aged 24 to 32 months (control) was 22.3 percent. In 1994, the proportion was 20.9 percent for firms 36 to 44 months old, and in 1993 it was 26.4 percent for firms 48 to 56 months old.

Table 5 shows the RD estimates with three alternative cutoff points: November 1993, November 1994, and November 1995. Despite not being statistically significant, except for a three-month bandwidth, the magnitudes of the coefficients in the first row are very similar to the first-stage estimates discussed above. Even more striking are the results with the threshold defined in November 1994, two years before the program was implemented. The coefficients are very large and precisely estimated in most cases. Finally, the third block in the table shows the estimates with the threshold defined one year before the provisional measure was drawn up. The point estimates are very similar to the ones in the first block as well as the RD estimates shown in tables 1 and 4 (mainly the third and fourth blocks).

Figure 6 summarizes the first-stage estimates shown in table 4 and the placebo effects. To build the figure, I used point estimates obtained in an eight-month bandwidth. The dark-gray bar refers to the proportion of eligible firms that were formal before the program. For the first two bars, the mean was obtained through the sample of firms that were 12 to 20 months old. For the placebo estimates, I computed the means using eligible firms 24 to 32 months old for November 1995, 36 to 44 months old for November 1994, and 48 to 56 months old for November 1993. The light-gray bars show the mean of the 'treatment' group computed as the sum between the mean of the control and the effect size. The second light-gray bar refers to the estimate using a dummy for October 1996.

As can be seen, the dark-gray bars look very stable across the years. This suggests that the formalization rates among 'control' firms do not change with the time the control firms are in business. The figure also shows that the height of the light-gray bars and the mean for the 'treatment' groups are similar, even in the years that preceded the creation of the program. In fact, the 95 percent confidence intervals overlap in all cases, suggesting that the effect of the program on formalization is not statistically different from placebo estimates.

Figure 6 – Summary of Impact and Placebo Effects on Formalization Rates



Based on these results, it seems that both RD and DID estimates may be picking up, at least partially, something other than just the effect of the program. In the best case scenario, the RD estimates on formalization rates should be seen as upper bound effects of the SIMPLES program.

5. Conclusion

The paper revisited the impact of the tax simplification program for SMEs (SIMPLES) in Brazil to reconcile the conflict of findings and conclusions in Fajnzylber et al. (2011) and Monteiro and Assunção (2012). It then checked the validity of the identification strategy used in both studies, but the focus was on the RD as a credible design to assess the impact of this program. The paper used the ECINF data from 1997 as in the original papers, and found that the main reason for disagreement between the studies was to do with the month used as cutoff point to identify the 'pre and-post-dummy' variable. While Fajnzylber et al. used the date the provisional measure was sent to the Congress to vote, Monteiro and Assunção (2012) used the date the law was enacted. The difference is just one month, but it seems to be key to explain the difference in findings between the two papers. Firms that formalized in November 1996 apparently anticipated the approval of the law by the Congress. This cluster of firms that came into business in November 1996 seems to drive most of the large results found by Fajnzylber et al. (2011).

To check whether the jump in formalization rates was observed among young firms, the study undertook a couple of validity tests that have become standard among RD design practitioners. A McCrary density test to check for perfect manipulation of the assignment variable and a simple test of difference in means were performed to check balance around the cutoff using different bandwidth sizes. The McCrary test pointed to a perfect manipulation of the assignment variable around the cutoff point. This result usually suggests that 'treated' and 'control' firms can be systematically different. In fact, a joint test for difference in means suggested that eligible firms created before and after the program do not seem to look much alike on average. But these results seem to be highly influenced by spikes (heaps) in the number of firms created in October each year.

Because data heap could be confounding both the RD and DID estimates, a couple of robustness checks were performed to attest the impact of the program on formalization rates. First, RD regressions were re-run, controlling for data heaping. Second, placebos tests with different cutoff points were undertaken to check if there is some data seasonality by the end of each year that could be confounding the estimates.

It turned out that the available evidence of the impact of the program does not seem to be very robust, given the similar point estimates between what is supposed to be the effect of the program and those found in the placebo tests. The results seem to indicate the program was not effective in increasing the formalization rates of small firms; however, one could have a more cautious interpretation of the data and conclude that the microdata used in both Fajnzylber et al. (2011) and Monteiro and Assunção (2012) do not seem to be ideal to inform the impact of SIMPLES on firms' formalization decision and performance.

References

Barreca, A. I., Lindo, J. M., and Waddell, G. R. 2015, 'Heaping-induce bias in regression-discontinuity designs,' *Economic Inquiry*, pp. 1-26.

Bruhn, M. and McKenzie, D., 2013, 'Entry Regulation and Formalization of Microenterprises in Developing Countries,' Policy Research Working Paper 6507, World Bank.

de Soto, Hernando, 1989. *The Other Path: The Invisible Revolution in the Third World*. Harper & Row, New York, NY.

de Soto, Hernando, 2000, *The Mystery of Capital*: *Why Capitalism Triumphs in the West and Fails Elsewhere Else*. New York: Basic Books.

Paula, A. de and Scheinkman, J. A. 2010, Value-Added Taxes, Chain Effects, and Informality, *American Economic Journal: Macroeconomics*, Vol. 2, pp. 195-221.

Monteiro, J. C. M. and Assunção, J. J. 2012, Coming Out of the Shadows? Estimating the Impact of Bureaucracy Simplification Tax Cut on Formality in Brazilian Microenterprises, *Journal of Development Economics*, Vol. 99, pp. 105-115.

Fajnzylber, P., Maloney, W. F., and Montes-Rojas, G. W. 2011, Does Formality Improve Micro-Firm Performance? Evidence from the Brazilian SIMPLES Program, *Journal of Development Economics*, Vol. 94, pp. 262-276.

Lee, D. S. and Lemieux, T. 2010, Regression Discontinuity Designs in Economics." *Journal of Economic Literature*, 48 (2): 281-355.

Appendix

Figure A.1 - Proportion of Formal Firms Before and After SIMPLES (Monteiro and Assunção code)

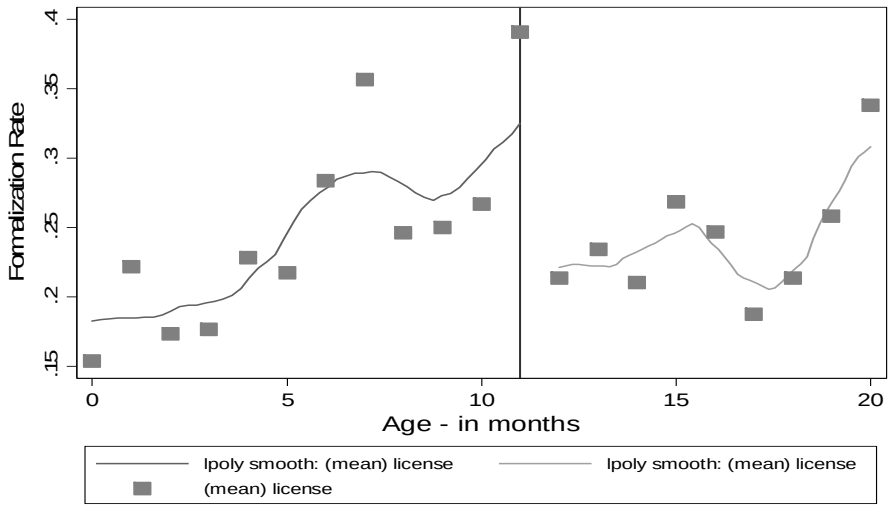


Figure A.2 - Proportion of Formal Firms Before and After SIMPLES (Monteiro and Assunção code)

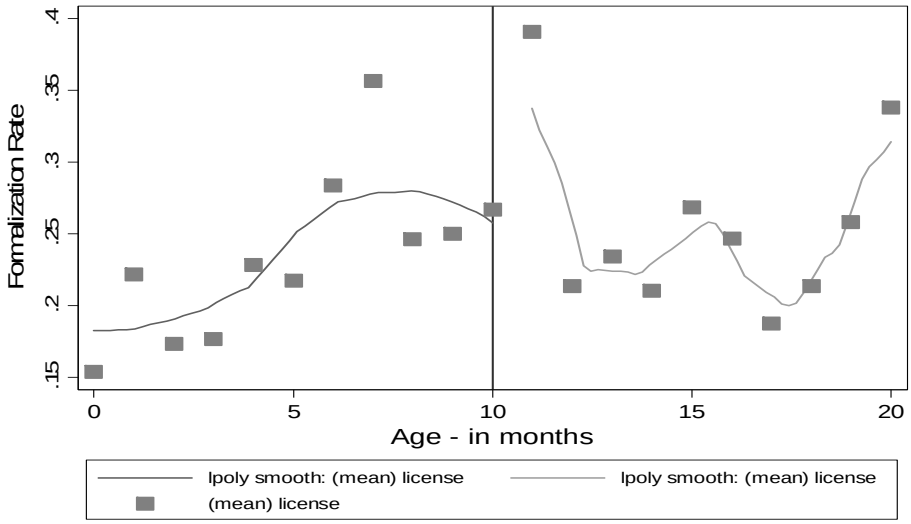


Figure A.3 – McCrary Density Test for the Manipulation of the Assignment Variable (Cutoff = Dec 1996)

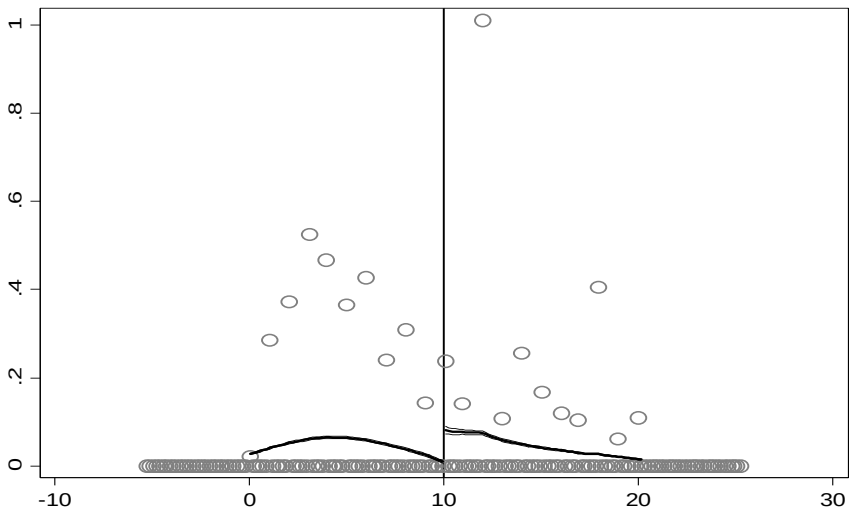


Figure A.4 – McCrary Density Test for the Manipulation of the Assignment Variable (Cutoff = Dec 1996)
Excludes Firms Created in October 1996

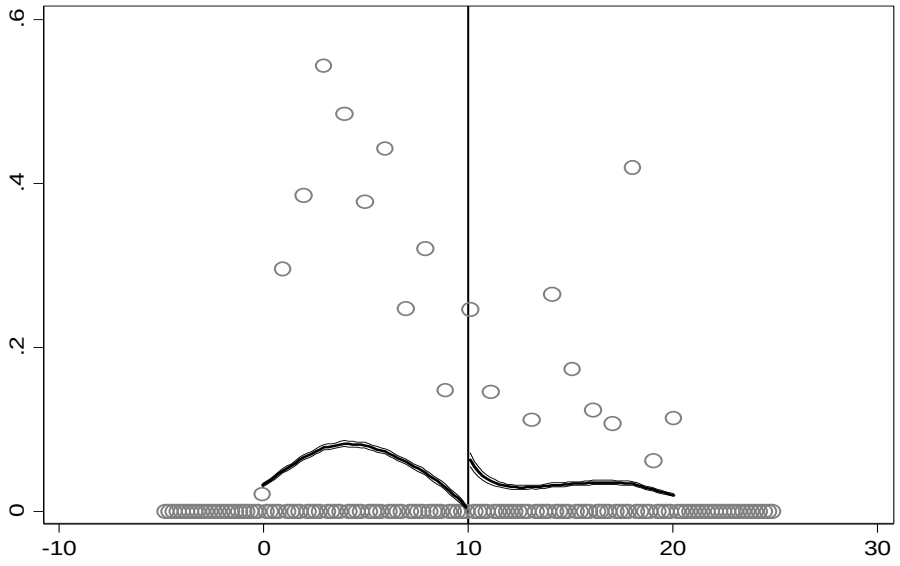


Figure A.5 – Formalization Rates Before and After November 1993

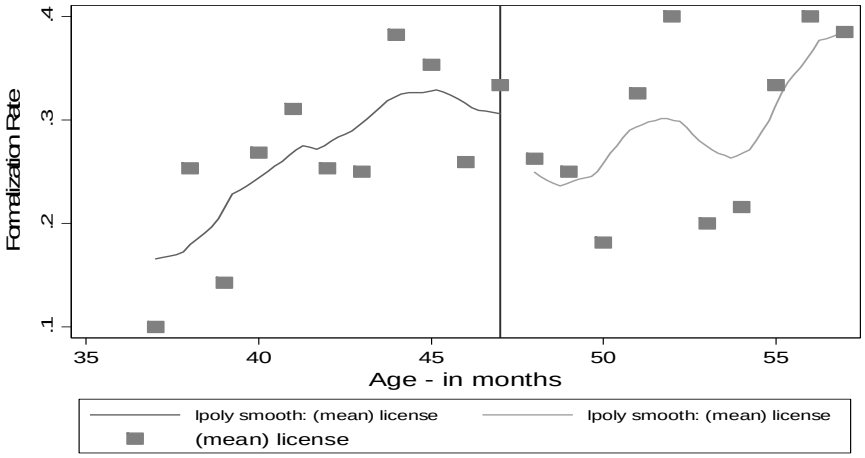


Figure A.6 – Formalization Rates Before and After November 1994

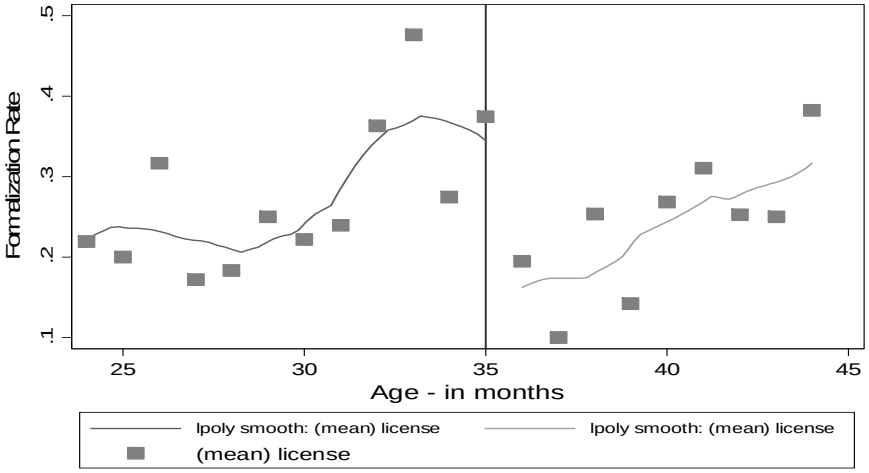


Figure A.7 – Formalization Rates Before and After November 1995