

Do Lower Taxes Reduce Informality?

Evidence from Brazil

Rudi Rocha (UFRJ) Gabriel Ulyssea (PUC-Rio)
Laísa Rachter (FGV)

WIDER Development Conference
Public Economics for Development

Maputo, July 6, 2017

Motivation

- Firms in developing countries are predominantly small (≤ 2 employees), while informality amongst them is very high.
 - Consequences for development: smaller tax base and misallocation of resources.
- The question of how to induce informal firms to formalize has motivated a large literature as well as an intense policy debate:
 - High registration costs are the usual suspects \rightarrow governments around the world have devoted substantial resources to reduce them.
 - The existing literature, however, shows little empirical support for the effectiveness of such reforms...
 - Perhaps the ongoing costs of formality constitute the relevant constraint.
- Little empirical evidence on the impacts of reducing taxes on firm informality.

This paper

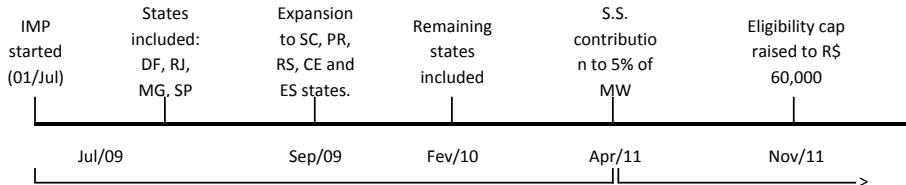
- This paper estimates the effects of reducing the ongoing costs of formality on firm formalization.
- More specifically:
 - We analyze a large-scale formalization program implemented in Brazil, the Individual Micro-Entrepreneur Program (IMP).
 - We exploit a difference-in-differences strategy, exploiting variation in exposure to the policy across industries, regions and time.
 - We can also assess whether reducing registration costs is a sufficient condition to induce firms to formalize.
 - We analyze if the formalization effects come from the formalization of existing informal businesses or from the creation of new formal ones.

- ➊ Related literature
- ➋ Institutional setting
- ➌ Data
- ➍ Empirical strategy
- ➎ Results:
 - Effects on the creation of new formal firms.
 - Intensive vs extensive margin?
 - General characterization: heterogeneity by income and timing.
 - Cost-benefit analysis and discussion.

- Cross-country evidence on the effects of entry costs and overall regulatory costs:
 - E.g. Djankov et al. (2002), Botero et al. (2004), and Djankov et al. (2010).
- Papers that estimate the effects of reducing registration costs but with the costs of being formal unchanged:
 - Non-experimental setting: Kaplan et al. (2011) and Bruhn (2011).
 - Experimental setting: De Mel et al. (2013), De Andrade et al. (2013), De Giorgi and Rahman (2013).
- Reducing the costs of being formal (SIMPLES), but quite endogenous:
 - Monteiro e Assunção (2012) and Fajnzylber et al (2011).

- The Individual Micro-Entrepreneur Program (henceforth IMP) aimed at reducing formalization costs in two dimensions:
 - ❶ The costs of entering the formal sector.
 - ❷ The costs of remaining formal: taxes and red tape.
- Program was designed to change both costs at once. However, there were two distinct phases: IMP1 and IMP2.
- Eligibility to the program was based on:
 - (i) Size (up to one employee).
 - (ii) Annual revenue (R\$36K \rightarrow R\$60K).
 - (iii) Industry.

Timeline



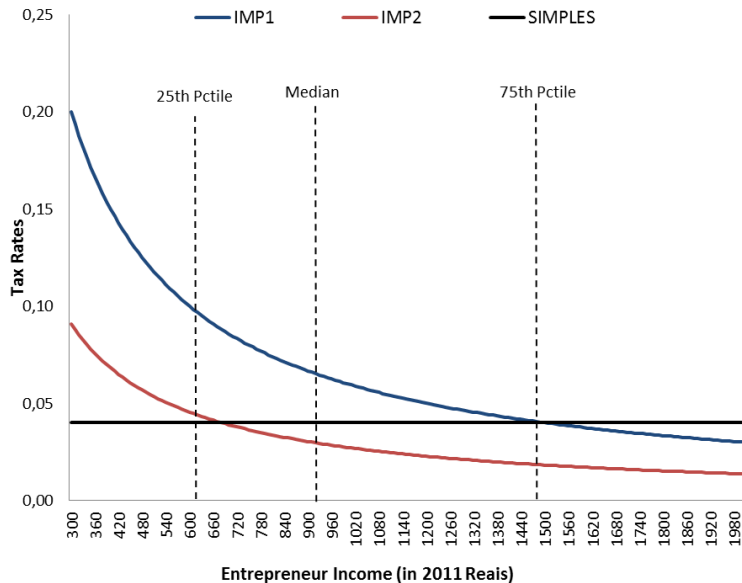
IMP Phase 1 (IMP1):

- * Eligibility cap: R\$36,000 per year
- * No entry costs
- * Fixed tax at 11% of MW

IMP Phase 2 (IMP2):

- * Eligibility cap: R\$60,000 per year
- * No entry costs
- * Fixed tax at 5% of MW

Tax rates under the different phases



RAIS:

- Administrative microdata set from the Ministry of Labor that contains the universe of formal firms and workers.
- Collapsed into a quarterly panel of industry-by-region cells, from January 2006 to August 2012.
- Main outcome variable: number of firms (up to 1 employee).

Brazilian Monthly Employment Survey (PME):

- Rotating panel 4-8-4, from January 2006 to August 2012.
- Filters: (i) age 21 to 65; (ii) worked at least 20 hours/week; (iii) exclude agriculture, public sector and domestic activities.
- Main outcome variable: indicator of whether the individual is an entrepreneur who contributes to the social security system.

We run two sets of regressions:

- At the region-by-industry level (RAIS)
 - They seek to identify aggregate effects on the number of formal firms (firm entry).
- At the individual level (PME)
 - They seek to identify whether the above effects come from: (i) formalization of informal entrepreneurs; (ii) creation of new small formal businesses.
 - We restrict the sample so that individuals have at least one observation before and one after either IMP1 or IMP2.

In both cases, we exploit variation in eligibility across industries, regions and time to identify these effects in triple (double) differences estimation strategy.

Regressions at the industry-by-state level (RAIS)

At the industry level, our benchmark specification is a standard triple (double) difference estimator:

$$\begin{aligned} \log(N_{\text{srt}}) = & \beta_1 \text{Post1}_{\text{rt}} \times \text{Eligible}_s + \beta_2 \text{Post2}_t \times \text{Eligible}_s + \text{DD Controls} \\ & + \lambda_t + \nu_s + \xi_r + \text{Trends}'_{\text{srt}} \theta + \epsilon_{\text{srt}} \end{aligned}$$

Where N_{srt} is the number of formal firms in industry s , region r in time t .

Regressions at the individual level (PME)

- We do not observe individuals' 5-digit industry; we thus define a potential of treatment variable:

$$\text{Intense}_s = \frac{\sum_{k \in s} (1 [\text{k is eligible}] \times N_k)}{N_s}$$

Where $k \in s$ denotes 5-digit industry within industry s . [► Table](#)

Regressions at the individual level (PME)

- We do not observe individuals' 5-digit industry; we thus define a potential of treatment variable:

$$\text{Intense}_s = \frac{\sum_{k \in s} (1 [\text{k is eligible}] \times N_k)}{N_s}$$

Where $k \in s$ denotes 5-digit industry within industry s . [► Table](#)

- We estimate:

$$Y_{\text{isrt}} = \beta_1 \text{Post1}_{\text{rt}} \times \text{Intense}_s + \beta_2 \text{Post2}_t \times \text{Intense}_s + \text{DD Controls} + Z_{\text{isrt}} + \phi_i + \epsilon_{\text{isrt}}$$

Where $Y_{\text{isrt}} = 1$ if the individual is a formal entrepreneur with up to one employee.

Identification

- It's a DID.
- Industry eligibility:
 - List of eligible industries from SIMPLES, introduced a decade before.
 - Parallel pre-trends + eligibility not correlated with industries' observables nor pre-trends. ▶ pre-trends ▶ regressions
- Some particular issues:
 - IMP1: Main concern is the 2009 economic shock → triple interaction.
 - IMP2: No relevant economic shocks, identification hinges on industry and time variations.

Results: effects on the number of formal firms (RAIS)

Dependent Variable: Log (# Formal Firms)		
	All Regions	PME Regions
	(1)	(2)
IMP1	-0.007 (0.008)	0.017 (0.015)
IMP2	0.048 (0.007)***	0.038 (0.009)***
Observations	379,030	114,575
R-squared	0.831	0.743
Sample	≤ 1 Employee	≤ 1 Employees

Where does the formalization effect come from?

- We now analyze individual transition rates into formal entrepreneurship from:
 - ① Intensive margin: informal entrepreneurship.
 - ② Extensive margin: formal and informal employee; unemployment.
- We also investigate:
 - ① Heterogeneous effects across entrepreneurs' income levels.
 - ② Timing of the effects.

Effects on formalization of informal entrepreneurs

Dependent Variable: Formal Entrepreneur (0/1)			
	(1)	(2)	(3)
IMP1	0.013 (0.035)	0.016 (0.033)	0.009 (0.028)
IMP2	0.067 (0.031)**	0.068 (0.032)**	0.044 (0.012)***
Number of individuals	5,231	5231	5231
R-squared	0.062	0.063	0.073
Time, Region, Industry and Entrepreneur FE	Yes	Yes	Yes
Control Economic Fluctuation	No	Yes	Yes
Control Convergence	No	No	Yes

$$\text{Effect} = \beta_{\text{IMP2}} \times E[\text{Intense}_{\text{si}}] = 0.044 \times 0.43 \times 100 = 1.9\text{p.p.}$$

$$\text{Formality Rate} = 20\%.$$

Heterogeneity by Income

Dependent Variable: Formal Entrepreneur (0/1)					
	(1)	(2)	(3)	(4)	(5)
IMP1	0.012 (0.028)	-0.027 (0.065)	0.016 (0.035)	0.156 (0.100)	-0.154 (0.168)
IMP2	0.044 (0.012)***	-0.010 (0.019)	0.020 (0.025)	0.042 (0.022)*	0.131 (0.031)***
Number of Individuals	5,231	1,309	1,307	1,308	1,307
R-squared	0.073	0.090	0.047	0.077	0.117
Mean(<i>Intense_s</i>)	0.435	0.498	0.414	0.361	0.472
Sample	Benchmark	1st Quartile	2nd Quartile	3rd Quartile	4th Quartile

$$\begin{aligned}
 \text{Effect}_{4\text{th}} &= \beta_{4\text{th}} \times E[\text{Intense}_{si} | i \in 4\text{th Quartile}] \\
 &= 0.131 \times 0.47 \times 100 = 6.1\text{p.p.}
 \end{aligned}$$

Formality Rate = 55.3%.

Effects on the extensive margin

Dependent Variable: Formal Entrepreneur (0/1)				
	(1)	(2)	(3)	(4)
IMP1	-0.005 (0.003)	0.016 (0.033)	0.023 (0.028)	0.012 (0.154)
IMP2	0.001 (0.002)	-0.005 (0.022)	0.008 (0.009)	-0.040 (0.032)
Number of Individuals	15,820	3,443	1,269	2,176
R-squared	0.003	0.026	0.037	0.097
Transition from:	Formal Employee	Informal Employee	Unemployed	Formal Entrepren.

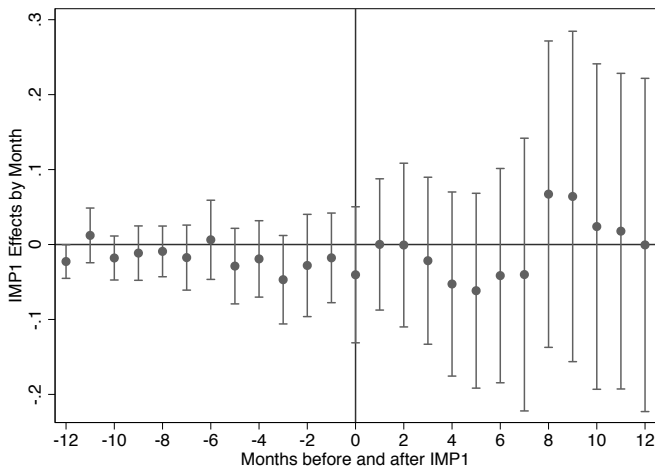
Notes: Significant at the *** 1 percent, **5 percent, and *10 percent levels.

- Falsification tests: we change the implementation dates, considering lags of one, two and three years relatively to the actual start date. ▶ Regressions
- We do not restrict our sample to individuals who have at least one observation before and one after either IMP1 or IMP2. ▶ Regressions
- We run the same specification as in the benchmark, but considering entrepreneurs with at most five employees. ▶ Regressions
- We restrict the sample using the annual income threshold (R\$36k until Nov/11 and R\$60k after that). ▶ Regressions
- Different definitions of treatment variable: ▶ Regressions

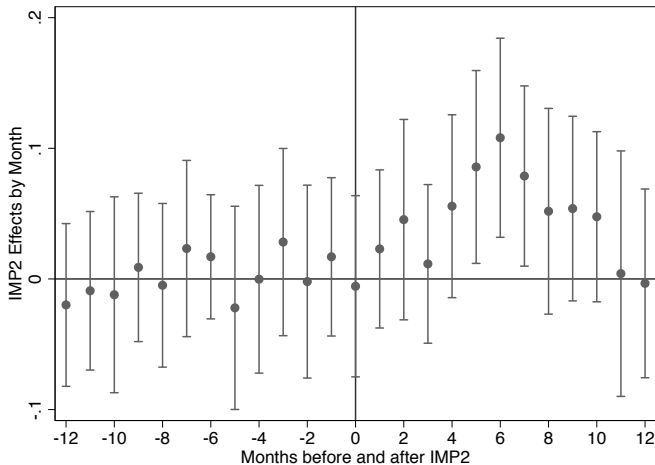
$$T_i^{50} = 1 [\text{Intense}_{s,i} \geq 0.5]$$

$$T_i^{90} = \begin{cases} 1 & \text{if } \text{Intense}_{s,i} \geq 0.9 \\ 0 & \text{if } \text{Intense}_{s,i} \leq 0.1 \end{cases}$$

Timing of the effects: IMP1

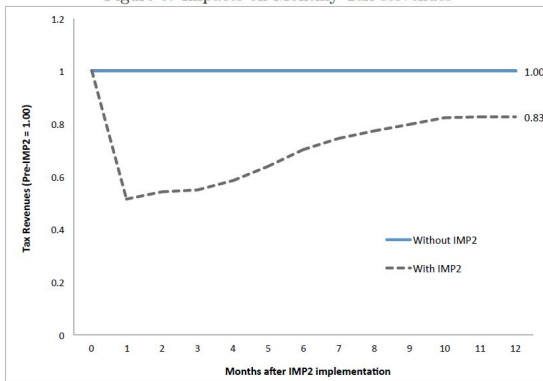


Timing of the effects: IMP2



Was the program cost effective?

Figure 6: Impacts on Monthly Tax Revenues



Notes: The black dotted line (with IMP2) is computed using the actual mandated tax expenditures defined in IMP2, times the counterfactual number of formal entrepreneurs, N_t^f , as described in the text: $N_t^f = \hat{\beta}_t N_{t-1}^i + N_{t-1}^f$, where we use all point estimates $\hat{\beta}_t$, regardless of their statistical significance. The blue solid line (without IMP2) is computed holding the pre-IMP2 stock of formal entrepreneurs constant at \bar{N}^f and using the mandated tax expenditures defined in IMP1.

Why don't we observe stronger formalization effects?

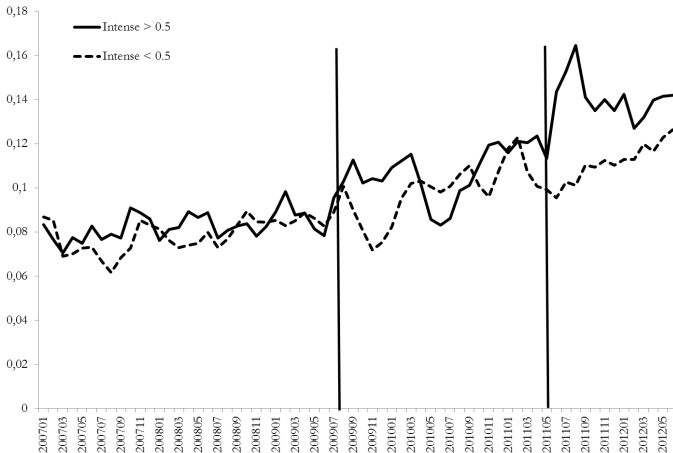
Dep. Variable: Entrepreneur's Log(Income)						
	(1)	(2)	(3)	(4)	(5)	(6)
IMP2	0.010 (0.025)	-0.041 (0.057)	-0.053 (0.065)	0.028 (0.033)	0.014 (0.059)	— —
Formal (Dummy)	— —	— —	— —	— —	— —	0.125 (0.347)
First Stage F-Stat	—	—	—	—	—	12.6
Observations	13,084	2,840	3,212	3,472	3,560	13,084
R-squared	0.027	0.141	0.086	0.072	0.063	
Number of Individ.	2,630	670	634	663	663	2,630
Sample	All	1st Quart.	2nd Quart.	3rd Quart.	4th Quart.	All
Model	OLS	OLS	OLS	OLS	OLS	2SLS

- Reducing taxes has a positive and sizable effect on formal firm creation and on the share of formal entrepreneurs:
 - We find an avg. effect of 10% increase in the baseline formalization rate.
 - The increase in formal entrepreneurship comes from the formalization of informal firms, not due to the creation of new formal businesses.
 - Formalization effects are stronger for entrepreneurs in upper income quartiles, who perceived the highest reductions in tax rates.
 - We find transitory effects.
- Similarly to previous studies, we find that reducing registration costs alone is not enough to induce greater formalization.
- The implied formalization elasticity is low and the program lead to net losses in tax revenues.

Tabela: Variable Intense: Tabulation for Different Samples

Var. Intense	Frequencies for Different Sub-samples			
	Informal Entrepren.	Inf. Employee	Form. Employee	All
0.000	0.219	0.197	0.192	0.193
0.090	0.262	0.167	0.180	0.200
0.126	0.015	0.038	0.058	0.046
0.375	0.015	0.016	0.015	0.016
0.389	0.005	0.009	0.011	0.010
0.604	0.062	0.058	0.067	0.067
0.699	0.015	0.040	0.041	0.033
0.751	0.001	0.009	0.024	0.016
0.822	0.003	0.011	0.017	0.012
0.838	0.056	0.118	0.138	0.119
0.895	0.089	0.051	0.011	0.034
0.907	0.043	0.058	0.036	0.043
0.917	0.002	0.002	0.003	0.003
0.951	0.018	0.037	0.018	0.021
1.000	0.196	0.189	0.188	0.187

Pre-trends: Informal-Formal 12 months transitions



Regressions at the Industry Level (PNAD 2009)

Dependent variable: Eligible industry in 2009			
$\Delta_{2009-2008}(\text{Log Wages})$	-0.389 (0.289)		-0.405 (0.288)
$\Delta_{2009-2008}(\text{Formality Rate})$		-0.110 (0.738)	0.115 (0.750)
Share Male	-0.180 (0.309)	-0.186 (0.313)	-0.180 (0.310)
Avg. Schooling (Yrs)	0.070 (0.052)	0.067 (0.051)	0.070 (0.051)
Share White	-1.207 (0.844)	-1.154 (0.832)	-1.225 (0.827)
Avg. Age	-0.014 (0.034)	-0.014 (0.035)	-0.014 (0.034)
Share Service	0.265 (0.250)	0.263 (0.249)	0.264 (0.251)
Share Manufacturing	0.621 (0.195)***	0.605 (0.189)***	0.624 (0.187)***
Number of Industries	118	118	118
R-squared	0.287	0.281	0.287

Eligible vs. Non-Eligible Industries: Descriptive statistics

	Non-Eligible Mean	Eligible Mean	Difference P-value
Size distribution (shares)			
No employees	0.815	0.791	0.000
One employee	0.055	0.051	0.113
2 to 5 employees	0.079	0.098	0.000
6 to 10 employees	0.024	0.031	0.000
11+ employees	0.028	0.029	0.415
Formal firms (has CNPJ) [†]	0.301	0.324	0.019
Contributes to S.S. (share)	0.300	0.314	0.005
Industry Composition			
Service	0.072	0.215	0.000
Commerce	0.456	0.142	0.000
Manufacturing	0.020	0.254	0.000
Transportation	0.053	0.094	0.000
Construction	0.313	0.000	0.000
Log-Earnings	7.0	7.0	0.293
Male (share)	0.732	0.543	0.000
Avg. schooling (years)	8.4	9.2	0.000
White (share)	0.526	0.552	0.000
Age	43.1	42.6	0.000
Hours per week	43.9	43.8	0.512

Effects on rates at the industry-by-region level (PME)

[illegible]

Robust standard errors in parentheses (cluster(industry) for all specs)

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Source: PME, monthly data from 2006 through 2012.

Robustness

- Falsification tests: we change the implementation dates, considering lags of one, two and three years relatively to the actual start date. ▶ Regressions
- We do not restrict our sample to individuals who have at least one observation before and one after either IMP1 or IMP2. ▶ Regressions
- We run the same specification as in the benchmark, but considering entrepreneurs with at most five employees. ▶ Regressions
- We restrict the sample using the annual income threshold (R\$36k until Nov/11 and R\$60k after that). ▶ Regressions
- Different definitions of treatment variable: ▶ Regressions

$$T_i^{50} = 1 [\text{Intense}_{s,i} \geq 0.5]$$

$$T_i^{90} = \begin{cases} 1 & \text{if } \text{Intense}_{s,i} \geq 0.9 \\ 0 & \text{if } \text{Intense}_{s,i} \leq 0.1 \end{cases}$$

Effects on Formalization of Informal Entrepreneurs: Full sample

Dependent Variable: Formal Entrepreneur (0/1)			
	(1)	(2)	(3)
IMP1	-0.009 (0.019)	-0.011 (0.020)	-0.007 (0.019)
IMP2	0.061 (0.031)*	0.061 (0.031)*	0.039 (0.011)***
Number of individuals	61,148	61,148	61,148
R-squared	0.029	0.029	0.035
Time, Region, Industry and Entrepreneur FE	Yes	Yes	Yes
Control Economic Fluctuation	No	Yes	Yes
Control Convergence	No	No	Yes

Notes: Significant at the *** 1 percent, **5 percent, and *10 percent levels.

Formalization of Informal Entrepreneurs: Entrepreneurs with up to 5 Employees

Dependent Variable: Formal Entrepreneur (0/1)			
	(1)	(2)	(3)
IMP1	0.002 (0.041)	0.007 (0.038)	-0.000 (0.032)
IMP2	0.067 (0.034)*	0.068 (0.034)*	0.043 (0.013)***
Number of ind	5,231	5,231	5,231
R-squared	0.064	0.064	0.075
FE	Yes	Yes	Yes
Economic Trends	No	Yes	Yes
Convergence	No	No	Yes

Notes: Significant at the *** 1 percent, **5 percent, and *10 percent levels.

Formalization of Informal Entrepreneurs: Using Annual Income

► Back

Dependent Variable: Formal Entrepreneur (0/1)			
	(1)	(2)	(3)
IMP1	0.007 (0.024)	0.010 (0.024)	0.005 (0.024)
Number of indiv.	2,639	2,639	2,639
R-squared	0.038	0.039	0.042
IMP2	0.054 (0.024)**	0.053 (0.024)**	0.034 (0.010)***
Number of indiv.	2,318	2,318	2,318
R-squared	0.071	0.072	0.079
FE	Yes	Yes	Yes
Economic Trends	No	Yes	Yes
Convergence	No	No	Yes

Notes: Significant at the *** 1 percent, **5 percent, and

Formalization of Informal Entrepreneurs: Different treatment variables

Dependent Variable: Formal Entrepreneur (0/1)						
	Treatment = T ⁵⁰			Treatment = T ⁹⁰		
	(1)	(2)	(3)	(4)	(5)	(6)
IMP1	-0.012 (0.017)	-0.013 (0.017)	-0.012 (0.018)	0.009 (0.020)	0.010 (0.020)	0.016 (0.020)
IMP2	0.066 (0.027)**	0.065 (0.026)**	0.040 (0.011)***	0.052 (0.026)*	0.051 (0.025)*	0.047 (0.021)**
Number of indiv.	54,227	54,227	54,227	38,346	38,346	38,346
R-squared	0.030	0.030	0.036	0.023	0.023	0.025
FE	Yes	Yes	Yes	Yes	Yes	Yes
Trends	No	Yes	Yes	No	Yes	Yes
Convergence	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Significant at the *** 1 percent, **5 percent, and *10 percent levels.

