UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

Essays in Applied Economics

Ensaios em Economia Aplicada

Danilo Paula de Souza

Orientador: Prof. Dr. Mauro Rodrigues Júnior
Co-orientador: Prof. Dr. Marcos Yamada Nakaguma

São Paulo - Brasil
2020

Prof. Dr. Vahan Agopyan
Reitor da Universidade de São Paulo

Prof. Dr. Fabio Frezatti
Diretor da Faculdade de Economia, Administração e Contabilidade

Prof. Dr. Jose Carlos de Souza Santos
Chefe do Departamento de Economia

Prof. Dr. Ariaster Baumgratz Chimeli
Coordenador do Programa de Pós-Graduação em Economia

DANILO PAULA DE SOUZA

Tese apresentada ao Departamento de Economia da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo como requisito parcial para a obtenção do título de Doutor em Ciências.

Área de concentração: Teoria Econômica

Orientador: Prof. Dr. Mauro Rodrigues Júnior
Co-orientador: Prof. Dr. Marcos Yamada Nakaguma

Versão Corrigida
(versão original disponível na Faculdade de Economia, Administração e Contabilidade)

São Paulo - Brasil
2020

Catalogação na Publicação (CIP)
Ficha Catalográfica com dados inseridos pelo autor

Souza, Danilo Paula de.
Essays in Applied Economics / Danilo Paula de Souza. - São Paulo, 2020.
135 p.

Tese (Doutorado) - Universidade de São Paulo, 2020.
Orientador: Mauro Rodrigues Junior.
Co-orientador: Marcos Yamada Nakaguma.

1. Crescimento econômico e qualidade da educação. 2. Economia política. 3. Propaganda política negativa. 4. Economia do trabalho. 5. Custos trabalhistas e comportamento das firmas. I. Universidade de São Paulo. Faculdade de Economia, Administração e Contabilidade. II. Título.

A todos aqueles que insistem em se perguntar
o porquê do porquê das coisas.

AGRADECIMENTOS

Começo esta seção de agradecimentos fazendo o inevitável, reconhecendo todo apoio e suporte que recebi da minha família ao longo de todo esse período de pós-graduação que se estende para além do doutorado. Agradeço meus irmãos, Claudio e Wanderley, e minha cunhada, Larissa, por todos os almoços de domingo que tornaram minhas semanas um pouco mais leves. À minha mãe, Leila, um agradecimento especial por todo o suporte e amor incondicional desde sempre. E ao meu pai, Francisco, que em vida me ensinou muito sobre trabalho, disciplina e a enorme importância de ser corintiano. Amo vocês, do fundo do meu coração.

Agradeço os meus eternos amigos da época do colégio, sem os quais nada disso seria possível e sequer teria sentido. Nelson, Beatriz, Gabriela, Bruna, Douglas, Fabio, Gabriel, Jaqueline, Letícia, Marcelo, Victor e Vitor, vocês foram e continuam sendo importantíssimos para mim e parte relevante de todas minhas conquistas. Muito obrigado!

Um agradecimento especial aos colegas de mestrado e doutorado, que foram muito importantes na confecção desse trabalho e, principalmente, no momento de descontração pós-almoço na salinha da pós-graduação. Obrigado a todos vocês, em especial Raphael Bruce, Rafael Tavares, Tiago Ferraz, Ricardo Sabbadini, Bruno Palialol, Raí Chicoli.

Não poderia deixar de agradecer ao meu orientador Mauro Rodrigues por todo o apoio acadêmico desde o mestrado e ao meu coorientador Marcos Nakaguma por sempre se mostrar disponível e pelos valiosos ensinamentos. Sou grato também ao professor Raphael Corbi, com quem aprendi muito ao longo dos últimos dois anos.

Por fim, mas não menos importante, agradeço à parceria e apoio incondicional da minha namorada, Camila. Sua contribuição vai muito além das sugestões em relação à tese. Com você percebi que de fato o tempo é relativo e que um relacionamento leve e construtivo para os dois lados é sim possível. Te amo.

Agradeço o apoio financeiro da CAPES e CNPq durante parte importante desse projeto.

“ **Somewhere, something incredible is waiting to be known**”

Carl Sagan

RESUMO

Esta tese de doutorado é composta por três artigos que abordam temas não relacionados entre si, mas diretamente ligados a três grandes áreas da pesquisa econômica: crescimento econômico, economia política e economia do trabalho. No primeiro artigo busca-se reconciliar as predições teóricas da literatura de crescimento econômico acerca da relação entre aumento de capital humano e PIB per capita com as evidências empíricas dos últimos 40 anos. O modelo de gerações sobrepostas proposto sugere a existência de um *trade-off* entre quantidade e qualidade de educação que surge através da escolha ocupacional dos indivíduos. Dessa forma, o impacto positivo de um aumento nos anos de estudo de um país sobre seu estoque de capital humano seria compensado no longo-prazo por uma redução do capital humano médio dos professores. O artigo contribui com uma literatura majoritariamente empírica que encontra na qualidade da educação uma variável relevante para explicar a ausência de correlação entre aumento de anos de estudo e crescimento econômico. O segundo artigo investiga como as estratégias de campanha eleitoral em relação à propaganda negativa são afetadas pelas características da disputa, dos candidatos participantes e das instituições eleitorais. Para entender determinantes e incentivos, constrói-se um modelo estilizado de decisão em relação ao tipo de campanha que prevê maior probabilidade de haver propaganda negativa por parte de um candidato à medida que o suporte inicial do oponente cresce e a competição eleitoral se reduz, além de prever diferenças significativas na probabilidade de ataque entre candidatos que disputam a cadeira em cidades com apenas um turno e em cidades com a possibilidade de dois turnos. As predições do modelo são testadas empiricamente com uma base de dados única de processos judiciais sobre o tipo de propaganda política realizada pelos candidatos nas eleições municipais brasileiras de 2012 e 2016. Os resultados corroboram as predições do modelo e documentam aspectos do processo de decisão dos candidatos políticos em relação à realização de propaganda negativa. O terceiro e último artigo busca documentar a estrutura, o funcionamento e os efeitos da justiça trabalhista brasileira no comportamento das firmas. Estimações de forma reduzida utilizando o universo de processos trabalhistas do maior tribunal do país de 2008 a 2013 mostram que um maior custo trabalhista implica em uma redução da taxa de crescimento do nível de emprego e do salário médio de novos contratados, além de uma redução da probabilidade de sobrevivência das firmas. Constrói-se então um modelo de *search-matching*, calibrado com os dados brasileiros, com a finalidade de avaliar como o nível de emprego e outras variáveis reagiram às mudanças introduzidas pela Reforma Trabalhista de 2017 ao repassar os custos do processo para a parte perdedora. De maneira similar aos resultados de forma reduzida, a análise de contrafactual mostra que ao reduzir as regulações trabalhistas e os custos esperados com processos trabalhistas, a reforma trabalhista pode ter aumentado o nível de emprego da economia.

Palavras-chaves: crescimento econômico, qualidade da educação, convergência condicional, propaganda política, propaganda negativa, instituições eleitorais, justiça trabalhista, custos trabalhistas, comportamento das firmas

ABSTRACT

This doctoral dissertation consists of 3 articles that address topics not related to each other, but directly linked to three major areas of economic research: economic growth, political economy, and labor economics. The first article tries to reconcile the theoretical predictions of economic growth literature on the relationship between human capital increase and GDP per capita with the empirical evidence of the last 40 years. The overlapping generations model suggests the existence of a trade-off between quantity and quality of education that arises through the occupational choice of individuals. Thus, the positive impact of an increase in a country's average years of schooling on its human capital stock would be compensated by a decrease in the average human capital of teachers in the long-run. The article contributes with a mostly empirical literature that finds in the education quality an important variable to explain the lack of correlation between increases in average years of schooling and economic growth. The second article analyzes how electoral campaign strategies regarding negative advertising are affected by the dispute and candidates' characteristics, and by electoral institutions. In order to understand determinants and incentives, we build a stylized model of decision on the campaign type which predicts that a candidate is more likely to attack his opponent as the opponent's initial support rises and the electoral competition decreases. The model also predicts significant differences in the likelihood of attack among candidates running for the seat in municipalities with a single-ballot system and municipalities with a dual-ballot system. Model predictions are empirically tested using an unique dataset of lawsuits on the type of political advertising carried out by candidates in the 2012 and 2016 Brazilian municipal elections. Results confirm the model predictions and document aspects of the political candidates' decision to conduct negative advertising. The third and last article documents the structure, functioning, and the effects of the Brazilian labor justice on firms' behavior. Reduced form estimations using the universe of labor lawsuits filed in the country's largest labor court from 2008 to 2013 show that a greater labor cost leads to a decrease in the growth rate of firm's employment and in the growth rate of the average wage of new hires. There is also a negative effect on the likelihood of firms' survival. A search-matching model is built and calibrated with Brazilian data in order to assess how the level of employment and other variables reacted to the changes introduced by the 2017 Labor Reform by shifting lawsuit costs to the losing party. Similar to the reduced form evidence, the counterfactual analysis shows that by reducing labor regulations and the expected costs with labor lawsuits, the Labor Reform may have increased the employment level of the economy.

Key-words: economic growth, education quality, conditional convergence, political advertising, negative advertising, electoral institutions, labor justice, labor costs, firm behavior

TABLE OF CONTENTS

1 CONDITIONAL CONVERGENCE AND THE DYNAMICS OF TEACHERS' QUALITY

Com Mauro Rodrigues

1.1 Introduction

The widely accepted idea of human capital as one major driver of economic growth and, thus, essential to individuals well-being has led to an historically unprecedent expansion in enrollment rates and in average years of schooling of most countries in the last 40 years. Still, cross-country data hasn't been clear on the importance of this educational attainment evolution to explain economic growth in the period. In the last decades, many have tried to understand these evidence empirically, but few have tried to reconcile it with the theory. Our goal is to fill this gap by proposing a model with endogenous education quality.

Cross-country data indeed shows that the vast majority of countries have approached the United States in education attainment in the past decades, but also shows something closer to a stable GDP distribution in the same period (figure 1.1.1)[1.1.1]. The widely cited work of Pritchett (2001) brings more evidence on this matter using cross-country regressions with level-on-level and growth-on-growth specifications, despite the contrary predictions of the augmented Solow model and the empirical findings on its favor in the early 90's (BARRO, 1991; MANKIW; ROMER; WEIL, 1992).

Delgado, Henderson e Parmeter (2014) present an overview of the literature findings in this cross-country setting. More importantly, however, the authors estimate a non-parametric model which suggests that this apparently statistically insignificant relation between human capital and growth is indeed positive and

[1.1.1] Yet, there were some exceptions: east asian countries like Singapore and Korea indeed experienced rapidly income and productivity growth in this period while other countries like Equatorial Guinea benefited from mineral and oil rents.

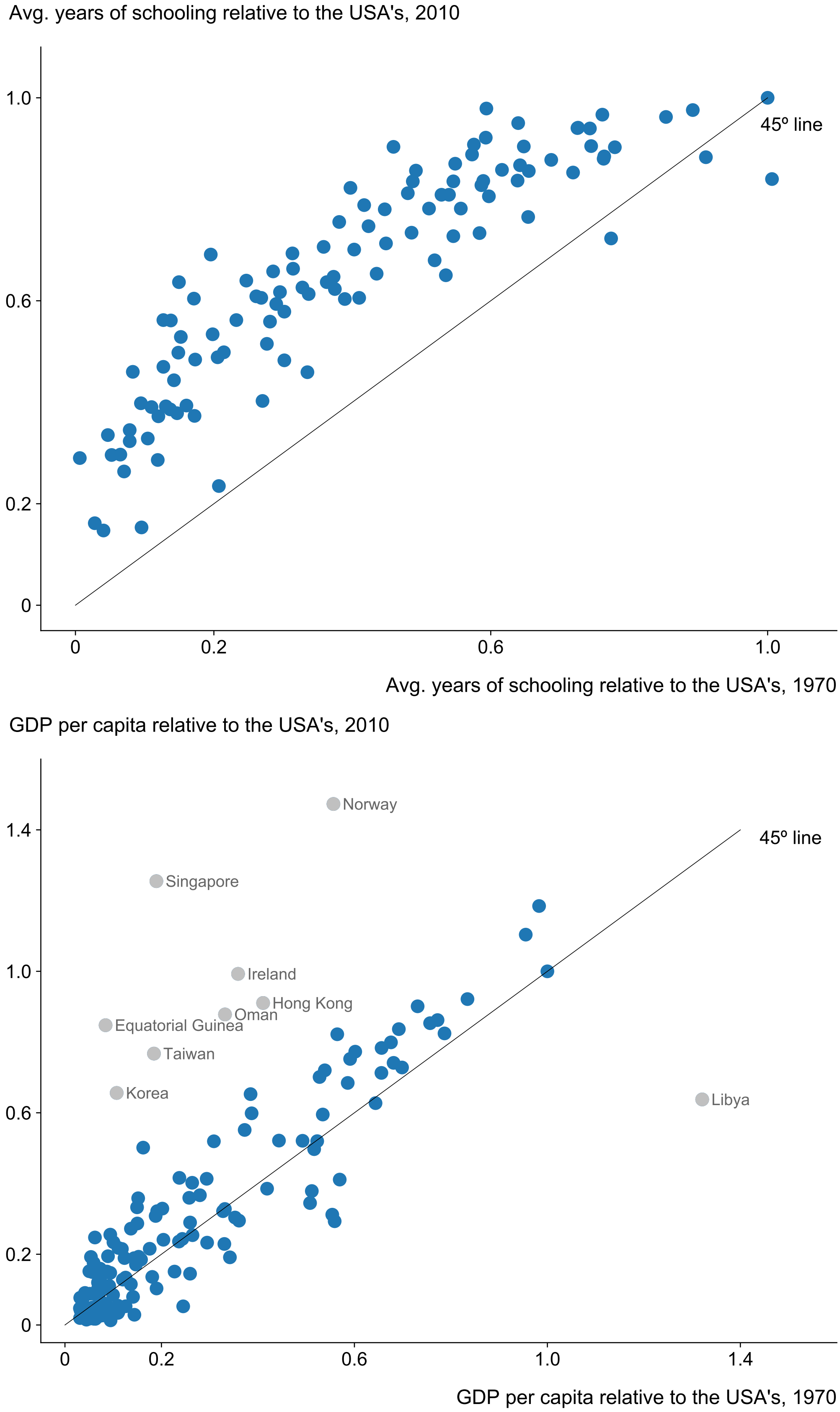


Figura 1.1.1 – Evolution of average years of schooling and GDP per capita, 1970-2010. *Source: Barro-Lee Database (average years of schooling) and The Maddison Project (GDP per capita).*

significant when measures of education quality are taken into account. Similarly, Hanushek (2013) shows that when controlling for cognitive skills acquired during education process, the initial value of years of schooling does not have any effect on average annual growth rate of GDP per capita in the 1960-2010 period[1.1.2]. These findings seem to point to education quality as one important missing factor in order to reconcile what we observe in the educational attainment and GDP data. Nevertheless, a macroeconomic model of how education quality and quantity relate and how this interaction could help explain the cross-country data is still missing.

Existing growth models that account for education quality, such as Tamura (2001) and Manuelli e Seshadri (2014), usually take it as exogenous or as a policy variable that responds to an arbitrary central planner objective function. This paper contributes by proposing an overlapping generations model with occupational choice in which education quality is endogenous and measured by the average human capital of individuals that have chosen the teacher career[1.1.3]. This correspondence between education quality and teachers quality is in accordance with the economics of education literature. Cheety, Friedman e Rockoff (2014) show that the quality of teaching is of great importance in the students' future through an estimation of long-term impacts of high value-added teachers. Moreover, Rivkin, Hanushek e Kain (2005) and Rockoff (2004) show that, when dealing with the determinants of students' academic achievement, teacher fixed effects are very important predictors of student outcome[1.1.4].

Since human capital is assumed to be compulsory and taken in public schools in our model, teachers are hired and paid by the government. The amount spent in education is restricted by the government budget constraint, which also define the

[1.1.2] Schoellman (2012), Hanushek e Woessmann (2012), and Manuelli e Seshadri (2014) are more recent contributions to the literature that assess the role of human capital quality on explaining output per worker and wages variability.

[1.1.3] Gilpin e Kaganovich (2012) build a similar model to address the evolution of teacher-student ratio and teacher quality in the post-WWII US economy.

[1.1.4] According to Rivkin, Hanushek e Kain (2005), *"the results demonstrate quite clearly that the observable school and teacher characteristics explain little of the between-classroom variation in achievement growth despite the fact that a substantial share of the overall achievement gain variation occurs between teachers. Importantly, even though the sample includes just schools with a single teacher per grade, the inclusion of school rather than teacher fixed effects reduces the explanatory power by over forty percent"* (p. 421).

government's capacity to attract individuals with higher levels of innate ability to the teacher career. A tradeoff between quantity and quality of education arises in this environment given that, by hiring more teachers, the government would have to pay less for each teacher. A similar tradeoff that relies on an exogenous tax rate also appears in Gilpin e Kaganovich (2012), but one could argue that by simply raising taxes the government would be able to mitigate the problem. To address this issue we propose a mechanism in which the tax rate is endogenous and result of the median voter optimization problem.

The model is calibrated using data from a set of 9 Latin America countries and the microeconometric evidence on education rates of return. We then simulate the model in order to evaluate the aggregate dynamic effects of the evolution in average years of schooling in the 1970-2010 period. The increase in the average years of schooling in these countries from 3.98 in 1970 to 8.18 in 2010 required a proportional increase in the fraction of teachers in the labor force from 2.95% to 6.06% in the model. This higher number of teachers reduced the attractiveness of the teacher carrer which also led to a more than 50% drop in the new steady-state value of education quality. As a result of this quality drop, the long-run impact that improving the average number of years of schooling had on the stock of human capital was close to zero. Sensitivity analysis also showed that the long-run relevance of a lower education quality was greater as the country became more educated.

Our paper is closely related to a mostly empirical literature that puts in doubt the importance of education in explaining economic growth. Since the contributions of Benhabib e Spiegel (1994), Bils e Klenow (2000) and Pritchett (2001), papers in this literature have presented some potential explanations to the zero correlation between education and growth: education data of poor quality (TEMPLE, 2001; FUENTE; DOMENECH, 2006; COHEN; SOTO, 2007), model misspecification (KALAITZIDAKIS et al., 2001; DELGADO; HENDERSON; PARMETER, 2014), result of the more educated individuals being alocated in low productivity sectors (ROGERS, 2008; SCHUNDELN; PLAYFORTH, 2014), and also that education could be relevant to economic growth only after reaching some threshold (AHSAN; HAQUE, 2017). Our main contribution, however, is theoretical by presenting a simple and tractable model, with closed form solutions, which may help to reconcile the cross-country data on average years of schooling and GDP *per capita*. In spite of not being the first model to present an endogenous trade-off between education quantity and quality (GILPIN; KAGANOVICH, 2012) or showing the importance of education quality in the GDP *per capita* convergence process (TAMURA, 2001), our paper is the first, at the best of our knowledge, to do both and also endogeneize the tax income rate in order to understand the macroeconomic implications of a country's policy to increase his population's average years of schooling alone.

The rest of this paper is organized as follows. Section 2.2 presents all the model steps and the problem that the median voter solves. Section 1.3 draws some comparative dynamics to help us understand the underlying mechanisms through which aggregate variables adjust to policy shocks. In sections 1.4 and 1.5 we calibrate and simulate the model in order to reproduce what we observed in the last decades in the cross-country data. Section 1.6 concludes.

1.2 Model

Consider an overlapping-generations economy in which growth is the result of human capital accumulation. Individuals live for two periods and have to decide between a job as a teacher and a job in the private market[1.2.1] in the second period. Teachers' wage is limited by the government budget constraint, which we assume to be balanced in every period. A single homogenous good is produced using the stock of human capital, which in turn depends on the average human capital of teachers. The income tax rate is result of the median voter optimization problem.

1.2.1 Individuals

As in Galor e Moav (2000) and Galor e Moav (2004), there is a continuum of measure 1 of individuals being born in every period. Individuals accumulate

[1.2.1] One can think these jobs in the private market as all the other occupations that directly contributes to the production of goods and not only to the production of human capital, as teachers.

human capital in the first period and choose a career, work, and consume in the second period. Accumulation of human capital is compulsory and taken in public schools in the first period. Individuals are assumed to derive utility from its own consumption and from the human capital of their offspring:

$$u_t^i = \log(c_t^i) + \phi \log(h_{t+1}^i) \tag{1.2.1}$$

where $\phi > 0$ measures how altruistic are the parents. Income is the result of the occupational choice since we assume that the private market pays a higher wage for individuals with more human capital whereas teachers' wage is the same for all hired individuals[1.2.2]. Hence, there is a budget constraint for private market workers and a different one for teachers, which are given, respectively, by

$$c_t^i \leq (1-\tau) w_t^M h_t^i$$

$$c_t^i \leq (1-\tau) w_t^T$$

Individuals, however, may differ in their innate ability a^i, which is assumed to be distributed uniformly over the interval $[0, \bar{B}]$ (GALOR; MOAV, 2000; GILPIN; KAGANOVICH, 2012). Innate ability is a component of human capital accumulation function along with the proportion of first period spent in school s_{t-1}, and education quality h_{t-1}^T. Thus,

$$h_t^i = Z a^i (s_{t-1})^\eta (h_{t-1}^T)^v \tag{1.2.2}$$

where $\eta, v \in (0,1)$, and Z is a constant productivity term. We assume that s_t is a function of the proportion of teachers in the labor force and other variables that are also not under the control of individuals[1.2.3]. Furthermore, we assume that s_t increases with the proportion of teachers in the labor force, which is the same as

[1.2.2] This could be the result of teachers being hired on the basis of public competition in which only the most skilled applicants get the job.

[1.2.3] By assuming that s_t is not result of individuals' choice, we are considering the case in which something like *compulsory-schooling laws* (ACEMOGLU; ANGRIST, 2000) are in place.

assuming that an increasing number of teachers in the economy will lead to more teachers available for students and therefore to the feasibility of spending more time in school. In essence, a higher number of teachers is a sufficient condition to increase average years of schooling *ceteris paribus*.

Given this structure of preferences and human capital accumulation, individuals will choose what career to take as a response to their level of human capital and so to their level of innate ability. The analysis of the individual who is indifferent between choosing the teacher career and the private market career gives us the innate ability threshold $a_t^\star$ which define who would apply to be a teacher and who would not.

By making the utility level of a private market worker equal to the utility level of a teacher we reach the innate ability threshold $a_t^\star$:

$$a_t^\star = \frac{w_t^T}{w_t^M} \cdot \frac{1}{Z.(s_{t-1})^\eta.(h_{t-1}^T)^v} \tag{1.2.3}$$

Therefore, everyone who has a level of innate ability such that $a^i > a_t^\star$ will choose to be part of the private market in t. Every individual i with $a^i < a_t^\star$ would prefer to be a teacher, but we impose as a restriction that the fraction of teachers in the labor force is constant and equal to θ for every t. This restriction can be thought as a physical restriction: for a given educational infrastructure (e.g., a given number of schools), the government must hire θ teachers to make its educational system work properly. It can not hire less than θ and would not be able to accommodate more than θ. In addition, these θ teachers are hired in a proccess that occurs in every period and offers only θ jobs, which, through competition among all applicants, are filled by the θ most qualified individuals. Figure 1.2.1 shows what this assumption implies: all individuals with a sufficiently low level of innate ability will end up working in the private market in spite of their initial desire of being teachers.

Assumption 1. *In order to have a proportion of θ teachers in the population for every t, $a_t^\star$ must lie inside innate ability interval such that $\theta\bar{B} < a_t^\star < \bar{B}$.*

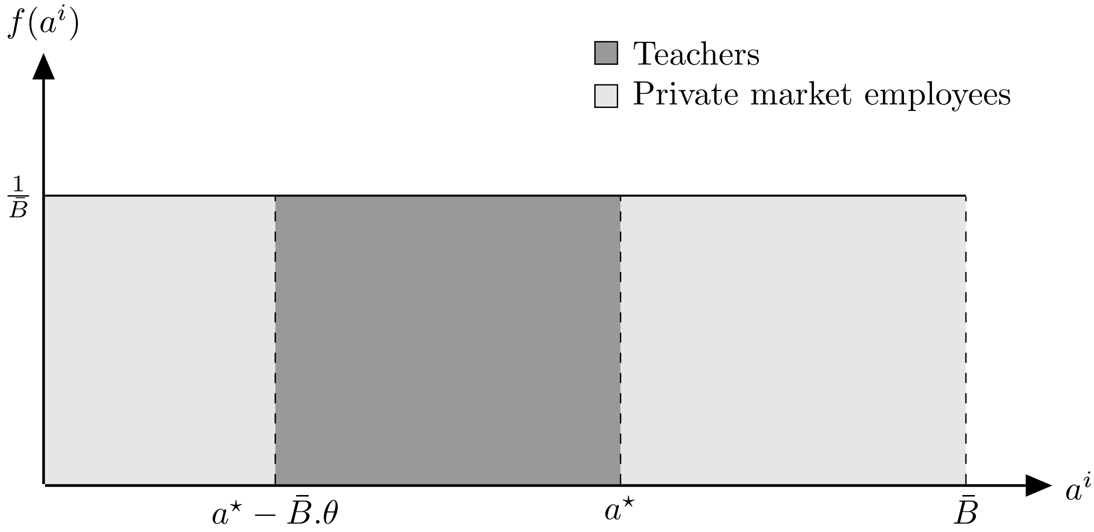


Figura 1.2.1 – Probability density function of innate ability. The dark grey area is the proportion θ of teachers in the population. Every individual whose innate ability is outside this dark grey area will be a private market employee.

1.2.2 Production of final output

There is a single homogeneous good being produced every period according to a constant-returns-to-scale production function. This technology uses the stock of human capital of private market workers as the only input. The output produced in t is described by:

$$Y_t = F(H_t) = A_t \cdot H_t \tag{1.2.4}$$

where H_t is the stock of human capital and A_t is the exogenous total factor productivity (TFP).

The representative firm operates in a perfectly competitive environment. Taking the wage rate of private market employees w_t^M as given, the producer in t choose the level of H_t so as to maximize profits. As result of this optimization, w_t^M is set to equal the marginal productivity of human capital, which is also equal to the TFP level A_t.

1.2.3 Government

Teachers are hired by a government that collects a τ fraction of all individuals labor income and spends all of its tax revenue paying teachers. The government budget constraint is given by:

$$\tau.w_t^M.H_t + w_t^T.\tau.\theta = \theta.w_t^T$$

where w_t^T is the teachers' wage, which can be set as a function of other variables and parameters so that

$$w_t^T = \frac{1}{\theta} \cdot \frac{\tau}{1-\tau} \cdot Y_t \tag{1.2.5}$$

Notice that in spite off w_t^T being the same for all teachers, it evolves according to the whole economy: teachers' wage will be higher the higher economy's Y_t. Therefore, given the model parameters and government budget constraint, w_t^T will evolve endogenously as will the ratio between w_t^M and w_t^T.

1.2.4 Aggregation and human capital dynamics

The stock of human capital and the average human capital of teachers are a direct result of the innate ability distribution and the innate ability threshold. As showed in figure 1.2.1, every individual with $a^i \in \left[0; a^\star - \bar{B}\theta\right)$ or $a^i \in \left(a^\star; \bar{B}\right]$ will end up working in the private market, whereas everyone with $a^i \in \left[a^\star - \bar{B}\theta; a^\star\right]$ will end up working as a teacher. Thus, we obtain the stock of human capital, H_t, and the average human capital of teachers, h_t^T, by integrating equation (1.2.2) in these innate ability intervals.

$$H_t = \int_0^{a_t^\star - \bar{B}\theta} h_t^i(a^i)dF(a^i) + \int_{a_t^\star}^{\bar{B}} h_t^i(a^i)dF(a^i)$$

$$h_t^T = \frac{1}{\theta}\int_{a_t^\star - \bar{B}.\theta}^{a_t^\star} h_t^i(a^i)dF(a^i)$$

Hence, it follows from (1.2.3)-(1.2.5) (see Appendix A) that

$$H_t = \frac{Z\bar{B}}{2} \cdot [(1+\theta^2)(1-\tau)] \cdot (s_{t-1})^\eta \cdot (h_{t-1}^T)^v \tag{1.2.6}$$

$$h_t^T = \frac{Z\bar{B}}{2\theta} \cdot [1-(1+\theta^2)(1-\tau)] \cdot (s_{t-1})^\eta \cdot (h_{t-1}^T)^v \tag{1.2.7}$$

Since the aggregate human capital of both teachers and private market workers is obtained by taking the integral of individual human capital function over the whole ability interval, we can represent it by the sum $H_t + \theta h_t^T$, which is equal to $\frac{Z\bar{B}}{2}(s_{t-1})^\eta(h_{t-1}^T)^v$. Equations (1.2.6) and (1.2.7) imply that H_t and h_t^T are then constant fractions of the whole population's aggregate human capital given by the model's parameters. Intuitively, this $(1+\theta^2)(1-\tau)$ fraction in H_t and h_t^T represents the incentives that individuals face in the occupational choice process, which can also be represented by the relative wage between private market workers and teachers.

$$\frac{w_t^M}{w_t^T} = \frac{1}{\tau} \cdot \frac{2\theta}{ZB(1+\theta^2)} \cdot \frac{1}{(s_{t-1})^\eta(h_{t-1}^T)^v} \tag{1.2.8}$$

The case of a higher θ is intuitive, given that the same government budget would pay less for each teacher and the teacher career would be less attractive in t, including for those with higher innate ability values[1.2.4]. Therefore, H_t would be higher as result of these most skilled individuals applying for jobs in the private market and moving away from the teacher career, which would also make the average teacher less qualified.

The effect of a higher income tax rate can also be viewed as result of changes

[1.2.4] Take the partial derivative of equation (1.2.8) with respect to θ to see that a higher θ implies a higher relative wage between private market workers and teachers.

in the relative attractiveness of both careers. In spite of having a negative effect on the disposable income of both private market workers and teachers, equation (1.2.8) shows that an increase in τ implies a decrease in the relative wage between private market workers and teachers. This change in the relative wage represents a worsening in the outside options of potential teachers, which in turn leads to an increase in the average quality of education as result of the most skilled individuals applying for jobs as teachers. In line with this intuition, Figlio (1997) presents evidence that the average starting teacher salary is positively related to the average quality of teachers regionally. Moreover, Nagler, Piopiunik e West (2015) show that Florida teachers entering the profession during recessions (when the outside options are worse) have a higher average (value-added) quality.

1.2.5 Median-voter choice on τ

In spite of the different model framework, we follow the idea of an income tax rate being result of the median voter optimization problem as in Persson e Tabellini (1994) and Jaimovich e Rebelo (2017). We analyze the case in which the median voter is a private market employee and face a constant and exogenous s[1.2.5]. This individual faces the following trade-off when chosing the tax rate. On the one hand, a higher τ implies a lower consumption level by reducing his disposable income. On the other hand, a higher τ means that the government would be able to pay teachers more, increasing the average teachers quality and the human capital of children, which also increases the utility level of all parents.

We can describe the optimization problem that this median voter faces as the choice of τ that maximizes his utility level subject to his budget constraint, equations (1.2.2) and (1.2.7), which in turn can be written as:

[1.2.5] Since we assume that the proportion of the individual's first period spent in school is a function of an exogenous θ and other constant institutional features, we will focus in the case of a constant s from now on.

$$
\begin{aligned}
\underset{\tau}{\text{maximize}} \quad & \log(c_t^i) + \phi \log(h_{t+1}^i) \\
\text{subject to} \quad & c_t^i \leq (1-\tau) w_t^M h_t^i, \\
& h_t^i = Z a^i (s_{t-1})^\eta (h_{t-1}^T)^v, \\
& h_t^T = \frac{Z\bar{B}}{2\theta}[1-(1+\theta^2)(1-\tau)](s_{t-1})^\eta (h_{t-1}^T)^v, \\
& s_t = s_{t-1} = s
\end{aligned}
$$

The solution to this problem is such that:

$$
\tau^\star = \frac{\phi v}{1+\phi v} + \frac{1}{1+\phi v} \cdot \frac{\theta^2}{1+\theta^2} \tag{1.2.9}
$$

From equation (1.2.9) one can see that $\tau^\star$ is a function of only three objects: (i) the parental altruism parameter, (ii) the education quality rate of return, and (iii) the proportion of teachers in the labor force. These parameters are all related to how important teachers are for the children's accumulation of human capital and how important children's human capital are for parents, which in turn justifies an optimal $\tau > 0$. Moreover, note that $\tau^\star$ is not indexed in i and thus all private market workers, regardless their innate ability level, would choose the same $\tau^\star$. Thus, in addition to the fact that the proportion of these workers in the labor force is always greater than 0.5 in the cross-country data, the median voter will indeed be represented by a private market worker and his choice given by equation (1.2.9).

It would be important to know, however, how the median voter choice of τ responds to changes in the parameters, specifically how it responds to an increase in the proportion of teachers in the labor force.

Proposition 1. *The median voter will choose a higher income tax rate ($\tau^\star$) the higher the proportion of teachers in the labor force (θ). However, $\tau^\star$ increases in decreasing rates as the return on education quality (v) increases.*

Proof. See Appendix B.

This result demonstrates that even a private market worker recognizes the importance of atracting skilled individuals to the teacher career. In spite of negatively affecting his disposable income, the median voter would vote for a higher income tax rate in response to an increase in θ in order to prevent teachers' salaries from falling too much. The magnitude of the increase in the optimal τ, however, depends on the education quality rate of return, which shows the importance of the average human capital of teachers mechanism even in the choice of optimal τ.

1.2.6 Competitive equilibrium

A competitive equilibrium in this OLG economy is such that the following condititions are satisfied: (i) each adult makes a career decision taking as given his own innate ability level and the innate ability threshold, (ii) the representative firm producing the final output maximizes profits taking wages as given, (iii) the government collects taxes and spends its revenue paying teachers, (iv) markets clear, and (v) τ is result of the median voter optimization. Thus, a formal definition of the competitive equilibrium can be presented as:

Definition 1. *A competitive equilibrium can be represented by sequences of the aggregate variables* $\{Y_t, H_t, h_t^T\}_{t=0}^{\infty}$*, innate ability threshold* $\{a_t^\star\}_{t=0}^{\infty}$*, wages* $\{w_t^M, w_t^T\}_{t=0}^{\infty}$ *and the optimal value for* $\tau^\star$ *such that equations (1.3.1) and (1.2.5) holds,* $w_t^M = A_t$*, human capital variables* $\{H_t, h_t^T\}_{t=0}^{\infty}$ *evolves according to equations (1.2.6)-(1.2.7), and* $\tau^\star$ *is given by equation (1.2.9).*

Additionally, the steady-state competitive equilibrium is such that all conditions of definition 1 are met and the average teachers' human capital is constant, i.e., $h_t^T = h_{t-1}^T = h^T$.

Proposition 2. *Consider two economies that are identical in all aspects except for their initial level of teacher's human capital. Given these, both economies will converge to the same unique steady-state equilibrium.*

Proof. See Appendix B.

1.3 Steady-state and comparative dynamics results

Our main interest in this section is to understand what are the policy implications of increasing average years of schooling by hiring more teachers. We would like to answer (i) what are the mechanisms through which government decisions affect individuals' occupational choice, (ii) if a policy to enhance human capital's quantity, such as the ones that we observed in the last decades in developing economies, affect the human capital's quality and long-run output equilibrium, and (iii) how responsive are these short-run and long-run effects to different education quantity and education quality rates of return.

1.3.1 Results on innate ability threshold

Note that all model's dynamics come from the evolution of teachers' average quality, which is deeply affected by the process of occupational choice and by the parameters which define who will be a teacher in the innate ability's distribution. In order to assess this question, use equation (1.2.3) and substitute equations (1.2.4), (1.2.5), and (1.2.6) to reach:

$$a^{\star} = \frac{(1+\theta^2)}{\theta} \cdot \frac{\bar{B}}{2} \cdot \tau \tag{1.3.1}$$

The occupational choice process takes place in the beginning of each period, before median voter chosing the optimal $\tau^{\star}$. For this reason, it is more informative to present $a^{\star}$ as a function of τ than replacing it by its $\tau^{\star}$ counterpart as in equation (1.2.9).

Proposition 3. *A higher proportion of teachers in the economy will make the teacher career less attractive to all individuals in the short and long-run, ceteris* *paribus. As a result, there will be a change in the pool of applicants for the teacher career, which now will come from a lower part of the innate ability distribution.*

Proof. See Appendix B.

Proposition 3 means that a greater θ implies a shift in the part of the innate ability distribution from which teachers come from. The same government budget will have to be divided by a larger number of teachers, which implies that the government will have to pay less for each teacher. By doing so, the individual that was indifferent between the career options will now strictly prefer to work in the private market. In fact, a greater θ means that potential good teachers will face a relative better outside option in the market. This leads to a selection problem in which the most qualified individuals choose not to apply for a job as a teacher. The greater the θ, the greater this selection issue.

1.3.2 Results on aggregate variables

We begin by taking the partial derivative of equation (1.2.7) and of its steady-state counterpart with respect to θ to understand how the proportion of teachers in the labor force affects teachers' quality in the short and long-run. This lead us to the following result:

Proposition 4. *A public policy that aims to increase the average years of schooling by hiring more teachers (i.e., increasing θ) has a negative impact on the average human capital of teachers (education quality) in both the short and long-run.*

Proof. See Appendix B.

This result is intuitive and is closely related to the negative impact that a higher θ has on $a^\star$. There is, however, another force working in the same direction: even if $a^\star$ had not changed, a higher θ would mean that the new job positions would be filled by relative less skilled candidates. This would imply a lower average quality of hired teachers, given the lower competition among candidates in the hiring process. Thus, this proposition shows the mechanism through which the education quantity-quality trade-off arises: an increase in average years of schooling by increasing the proportion of teachers in the labor force would be followed by a decrease in the quality of education as measured by the average human capital of teachers. Figure 1.3.1 ilustrates this proposition.

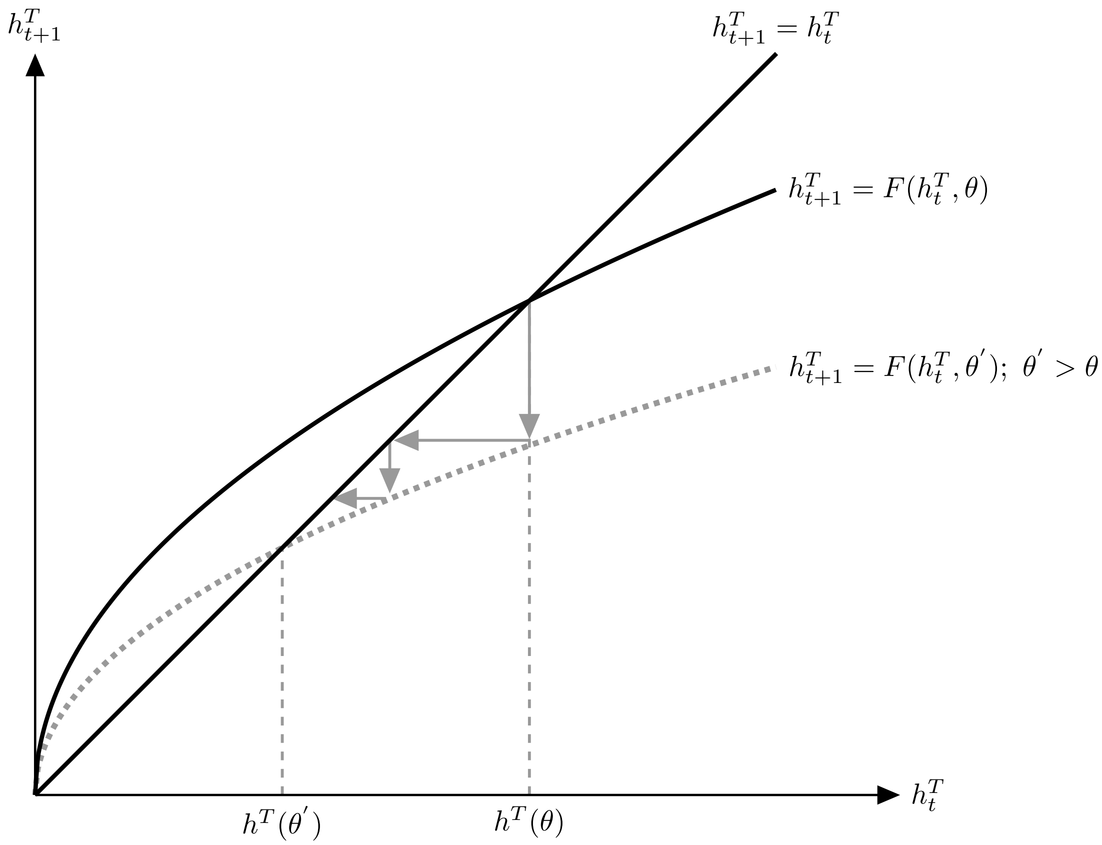


Figura 1.3.1 – Dynamics of the average human capital of teachers and the number of teachers.

The effect of this θ increase on H_t is unambiguously positive in the short-run[1.3.1], which is result of proposition 3 and intuitively of some of the most skilled individuals that were previously chosing the teacher career moving to the private market. On the other hand, the effect of a higher θ on the long-run value of H_t is result of two forces working on opposite directions: skilled individuals are more

[1.3.1] Take the derivative of equation (1.2.6) with respect to θ to see that $\partial H_t / \partial \theta > 0$.

attracted by the career on the private market, which makes H_t higher, but at the same time H_t is pulled down by the lower average teachers' quality.

To understand this long-run effect, lets take the partial derivative of the human capital stock of steady-state (H) with respect to θ. After multiplying both sides by the ratio between these variables we reach the elasticity of H with respect to θ, which can be written as (see Appendix A for the derivation):

$$\xi_{H\theta} = \frac{1}{1-\upsilon}(\eta\xi_{s\theta} - \upsilon) \tag{1.3.2}$$

where $\xi_{s\theta}$ is the elasticity of the proportion of the first period spent in school with respect to θ. Leaving aside the need for some teachers to work outside the classroom for the school to function properly, we can see the total number of teachers in the economy as a linear function of the average years of schooling such as:

$$teachers = \left(\frac{teachers}{pupils}\right) \cdot \left(\frac{pupils}{years\ of\ schooling}\right) \cdot years\ of\ schooling$$

where ($teachers/pupils$) is equivalent to the inverse of the well know pupil-teacher ratio and ($pupils/years$) is related to the size of the school age population. Since we are not interested in exploring the effects of changes in these ratios, we represent both as a constant term. Noting that s can be represented by the ratio between the average years of schooling and the total number of years of childhood, we can finally divide both sides of the equation above by this total number of years and by the population size of each generation, which is always one, to reach the linear form for s:

$$s = K \cdot \theta \tag{1.3.3}$$

where K represents all these education sector institutional features. From this point it is straightforward to rewrite equation (1.3.2) as a function of the education quantity and education quality rates of return.

$$\xi_{H\theta} = \frac{\eta - v}{1 - v} \tag{1.3.4}$$

It is clear from this equation that the effect of an increase in θ on the steady-state level of human capital stock can be both positive and negative: the sign and magnitude of $\xi_{H\theta}$ depend directly on the values we choose for the rates of return η and v. The difference between the short-run and long-run effects leads us to the main result of the paper:

Proposition 5. *A public policy that aims to increase the average years of schooling by hiring more teachers have a positive impact on the stock of human capital (and thus on GDP per capita) in the period right after the shock but its long-run impact depends on the education quantity and education quality rates of return.*

Proof. See Appendix B.

If we shut down the human capital quality channel by making $v \to 0$, equation (1.3.4) shows that an increase in education quantity (represented by a higher θ) would imply a higher aggregate human capital and, therefore, a higher GDP per capita in the steady-state. This is what one would expect as result of the evolution of average years of schooling in most countries in the last decades. As discussed in the introduction, figure 1.1.1 shows that this is not the case and proposition 5 helps us to understand why. When the human capital quality channel is working ($v > 0$), a higher θ would not necessarily imply a higher aggregate human capital in the steady-state and may even imply a lower one. This is a direct result of the education quantity-quality tradeoff in which a higher number of teachers leads to an increase in the average years of schooling but also leads to a lower average quality of teachers. So, the GDP per capita behavior in figure 1.1.1 could be result of the increase in education quantity being accompanied at the same time by a decrease in education quality, making the contribution to economic growth of an increase in the average years of schooling close to zero.

Duflo (2001) assess an unusual policy of school construction and teacher hiring in Indonesia in which investment in education quantity led to an increase in individuals human capital, in spite of the possible negative effects on education quality. Still, the effect of the policy on individuals' wages is only positive and statistically significant in regions that had a below median preprogram education, which may be viewed as another evidence in favor of our proposition 5 if we assume smaller values for η in equation (1.3.2) as average years of schooling increases (HALL; JONES, 1999; CASELLI, 2005).

1.4 Calibration

We calibrate the model using data from the World Bank's World Development Indicators for a set of nine Latin America countries[1.4.1] in the years of 1970, 1990 and 2010. The resulting values for our baseline economy are summarized in table 1.4.1.

Tabela 1.4.1 – Calibrated parameters for our baseline economy

Description	Parameter	Value	Source
Education quantity rate of return	η	10%	Psacharopoulos (1994)
Education quality rate of return	υ	10%	Schoellman (2012)
Average years of schooling function parameter	K	4.50	Match Psacharopoulos, Valenzuela e Arends (1996)
Percentage of teachers in the labor force, 1970	θ_{1970}	2.95%	Match average years of schooling in 1970
Percentage of teachers in the labor force, 2010	θ_{2010}	6.06%	Match average years of schooling in 2010
Parental altruism	ϕ	0.24	Match $\tau^{\star}$
Upper bound of innate ability	$\bar{B}$	1.00	Normalization
Individual human capital function parameter	Z	4.54	Match Psacharopoulos, Valenzuela e Arends (1996)

We begin by calibrating the education quantity and education quality rates of return (υ and η, respectively) according to the microeconometric literature evidence. There is a well established consensus on the mincerian return to one more

[1.4.1] Argentina, Bolivia, Brazil, Colombia, Costa Rica, Ecuador, Honduras, Peru and Venezuela. According to the World Bank, these countries accounted for about 64% and 68% of Latin America's population and GDP, respectively, in 2010.

year of schooling being around $6\% - 14\%$ (e.g., Psacharopoulos (1994)), and thus we set $\eta = 0.10$. This consensus on the rate of return of schooling quantity, however, is not followed by a similar one on the value of the schooling quality's rate of return, given the inherent empirical issues on estimating returns to teacher quality while controlling for teacher quantity. However, there are a plenty of evidence that $\upsilon > 0$ (e.g., Cheety, Friedman e Rockoff (2014)), and evidence that differences in education quality are at least as important as differences in education quantity in accounting for differences in cross-country output levels (SCHOELLMAN, 2012). Following the latter, we set $\upsilon = 0.10$ in our baseline economy but also show the results for different υ values.

Our choice of $K = 4.5$ imply that, as in equation (1.3.3), the needed θ to generate the observed mean of average years of scholling in these nine Latin America countries in 1990 (BARRO; LEE, 2013) is the same as the 1989 observed mean of the percentage of teachers in the labor force in these countries as in Psacharopoulos, Valenzuela e Arends (1996). We then use this value for K to reach the θ values for 1970 and 2010 given the observed mean of average years of schooling for both years, which implies $\theta = 2.95\%$ for a mean of 3.98 in average years of schooling in 1970, and $\theta = 6.06\%$ for a mean of 8.18 in average years of schooling in 2010. The choice of a constant K means that we are keeping constant the institutional features of the education sector (e.g., the teacher-pupil ratio and the number of pupils per year of schooling), which enable us to assess the effects of increasing education quantity in isolation.

As shown in equation (1.2.9), the median voter choice on τ is directly related to the parental altruism parameter ϕ. We, therefore, choose $\phi = 0.24$ as to guarantee that $\tau^\star$ equals the mean of taxes on income, profits and capital gains as % of GDP in 2010[1.4.2]. This value is similar to the parental altruism parameter in the Lee e Seshadri (2019) model calibration. In their paper, however, this parameter represents how much the parental utility is responsive to children's consumption and not how much it is responsive to children's human capital, as in our model.

Finally, we normalize $\bar{B}$ to one as in Galor e Moav (2000), and choose the

[1.4.2] We have data on this income tax only for Argentina, Brazil, Colombia, Costa Rica, Honduras, and Peru.

productivity term Z as to the predicted steady-state ratio between teachers' wage and market employees' wage in 1990 equals the mean of earnings ratio between teachers and the comparison group of Psacharopoulos, Valenzuela e Arends (1996).

1.5 Numeric simulation

In this section, we feature the results for the initial steady-state of the baseline economy and simulate the short-run and long-run effects of an increase in average years of schooling. We begin by doing the former.

With both rates of return being equal to 10% and $\phi = 0.24$, the 2.95% of teachers in the labor force generates an innate ability threshold of 0.413. The 40% less skilled individuals in the economy were attracted by the teacher career, but only those with their innate ability above 0.384 and below 0.413 were actually hired. This low attractiveness of the teacher career is in accordance with the empirical evidence for the US, in which low ability individuals are more likely to enter teaching and high ability individuals who do enter teaching are more likely to leave the career at some point (HANUSHEK; PACE, 1995; PODGURSKY; MONROE; WATSON, 2004).

This $\theta = 2.95\%$ implies that average years of schooling are equal to 3.98 in our baseline economy, only 37% of the United States' average years os schooling in 1970. Moreover, in this baseline economy, the optimal income tax is equal to 2.435%, which is mostly the result of the median voter recognizing the importance of government offering a good salary for teachers in order to guarantee the quality of his children education.

Now we carry out two sets of simulations. First, we evaluate how the government decision to increase average years of schooling affects the occupational choice of individuals, the short-run/long-run stock of human capital, and the short-run/long-run level of education quality in the baseline economy. Second, we vary each rate of return (η and v) at a time to examine how the effects of the increase in education quantity are responsive to these parameters.

1.5.1 Policy implications for the baseline economy

Figure 1.5.1 shows the aggregated effects of an increase in the proportion of teachers in the labor force from 2.95% to 6.06%. This increase in θ is result of the observed increase in the average years of schooling of our calibrated economy from 3.98 years in 1970 to 8.18 years in 2010.

Just as predicted by proposition 3, a higher θ led to a lower innate ability threshold such that only individuals with $a^i \in [0.164; 0.224]$ actually became teachers. As a result of the lower attractiveness of the teacher career, the most skilled potential teachers moved to jobs in the private market, which led to an increase in H_t in the short-run. However, the lower attractiveness of the teacher career also made the average teachers' quality to fall in the short-run and to fall to less than half of the initial value in the new steady-state.

As a result of the impact that a lower teachers' quality have on human capital accumulation of future generations, the effect of a higher θ on the long-run value of H_t is zero. In spite of the positive effect in the short-run as a result of the most skilled individuals moving to jobs in the private market, the lower teachers' quality pulled down the positive effect that this movement could also have in the long-run. Thus, equation (1.2.4) implies that, aside some exogenous shock on TFP, the distribution of GDP *per capita* should remain the same even after a great increase in education quantity such as the one we observed in the $1970 - 2010$ period in Latin America.

1.5.2 Policy implications for different rates of return

This zero impact on the steady-state value of H_t is due to the trade-off between education quantity and education quality, and to the trade-off magnitude that is given by the rates of return as in equation (1.3.4). In order to assess how much the impact of a higher θ responds to changes in the rates of return, we replicate the exercise in the previous section with different values for υ and η.

Figure 1.5.2 shows that with $\eta = 0.10$, the effect of a higher θ on H_t goes

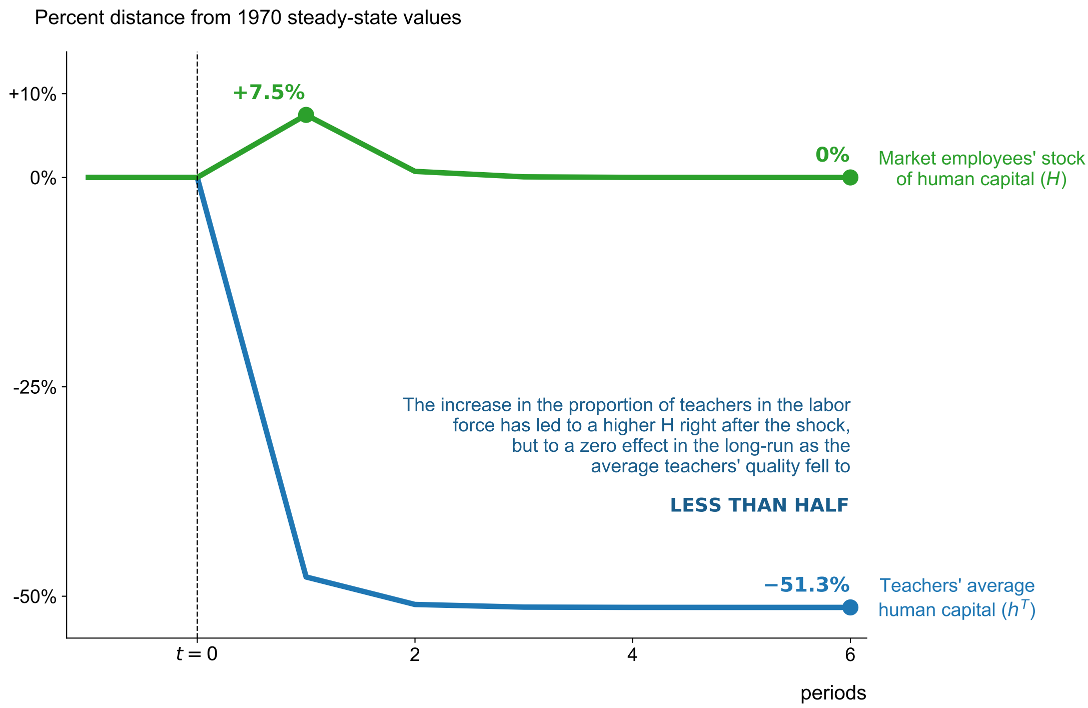


Figura 1.5.1 – Evolution of teachers's human capital and private market workers' human capital as result of an increasing θ policy.

from $+7.5\%$ with $v = 0$ to -8.6% with $v = 0.20$. Intuitively, a higher v makes education quality more relevant to the accumulation of children's human capital such that the negative impact of a decrease in the quality of teachers could be enough to make a larger θ implies a lower H_t in the new steady-state. This shows how important it is to endogenously include education quality in the model in order to understand the cross-country data in figure 1.1.1. Besides that, there is not a consensus on the magnitude of the rate of return of education quality and so it is also important to understand how the new steady-state of H_t varies with the magnitude of v.

Unlike the education quality's rate of return, there is a well established consensus on the magnitude of the education quantity's rate of return. However, there are also evidence that this rate becomes lower as the country's average years of schooling increases (PSACHAROPOULOS, 1994). Then, following Hall e Jones (1999) and Caselli (2005), we evaluate how the effect of a higher θ responds to a

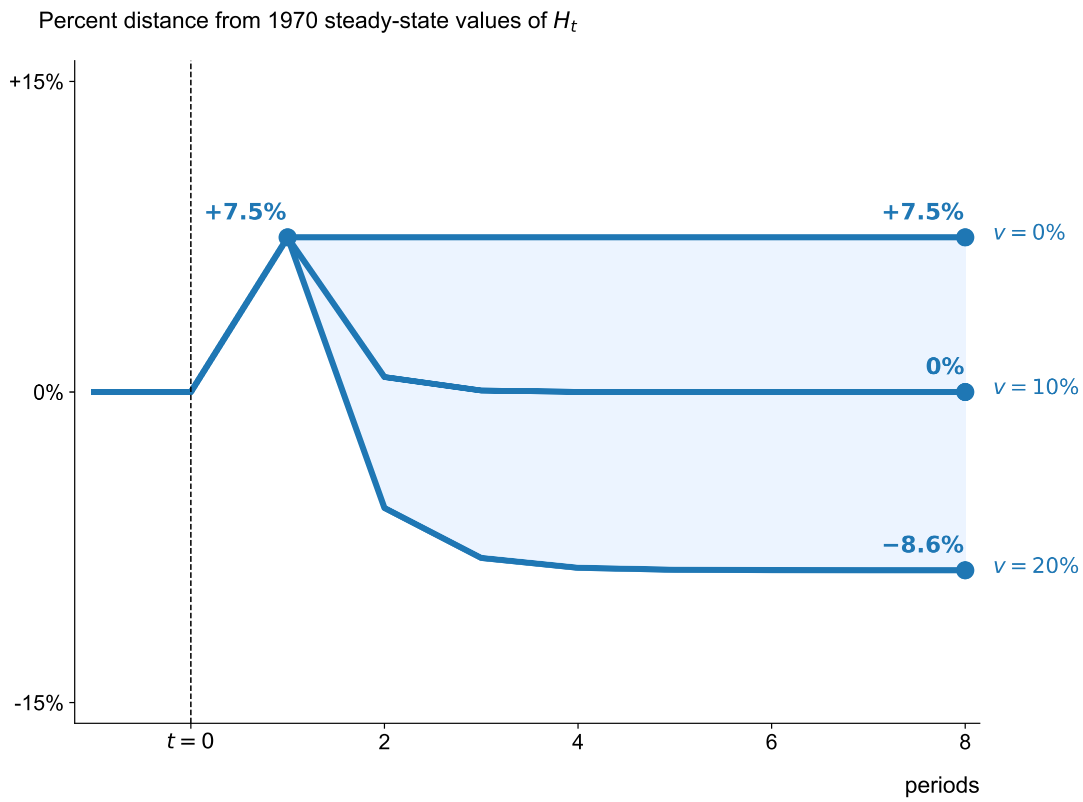


Figura 1.5.2 – Sensitivity of the private market workers' human capital to the rate of return of education quality.

return-to-schooling of 13.4%, 10.1% and 6.8%, which are, respectively, the average values for the sub-Saharan Africa, the world as a whole, and the OECD countries as reported in Psacharopoulos (1994).

Figure 1.5.3 shows the result of this exercise when we keep v constant at 10%. Regardless of the value of η, the impact of a higher θ is larger in the short-run, just as predicted by proposition 5. The impact in the long-run goes from +2.8% with $\eta = 13.4\%$ to -2.5% with $\eta = 6.8\%$. Thus, as one could have concluded from equation (1.3.4), if η decreases with an increasing average years of schooling, the effectiveness of a government policy to improve the quantity of education becomes lower as the country's population becomes more educated. If we keep v at 10% and assume the same piecewise linear function for η as in Hall e Jones (1999) and Caselli (2005), a country with an average years of schooling of more than 8 should expect a decrease in the stock of human capital as result of a policy to further increase the education quantity of the population, and not the other way around.

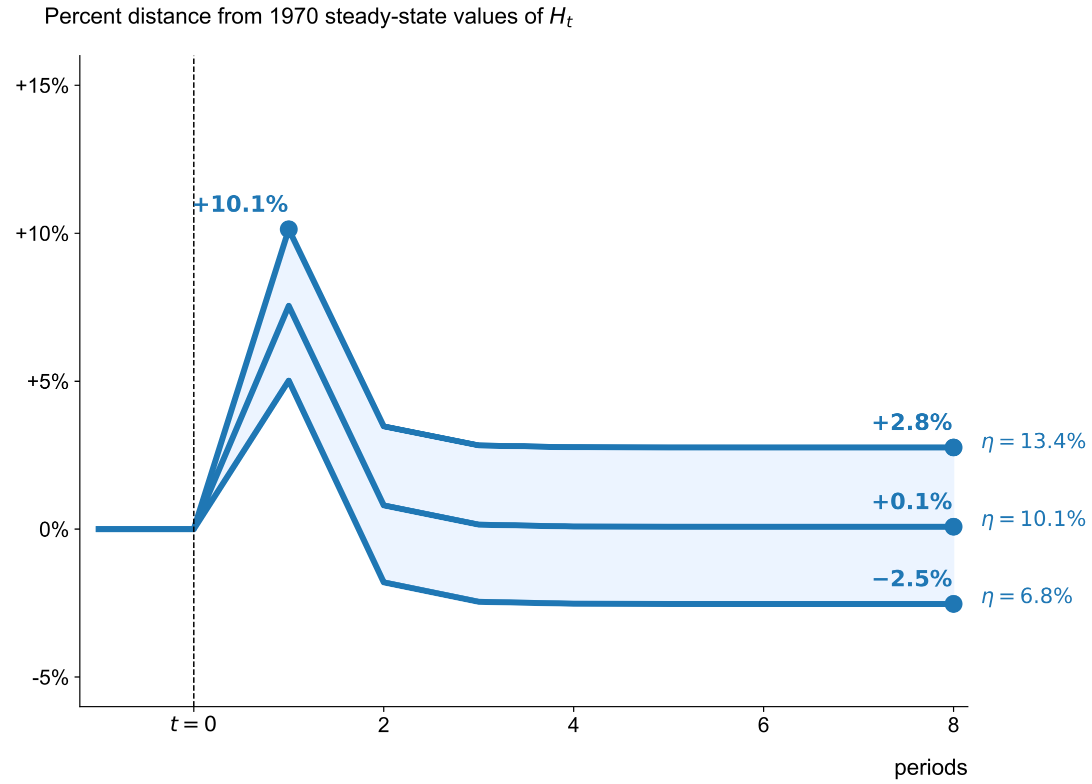


Figura 1.5.3 – Sensitivity of the evolution of private market workers' human capital to the rate of return of education quantity.

1.6 Conclusion

Over the last half century we have witnessed a rapid expansion in the average years of schooling of most countries, which made the world a less unequal place in education. We have not witnessed, however, a similar convergence pattern in the world's distribution of GDP *per capita*. In this paper we have developed an overlapping generations model in which the existence of a trade-off between the quantity of education and the quality embodied in this education could help us to understand what is behind the pattern that we observe in the cross-country data.

In this model, individuals are heteregenous in their innate ability and the quality of the human capital accumulated in public schools is measured by the average quality of those individuals who choose to be teachers in the adulthood. The income tax rate is result of the median voter optimization, which leaves only the number of teachers under government control. Given the balanced government budget and keeping institutional features of the education sector constant, the model predicts that a policy to increase the population's average years of schooling is followed by a decrease in the quality of teachers in both short and long-run. As a result, the long-run effect of this same policy on the country's stock of human capital is undetermined in spite of an unambigous short-run positive effect. Moreover, we show that the sign and magnitude of this long-run effect is directly related to the education quantity and education quality rates of return of the human capital accumulation function.

In order to understand what the model would predict in response to the observed increase in average years of schooling between 1970 and 2010 in Latin America we simulate the model and evaluates the short and long-run effects on the quality of education and in the country's stock of human capital. We show that in spite off a short-run increase of 7.5% in the stock of human capital, the long-run decrease of -51.3% in the average human capital of teachers is enough to zero the long-run impact on the stock of human capital. If we assume, however, that the return-to-schooling decreases as the average years of schooling increases, the long-run effectiveness of this education quantity policy also decreases. These results suggest, as pointed empirically by Delgado, Henderson e Parmeter (2014), that taking education quality into account could be an important step in order to reconcile classical growth model predictions and the cross-country data.

Appendix A: reaching final equations

Solving the median voter's problem. Given that adults utility is logarithmic in both consumption and children's human capital, the first-order condition with respect to

τ is necessary and suficient to reach the $\tau^\star$ that maximizes u_t^i. We can rewrite the optimization problem by substituting all constraints in the utility function such as

$$\underset{\tau}{\text{maximize}} \quad \log\left[(1-\tau)w_t^M h_t\right] + \phi\log\left[Za^i s^\eta (h_t^T)^v\right]$$

which is also equivalent to

$$\underset{\tau}{\text{maximize}} \quad \log\left(1-\tau\right) + \phi v \log\left(h_t^T\right)$$

given that the average teachers quality h_t^T is a function of τ and given that the private market worker take w_t^M, h_t, Z, a^i and s^η as given. Thus, the first-order condition is given by

$$\frac{-1}{1-\tau^\star} + \frac{\phi v(1+\theta^2)}{1-(1+\theta^2)(1-\tau^\star)} = 0$$

which finally implies

$$\tau^\star = \frac{\phi v}{1+\phi v} + \frac{1}{1+\phi v}\cdot\frac{\theta^2}{1+\theta^2}$$

How to reach equations (1.2.6) and (1.2.7). Equations (1.2.6) and (1.2.7) are the result of integrating the individual human capital production function for each group (market employees and teachers), given the innate ability threshold and the government's budget constraint. Thus,

$$
\begin{aligned}
H_t &= \int_0^{a^\star - \bar{B}\theta} h_t^i(a^i)dF(a^i) + \int_{a^\star}^{\bar{B}} h_t^i(a^i)dF(a^i) \\
&= \int_0^{a^\star - \bar{B}\theta} Za^i(s_{t-1})^\eta (h_{t-1}^T)^v f(a^i)da^i + \int_{a^\star}^{\bar{B}} Za^i(s_{t-1})^\eta (h_{t-1}^T)^v f(a^i)da^i \\
&= Z(h_{t-1}^T)^v(s_{t-1})^\eta \left(\int_0^{a^\star - \bar{B}\theta} a^i \frac{1}{\bar{B}} da^i + \int_{a^\star}^{\bar{B}} a^i \frac{1}{\bar{B}} da^i \right) \\
&= \frac{Z(s_{t-1})^\eta (h_{t-1}^T)^v}{\bar{B}} \left(\left[\frac{(a^i)^2}{2} \right]_0^{a^\star - \bar{B}\theta} + \left[\frac{(a^i)^2}{2} \right]_{a^\star}^{\bar{B}} \right) \\
&= \frac{Z(s_{t-1})^\eta (h_{t-1}^T)^v}{2\bar{B}} \left((a^\star)^2 - 2a^\star \bar{B}\theta + \bar{B}^2\theta^2 + \bar{B}^2 - (a^\star)^2 \right) \\
&= \frac{Z\bar{B}(1+\theta^2)(s_{t-1})^\eta}{2}(h_{t-1}^T)^v - Z(s_{t-1})^\eta (h_{t-1}^T)^v \theta a^\star \\
&= \frac{Z\bar{B}(1+\theta^2)(s_{t-1})^\eta}{2}(h_{t-1}^T)^v - Z(s_{t-1})^\eta (h_{t-1}^T)^v \theta \frac{w_t^T}{w_t^M} \frac{1}{Z(s_{t-1})^\eta (h_{t-1}^T)^v} \\
&= \frac{Z\bar{B}(1+\theta^2)(s_{t-1})^\eta}{2}(h_{t-1}^T)^v - \theta \frac{1}{w_t^M} \frac{1}{\theta} \frac{\tau}{1-\tau} w_t^M H_t \\
&= \frac{Z\bar{B}(1+\theta^2)(s_{t-1})^\eta}{2}(h_{t-1}^T)^v - \frac{\tau}{(1-\tau)} H_t
\end{aligned}
$$

and finally

$$
H_t = \frac{Z\bar{B}}{2} \Big[(1+\theta^2)(1-\tau) \Big] (s_{t-1})^\eta (h_{t-1}^T)^v
$$

Equation (1.2.7) just come from equation (1.2.6) and the fact that $H_t + \theta.h_t^T = \int_0^{\bar{B}} h_t^i(a^i).dF(a^i) = \frac{Z.\bar{B}}{2}(s_{t-1})^\eta.(h_{t-1}^T)^v$. Thus,

$$
\begin{aligned}
h_t^T &= \frac{1}{\theta} \left[\frac{Z.\bar{B}}{2}(s_{t-1})^\eta.(h_{t-1}^T)^v - H_t \right] \\
&= \frac{1}{\theta} \left[\frac{Z.\bar{B}}{2}(s_{t-1})^\eta.(h_{t-1}^T)^v - \frac{Z\bar{B}}{2} \Big[(1+\theta^2)(1-\tau) \Big] (s_{t-1})^\eta (h_{t-1}^T)^v \right]
\end{aligned}
$$

and finally

$$h_t^T = \frac{Z\bar{B}}{2\theta}[1-(1+\theta^2)(1-\tau)](s_{t-1})^{\eta}(h_{t-1}^T)^{v}$$

How to reach equation (1.3.2). We begin by writing the average teachers quality of steady-state

$$h^T = \left[\frac{Z\bar{B}}{2\theta}[1-(1+\theta^2)(1-\tau^\star)]s^{\eta}\right]^{\frac{1}{1-v}}$$

Note that the aggregated human capital of steady-state H is a function of h^T and given by

$$H = \frac{Z\bar{B}}{2}[(1+\theta^2)(1-\tau^\star)]s^{\eta}(h^T)^{v}$$

which can be simplified in order to isolate the exogenous parameters and the part that is dependent on θ:

$$\begin{aligned}
H &= \left[\frac{Z\bar{B}}{2}\right]^{\frac{1}{1-v}} \cdot \frac{s^{\frac{\eta}{1-v}}}{\theta^{\frac{v}{1-v}}} \cdot \left[(1+\theta^2)(1-\tau^\star)\right] \cdot \left[1-(1+\theta^2)(1-\tau^\star)\right]^{\frac{v}{1-v}} \\
&= \left[\frac{Z\bar{B}}{2}\right]^{\frac{1}{1-v}} \cdot \frac{s^{\frac{\eta}{1-v}}}{\theta^{\frac{v}{1-v}}} \cdot \left[(1+\theta^2)\left(1-\frac{\phi v(1+\theta^2)+\theta^2}{(1+\phi v)(1+\theta^2)}\right)\right] \cdot \left[1-(1+\theta^2)\left(1-\frac{\phi v(1+\theta^2)+\theta^2}{(1+\phi v)(1+\theta^2)}\right)\right]^{\frac{v}{1-v}} \\
&= \left[\frac{Z\bar{B}}{2}\right]^{\frac{1}{1-v}} \cdot \frac{s^{\frac{\eta}{1-v}}}{\theta^{\frac{v}{1-v}}} \cdot \frac{1}{(1+\phi v)} \cdot \frac{(\phi v)^{\frac{v}{1-v}}}{(1+\phi v)^{\frac{v}{1-v}}} \\
&= \left[\frac{Z\bar{B}}{2}\right]^{\frac{1}{1-v}} \cdot \frac{s^{\frac{\eta}{1-v}}}{\theta^{\frac{v}{1-v}}} \cdot \frac{(\phi v)^{\frac{v}{1-v}}}{(1+\phi v)^{\frac{1}{1-v}}} \\
&= \left[\frac{Z\bar{B}}{2}\right]^{\frac{1}{1-v}} \cdot \frac{(\phi v)^{\frac{v}{1-v}}}{(1+\phi v)^{\frac{1}{1-v}}} \cdot s^{\frac{\eta}{1-v}} \cdot \theta^{\frac{-v}{1-v}}
\end{aligned}$$

Thus, since $s = f(\theta)$, we can differentiate H with respect to θ such that

$$\begin{aligned}
\frac{\partial H}{\partial \theta} &= \left[\frac{Z\bar{B}}{2}\right]^{\frac{1}{1-v}} \cdot \frac{(\phi v)^{\frac{v}{1-v}}}{(1+\phi v)^{\frac{1}{1-v}}} \cdot \left[\frac{\eta}{1-v} s^{\frac{\eta}{1-v}-1} \frac{\partial s}{\partial \theta} \theta^{\frac{-v}{1-v}} - \frac{v}{1-v} \theta^{\frac{-v}{1-v}-1} s^{\frac{\eta}{1-v}}\right] \\
&= \left[\frac{Z\bar{B}}{2}\right]^{\frac{1}{1-v}} \cdot \frac{(\phi v)^{\frac{v}{1-v}}}{(1+\phi v)^{\frac{1}{1-v}}} \cdot \left[s^{\frac{\eta}{1-v}} \theta^{\frac{-v}{1-v}} \frac{\eta}{1-v} \frac{1}{s} \frac{\partial s}{\partial \theta} - s^{\frac{\eta}{1-v}} \theta^{\frac{-v}{1-v}} \frac{v}{1-v} \frac{1}{\theta}\right] \\
&= \left[\frac{Z\bar{B}}{2}\right]^{\frac{1}{1-v}} \cdot \frac{(\phi v)^{\frac{v}{1-v}}}{(1+\phi v)^{\frac{1}{1-v}}} \cdot s^{\frac{\eta}{1-v}} \cdot \theta^{\frac{-v}{1-v}} \cdot \left[\frac{\eta}{1-v} \frac{1}{s} \frac{\partial s}{\partial \theta} - \frac{v}{1-v} \frac{1}{\theta}\right] \\
&= H \cdot \frac{1}{1-v} \cdot \left[\frac{\eta}{s} \frac{\partial s}{\partial \theta} - \frac{v}{\theta}\right]
\end{aligned}$$

And finally, by multiplying both sides by θ/H,

$$\frac{\partial H}{\partial \theta} \cdot \frac{\theta}{H} = \frac{1}{1-v} \cdot \left[\eta \frac{\partial s}{\partial \theta} \frac{\theta}{s} - v\right]$$

which can be restated as

$$\xi_{H\theta} = \frac{1}{1-v}(\eta \xi_{s\theta} - v)$$

Appendix B: proof of propositions

Proof of Proposition 1. This proof follows directly from the partial derivatives of $\tau^\star$ with respect to θ and v:

$$\frac{\partial \tau^\star}{\partial \theta} = \frac{2\theta}{(1+\phi v)(1+\theta^2)^2} > 0$$

$$\frac{\partial^2 \tau^\star}{\partial \theta \partial v} = -\frac{2\theta}{(1+\phi v)^2(1+\theta^2)^2} \cdot \phi < 0$$

Note that □

Proof of Proposition 2. Since the whole model's dynamic is driven by h_t^T evolution, its behavior is a necessary and sufficient condition to define behavior of all real variables. Growth of h_t^T is given by:

$$g(h_{t-1}^T) \equiv \frac{h_t^T}{h_{t-1}^T} = \frac{Z\bar{B}}{2\theta}[1-(1+\theta^2)(1-\tau)]s^{\eta}(h_{t-1}^T)^{v-1}$$

By definition, h^T is a steady state if and only if $g(h^T) = 1$, i.e. $h_t^T = h_{t-1}^T = h^T$. Since $\lim_{h^T \to 0} g(h^T) = \infty$, $\lim_{h^T \to \infty} g(h^T) = 0$ and $g(.)$ is a continuous function, by the Intermediate Value Theorem there is a h^T such that $0 < g(h^T) = 1 < \infty$. Moreover, note that $g(.)$ is everywhere decreasing, i.e. $\partial g(h^T)/\partial h^T < 0 \; \forall h^T \in (0, \infty)$. Therefore, there can only exist a unique h^T such that $g(h^T) = 1$. Every country with the same structural and policy parameters will converge to this h^T, regardless the initial conditions.

□

Proof of Proposition 3. This proof follows directly from the partial derivative of $a^\star$ with respect to θ when (i) τ is taken as exogenous (short-run effect), and (ii) the median voter $\tau^\star$ is in place (long-run effect). The first case is straightforward,

$$\frac{\partial a^\star}{\partial \theta} = -\frac{(1-\theta^2)}{\theta^2} \cdot \frac{\bar{B}}{2} \cdot \tau < 0$$

but the second one need a little more algebra. Lets take equations (1.2.9) and (1.3.1) together and then take the partial derivative of the resulting equation with respect to θ.

$$\begin{aligned} \frac{\partial a^\star}{\partial \theta} &= \frac{\phi v}{1+\phi v} \cdot \left(-\frac{(1-\theta^2)}{\theta^2} \cdot \frac{\bar{B}}{2} \right) + \frac{1}{1+\phi v} \cdot \frac{\bar{B}}{2} \\ \frac{\partial a^\star}{\partial \theta} &= \frac{1}{1+\phi v} \cdot \frac{\bar{B}}{2} \cdot \left(-\frac{\phi v(1-\theta^2)}{\theta^2} + 1 \right) \end{aligned}$$

Notice that $\partial a^\star / \partial \theta < 0$ in the equation above if and only if $\phi v(1-\theta^2) > \theta^2$. This is indeed the case for the values we found in the cross-country data and the microeconometric evidence.

□

Proof of Proposition 4. We begin by taking the partial derivative of equation (1.2.7) with respect to theta to show the short-run effect:

$$\frac{\partial h_t^T}{\partial \theta} = \left[\frac{Z\bar{B}}{2} \cdot (s_{t-1})^{\eta} \cdot \left(h_{t-1}^T\right)^{v}\right] \cdot \left[\frac{-2\theta^2(1-\tau) - 1 + (1+\theta^2)(1-\tau)}{\theta^2}\right]$$

$$\frac{\partial h_t^T}{\partial \theta} = \left[\frac{Z\bar{B}}{2} \cdot (s_{t-1})^{\eta} \cdot \left(h_{t-1}^T\right)^{v}\right] \cdot \left[\frac{-\tau - \theta^2(1-\tau)}{\theta^2}\right] < 0$$

To see the long-run effect of the same increase in θ we first take the log in both sides of the steady-state counterpart of equation (1.2.7)

$$h^T = \left[\frac{Z\bar{B}}{2\theta}[1-(1+\theta^2)(1-\tau^{\star})]s^{\eta}\right]^{\frac{1}{1-v}}$$

$$\ln(h^T) = \frac{1}{1-v} \cdot \left[\ln\left(\frac{ZB}{2}\right) + \eta \ln(s) - \ln(\theta) + \ln(1-(1+\theta^2)(1-\tau))\right]$$

and, then, take the partial derivative of $\ln(h^T)$ with respect to $\ln(\theta)$, which is the same as calculating the elasticity of h^T with respect to θ.

$$\frac{\partial \ln(h^T)}{\partial \ln(\theta)} = \frac{1}{1-v} \cdot \left[\eta \frac{\partial \ln(s)}{\partial \ln(\theta)} - 1 + \frac{\partial \ln(1-(1+\theta^2)(1-\tau))}{\partial \ln(\theta)}\right]$$

$$\xi_{h^T\theta} = \frac{1}{1-v} \cdot \left[\eta \xi_{s\theta} - 1 - \frac{2\theta^2(1-\tau)}{1-(1+\theta^2)(1-\tau)}\right]$$

which can be finally rewriten as the equation below by incorporating the function that relates s and θ as in equation (1.3.3).

$$\xi_{h^T\theta} = \frac{1}{1-v} \cdot \left[\eta - 1 - \frac{2\theta^2(1-\tau)}{1-(1+\theta^2)(1-\tau)}\right]$$

Thus, since $\eta \in (0,1)$, it is clear that $\xi_{h^T\theta} < 0$, which means that a policy to increase average years os schoolling by hiring more teachers is followed by a decrease in the average human capital of teachers in both the short and long-run.

□

Proof of Proposition 5. We begin by taking the partial derivative of equation (1.2.6) with respect to theta to show the short-run effect:

$$\frac{\partial H_t}{\partial \theta} = \left[\frac{Z\bar{B}}{2} \cdot (s_{t-1})^{\eta} \cdot \left(h^T_{t-1}\right)^{v}\right] \cdot 2\theta(1-\tau) > 0$$

The long-run effect of the same increase in θ, however, follows directly from equation (1.3.4). The elasticity of the steady-state value of H_t with respect to θ can be both positive and negative depending on the values we assume for the education quantity and education quality rates of return (η and v respectively). □

2 NEGATIVE CAMPAIGNING: THEORY AND EVIDENCE

Com Marcos Nakaguma

2.1 Introduction

Political campaigns are at the core of the democratic process. Campaigns matter because they have a direct impact on voters' turnout decisions, choice of candidates and general political attitudes by providing relevant information about politicians, their past performances and policy proposals (JACOBSON, 2015). The political economy literature widely recognizes the role of information on policy outcomes and accountability and it is in this sense that the campaign strategies adopted by candidates are crucial: they affect both the type and quality of information revealed to voters.

While campaigns may facilitate access to information and stimulate political participation, they are often conducted with a clear intent to spread distorted and misleading information, focusing on personal defamatory attack[2.1.1]. "Dirty campaigns"and "mudslinging"are by no means a new phenomena in politics[2.1.2], but their potential negative impacts have increased significantly in recent years with the advent of social media and the ability of campaigns to target specific constituencies (ALLCOTT; GENTZKOW, 2017). Indeed, there is an increasing concern that fake news and slanderous attacks may pose serious threats to democracy (ANSOLABEHERE et al., 1994), so that it is important to understand the main factors that might influence candidates' decisions to "go dirty".

[2.1.1] According to CNN, "in the last days of 2016 US presidential campaign only 3% [of television ads] focused on positive messages about Clinton, and 5% were built around positive messages about Trump"(<http://edition.cnn.com/2016/11/08/politics/negative-ads-hillary-clinton-donald-trump/>)

[2.1.2] For instance, in the 1796 US presidential election, Alexander Hamilton accused Thomas Jefferson of sleeping with one of his slaves and in the 1828 US presidential election, John Quincy Adams' supporters called Andrew Jackson's mother a prostitute, his wife a whore and Jackson himself a murderer and a cannibal.

Most of the research on negative advertising has been carried out in the context of the US electoral system and the predominant view in the literature is that the two-party plurality system adopted in the United States generates more incentive for negative advertising (GHANDI; IORIO; URBAN, 2016). However, while there are a few studies comparing levels of campaign negativity across countries (WALTER, 2014; WALTER; BRUG; PRAAG, 2014), very little is known about the causal impact of the race characteristics on the decision to go negative. In fact, even from a theoretical point of view, it is not clear what pattern of behavior to expect as one moves from a single-ballot to a dual-ballot system, for example. Therefore, both theory and empirical evidence on this topic are much needed in order to understand mechanisms and establish basic stylized facts, highlighting which elements are more relevant to the problem.

This paper analyzes both theoretical and empirical aspects of candidates's campaign strategies by investigating the incentives candidates have as we change the race background. We begin by developing a stylized model of decision on dirty campaigning and evaluating its equilibrium properties and predictions. We define the winning probability as a function of candidates innitial support and the distribution of attacks, and then fully characterize the model equilibrium for the case of two and three candidates in both single-ballot and dual-ballot systems. The equilibrium analysis establishes how each candidate behave in a more competitive environment, against all opponents and as opponents become stronger, and in different electoral systems (single-ballot and dual-ballot).

We show that, in a two-candidate competition of a single-ballot race, both the front-runner and the runner-up are more likely to attack when the opponent's innitial support increases. This suggest that the return on attacking another candidate is higher the more potential votes the candidate has to lose. Next, the model predicts that, under single-ballot system, candidate 1 and 2 are more likely to attack each other and candidate 3 more likely to attack candidate 1. Moreover, candidates 1 and 2 are more aggressive towards each other when there isn't any other candidate in the dispute, i.e., when the electoral competition is higher. Finally, we show that candidates 2 and 3 are more likely to attack each other when we move from a single-ballot to a dual-ballot system. Intuitively, this could be result of an attempt to guarantee a place in the second round of election, which is present only on dual-ballot systems.

The model predictions are tested using both a fixed-effect regression methodology and a regression discontinuity design on a dataset of electoral races for mayor in Brazil. As a measure of dirty negative campaigning, we use detailed information collected from Brazil's Regional Electoral Courts databases on the "right of reply"lawsuits ("acoes direito de resposta") filed during the municipal elections of 2012 and 2016. The Brazilian electoral legislation protects candidates against slanderous, defamatory or untrue statements, granting the aggrieved candidate the right to demand space for a reply in the same media vehicle. For instance, a candidate attacked on television during the electoral campaign period has the right to demand the use of the offender's television time to set the record straight. A right of reply lawsuit has to be judged by an electoral court within 72 hours of being filed. We use these lawsuits to identify, for each municipality, the candidates between which there was an attack that motivated a legal action. We then merge this information with data on electoral outcomes and the characteristics of municipalities and candidates to create an unique dataset of candidates pairs per municipality.

We begin our empirical analysis by estimating linear probability models at the candidate-pair level where we use as dependent variable a dummy indicating whether a candidate i attacked another candidate j in a given municipality and year. Controlling for a number of candidate-pair characteristics, municipality and election-year fixed-effects, and focusing only on first-round elections, we obtain a number of robust stylized facts. First, we show that dirty campaigning is significantly more likely between candidate i and his opponent j as the opponent's innitial support increases. Second, we present evidence that under single-ballot system candidate 1 and 2 are more likely to attack each other and candidate 3 more likely to attack candidate 1. Third, we show that the probability of an attack involving the two best placed candidates increases significantly when one moves from a race with more than two candidates to another race with only two candidates disputing the seat ("duopoly").

Next, we exploit the fact that Brazil municipalities with less than $200,000$

voters use a single-ballot plurality rule, while municipalities above this threshold adopt a dual-ballot system, to implement a regression discontinuity design. Our analysis shows that there is an increase in dirty campaigning when one moves from a single-ballot to a runoff plurality system. We find that right of reply lawsuits between the 1st and 2nd candidates are about 50 p.p. more likely under runoff elections, while lawsuits between the 2nd and 3rd candidates are about 20 p.p. more likely to occur under this voting system. Finally, a number of robustness checks suggest that our findings are not the result of random chance or any other discontinuity around the $200,000$ voters threshold.

Our paper contributes to the political economy literature in several ways. First, by proposing a model of decision on dirty campaigning, we bring some light to how candidates campaign strategies react to the characteristics of the race such as electoral competition and institutions. To the best of our knowledge, this is the first model to fully characterize the equilibrium and derive theoretical predictions in so many cases. In fact, the theoretical literature has had a hard time trying to pin down specific results, since most of the predictions depend on the exact distribution of vote shares among candidates and on the specification of how negative campaigning affects all candidates[2.1.3].

Second, by taking advantage of a unique feature of the Brazilian electoral legislation, we propose the use of right of reply lawsuits as an objective proxy for campaign tone, a feature of political campaigns that is subjective and usually very difficult to measure. Third, while there is a large empirical literature studying the effect of negative campaigns (see Lau et al. (1999), Lau, Sigelman e Rovner (2007) and Lau e Rovner (2009)), relatively few papers have analyzed the main factors that influence the candidates' decisions to "go dirty". In particular, to the best of our knowledge, no previous study has estimated the causal impact of the type of electoral system on the campaign strategies of candidates using quasi-experimental data.

[2.1.3] Ghandi, Iorio e Urban (2016) show that negative campaigns should be expected to occur more frequently in races where there are only two viable candidates ("duopoly"). The intuition is that an attack directed to a certain candidate generates not only a benefit to the attacker, but also "positive externalities"to all other candidates.

Fourth, while the previous empirical literature on the Duverger's Law (FUJIWARA, 2011) has provided evidence suggesting that voters behave strategically under two-ballot plurality systems by voting more to the 3rd and lower ranked candidate, we show that candidates and parties also have an incentive to behave strategically. In particular, our results provide evidence that candidates campaign "harder"and more aggressively under the runoff system. Finally, our paper contributes to theoretical literature on contests with sabotage and negative campaigning (LAZEAR, 1989; SKAPERDAS; GROFMAN, 1995; HARRINGTON; HESS, 1996; CHEN, 2003; DESPOSATO, 2007) which studies the role played by "negative"activities in promotions and electoral contests. Our paper also tests empirically some predictions of these models and establishes some stylized facts that can serve as a guide to more future theoretical research.

The rest of this paper is organized as follows. The next section presents the model and the testable hypothesis that come from it. Section 2 discusses the institutional backgroung and provides a detailed description of our database, how the variables were constructed, and presents some stylized facts about negative advertising in Brazil. Our empirical strategy, results and robustness checks are then presented in section 3. We conclude in section 4.

2.2 Theoretical framework

2.2.1 Two-Candidate Competition

Consider an election with two candidates $N = 2$, $i \in \{1, 2\}$, each with initial support (strength) $s_i^o > 0$ with $s_1^o > s_2^o$. The electoral support of an agent can be interpreted as a measure of his political capital and is assumed to be common knowledge. Candidates decide simultaneously whether to attack or not their opponents, with $a_i \in \{0, 1\}$ denoting candidate i's binary decision to attack or not. For convenience, let $n_i = a_j$ indicate whether candidate i was the attacked or not by j.

An attack allows a candidate to steal a fraction $\phi \in (0, 1)$ of his opponent's initial support. Our analysis will mostly concentrate on the case where ϕ is relatively small, since the empirical evidence suggests that the effect of negative campaigning tends to be small on average. The final level of support of candidate i is given by:

$$v_i\left(n_i, n_j\right) = \left(1 - \phi n_i\right) s_i^o + n_j \phi s_j^o + \epsilon_i \tag{2.2.1}$$

where ϵ_i is an iid shock with Type I Extreme Value distribution that captures, in a reduced form fashion, all the uncertainty associated with the electoral process. These shocks are realized after the agents' decisions have been made and are ex-ante unknown. As in Baron (1994) and Grossman e Helpman (1996), our model assumes that voters are "impressionable"and, therefore, respond to campaigning.

Elections are held by plurality and the candidate with highest final support wins. Following McFadden (1974), the probability of candidate i winning is given by:

$$p_i\left(n_i, n_j\right) = \Pr(v_i\left(n_i, n_j\right) > v_j\left(n_i, n_j\right)) = \frac{\exp\left(s_i\left(n_i, n_j\right)\right)}{\exp\left(s_i\left(n_i, n_j\right)\right) + \exp\left(s_j\left(n_i, n_j\right)\right)} \tag{2.2.2}$$

where $s_i\left(n_i, n_j\right) = \left(1 - \phi n_i\right) s_i^o + n_j \phi s_j^o$.[2.2.1] The cost of an attack is given by a constant $c > 0$, which captures all the expenses involved, including those related to an ensuing litigation. Candidates seek to maximize their probability of winning the election net of attacking costs:

$$u_i\left(n_i, n_j\right) = p_i\left(n_i, n_j\right) - n_j c \tag{2.2.3}$$

Equilibrium Analysis.

We now proceed to characterize the Nash equilibria of the game. A (mixed) strategy for player i is given by $\sigma_i \in [0, 1]$ which represents the probability with which he attacks his opponent. For n_i fixed, the benefit of candidate i attacking j is given by:

$$\Delta_{ij}\left(n_i\right) = p_i\left(n_i, 1\right) - p_i\left(n_i, 0\right), \tag{2.2.4}$$

i.e. the difference between the probability of winning the election when he attacks, $p_i\left(n_i, 1\right)$, and does not attack his opponent, $p_i\left(n_i, 0\right)$. The best-response

[2.2.1] For simplicity, we assume that there are no undecided voters. However, the inclusion of this type of voters in the model would not change any of the main qualitative results of the analysis.

correspondence of candidate i can then be expressed as:

$$b_i(\sigma_j) = \begin{cases} \sigma_i = 1 & \text{if } \sigma_j \Delta_{ij}(1) + (1-\sigma_j)\Delta_{ij}(0) > c \\ \sigma_i \in [0,1] & \text{if } \sigma_j \Delta_{ij}(1) + (1-\sigma_j)\Delta_{ij}(0) = c \\ \sigma_i = 0 & \text{if } \sigma_j \Delta_{ij}(1) + (1-\sigma_j)\Delta_{ij}(0) < c \end{cases}, \tag{2.2.5}$$

where $\sigma_j \Delta_{ij}(1) + (1-\sigma_j)\Delta_{ij}(0)$ represents the candidate i's expected utility of attacking j, given that j attacks i with probability σ_j. It follows immediately that an equilibrium always exists.

The following lemma establishes some basic properties of the function $\Delta_{ij}(n_i)$.

Lemma 1. *The function $\Delta_{ij}(n_i)$ satisfies the following properties:*

i. The benefit of an attack for the front-runner is larger when he is attacked:

$$\Delta_{12}(0) < \Delta_{12}(1)$$

ii. The benefit of an attack for the runner-up is larger when he is not attacked:

$$\Delta_{21}(1) < \Delta_{21}(0)$$

iii. The runner-up is more aggressive than the front-runner in the following sense:

$$\Delta_{12}(1) < \Delta_{21}(0)$$

$$\Delta_{12}(0) < \Delta_{21}(1)$$

Interestingly, the front-runner is more willing to attack when he himself is attacked, $\Delta_{12}(0) < \Delta_{12}(1)$. Ceteris paribus, receiving an attack reduces candidate 1's support, which makes it more likely that an attack on candidate 2 is actually pivotal for the outcome of the election. The exact opposite result holds for the runner-up. The runner-up is less willing to attack when attacked, $\Delta_{21}(1) < \Delta_{21}(0)$. Note that in this case he is further away from the front-runner, which makes it less likely that an attack on candidate 1 is decisive. Finally, we show that the runner-up is generally the more aggressive candidate in that both the upper and lower bounds of the benefits of attacking are larger for the runner-up than for the front-runner, $\Delta_{12}(1) < \Delta_{21}(0)$ and $\Delta_{12}(0) < \Delta_{21}(1)$.

The next proposition[2.2.2] provides a characterization of the structure of the equilibria.

Proposition 1. *The unique equilibrium of the game is such that:*

i. Both candidates attack if and only if:

$$c < \min\{\Delta_{12}(1), \Delta_{21}(1)\}$$

ii. The candidates attack with probabilities σ_1 and σ_2 strictly between 0 and 1 if and only if:

$$\Delta_{21}(1) < c < \Delta_{12}(1)$$

iii. Only the runner-up attacks if and only if:

$$\Delta_{12}(1) < c < \Delta_{21}(0)$$

iv. Neither candidate attacks if and only if:

$$c > \Delta_{21}(0)$$

Our analysis shows that there are four different types of equilibria in this game. Figure 2.2.1 depicts the region of the parameters where each type of equilibrium exists assuming that $\Delta_{21}(1) < \Delta_{12}(1)$.[2.2.3] Observe that as the cost of attacking increases, we move in sequence across the following equilibria: (*i*) equilibrium where both attack, (*ii*) equilibrium in mixed strategies, (*iii*) equilibrium where only candidate 2 attacks and (*iv*) equilibrium where no one attacks. The equilibrium in mixed strategies arises in the region where candidate 1 attacks only if attacked while candidate 2 attacks only if not attacked, so that some degree of randomization is necessary. Also, note that the equilibrium where only candidate

[2.2.2] Formal proofs of all propositions will be available in the next version of the paper.

[2.2.3] When $\Delta_2(1) > \Delta_1(1)$, the equilibrium in mixed strategies does not exist but all the other three classes of equilibria remain.

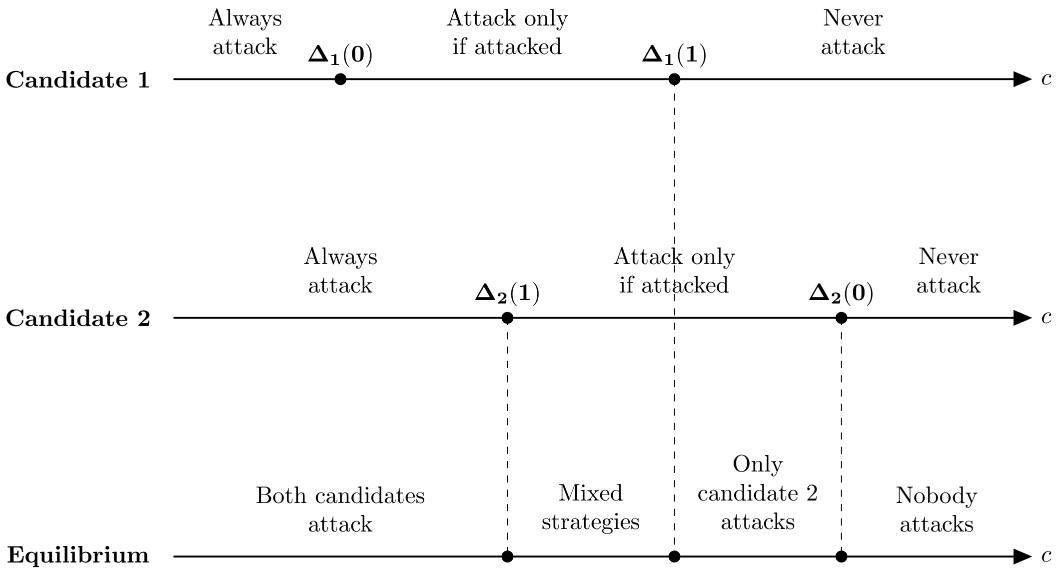


Figura 2.2.1 – Game equilibrium in a two-candidate competition under single-ballot system.

2 attacks occurs in the region where candidate 1 never attacks while candidate 2 attacks only if not attacked. Interestingly, our analysis shows that the likelihood of an attack is not necessarily decreasing in c.

We complement our analysis by deriving some comparative static results. The following lemma characterizes how the benefit of an attack $\Delta_{ij}(n_i)$ varies with the initial support levels, s_1^o and s_2^o.

Lemma 2. *There exists $\overline{\phi} \in (0, 1/2)$ such that if $\phi < \overline{\phi}$, then:*

i. The benefit of an attack for the front-runner decreases with s_1^o and increases with s_2^o, that is:

$$\frac{\partial \Delta_{12}(n_1)}{\partial s_1^o} < 0 \quad \textit{and} \quad \frac{\partial \Delta_{12}(n_1)}{\partial s_2^o} > 0 \quad \textit{for any } n_1 \in \{0, 1\}$$

ii. The benefit of an attack for the runner-up increases with his both s_1^o and s_2^o, that is:

$$\frac{\partial \Delta_{21}(n_2)}{\partial s_1^o} > 0 \quad \textit{and} \quad \frac{\partial \Delta_{21}(n_2)}{\partial s_2^o} > 0 \quad \textit{for any } n_2 \in \{0, 1\}$$

Intuitively, the benefit of an attack $\Delta_{ij}(n_i)$ increases with the initial support level of one's rival s_j^o. Moreover, in line with our previous discussion, we show that while the front-runner becomes less aggressive when his initial support increases, $\partial\Delta_{12}/\partial s_1^o < 0$, the opposite result holds for the runner-up, $\partial\Delta_{21}/\partial s_2^o > 0$. The next proposition tell us what we should expect to observe in equilibrium as a result of changes in the initial support levels.

Proposition 2. *There exists* $\overline{\phi} \in (0, 1/2)$ *such that if* $\phi < \overline{\phi}$*, then:*

i. The front-runner becomes more likely to attack when s_2^o *increases.*

ii. The runner-up becomes more likely to attack when s_1^o *increases.*

Thus, our analysis shows that the region of parameters where the front-runner attacks increases with the initial support of his opponent, s_2^o. Similarly, the region of parameters where we expect the runner-up to attack increases with s_1^o.

2.2.2 Three-Candidate Competition

We now consider the case with three candidates, with $s_1^o > s_2^o > s_3^o > 0$. Candidates decide simultaneously whether to attack or not each of their opponents, with $a_{ij} \in \{0, 1\}$ denoting candidate i's binary decision to attack j. For simplicity, we suppose that candidates can choose to attack at most one other competitor. Let $n_i = \sum_k a_{ki}$ represent the number of attacks received by candidate i, with $n = (n_1, n_2, n_3)$ and $n \in \mathcal{N} = \{0, 1, 2\}^3$. The final level of support of candidate i is given by:

$$v_i(n) = (1 - \phi n_i) s_i^o + \sum_{j \neq i} \frac{n_j \phi s_j^o}{2} + \epsilon_i \qquad (2.2.6)$$

where we assume that an attack on candidate j benefits the two other candidates equally, so that each gets $\phi s_j^o/2$. As before, ϵ_i is an iid shock with Type I Extreme Value distribution.

When elections are held by single-ballot plurality, the probability of winning is given by:

$$p_i^{SB}(n) = \frac{\exp(s_i(n))}{\sum_{k=1}^{3} \exp(s_k(n))} \tag{2.2.7}$$

We also consider the situation where elections are held by dual-ballot (runoff) plurality, in which case the probability of advancing to the second-round is:

$$p_i^{DB}(n) = p_i^{SB}(n) + p_j^{SB}(n) \frac{\exp(s_i(n))}{\exp(s_i(n))+\exp(s_k(n))} + p_k^{SB}(n) \frac{\exp(s_i(n))}{\exp(s_i(n))+\exp(s_j(n))}, \tag{2.2.8}$$

where $i, j, k \in \{1, 2, 3\}$, $i \neq j$, $i \neq k$ and $j \neq k$. Note that this expression gives the probability that candidate i finishes either in first or second place.

As before, candidates seek to maximize their probability of winning or advancing to the second round net of attacking costs:

$$u_i^{\ell}(a) = p_i^{\ell}(n) - (a_{ij} + a_{ik})\, c \tag{2.2.9}$$

where $\ell \in \{SB, DB\}$ and $a = (a_1, a_2, a_3)$ with $a_i = (a_{i1}, a_{i2}, a_{i3})$. Note that, under dual-ballot plurality, we assume that candidates attribute the same weight to the utilities of being placed first or second. However, our main qualitative results are robust to relaxing this assumption.

Equilibrium Analysis.

A (mixed) strategy for player i is given by $\sigma_i = (\sigma_{i1}, \sigma_{i2}, \sigma_{i3})$, with $\sigma_{ij} \in [0, 1]$ and $\sum_j \sigma_{ij} = 1$ for $i \in \{1, 2, 3\}$, where σ_{ij} represents the probability with which i attacks j. For a fixed vector $n \in \mathcal{N}$, the benefit of candidate i attacking j is given by:

$$\Gamma_{ij}(n) = p_i(n_i, n_j + 1, n_k) - p_i(n_i, n_j, n_k), \tag{2.2.10}$$

i.e. the difference between the probability of winning the election when i attacks j and when he does not attack j.

Let σ_{-i} represent the strategies of player i's opponents and note that it induces a probability distribution over $\mathcal{N}$. Denote by $\pi(n|\sigma_{-i})$ the probability of $n \in \mathcal{N}$ given σ_{-i}. Thus, the expected benefit of an attack of candidate i on j is given by:

$$U_i\left(j,\sigma_{-i}\right)=\sum_{n\in\mathcal{N}}\pi\left(n;\sigma_{-i}\right)\Gamma_{ij}\left(n\right) \tag{2.2.11}$$

The best-reply correspondence of candidate i, $b_i\left(\sigma_{-i}\right)$, can be expressed in this case similarly to (2.2.5) except that here candidate i's optimal strategy is to attack j if and only if $U_i\left(j,\sigma_{-i}\right)>c$ and $U_i\left(j,\sigma_{-i}\right)>U_i\left(k,\sigma_{-i}\right)$ for $k\neq j$. As before, the best-response correspondence is well-behaved, so that an equilibrium always exists.

While a full characterization of the equilibria of the model in an environment with three candidates is considerably more difficult, we are able to derive some general results. We start by analyzing the case of single-ballot pluraliy. The following proposition establishes basic characteristics of the benefit function $\Gamma_{ij}^{SB}\left(n\right)$ and derives implications for the equilibrium.

Proposition 3. *Under single-ballot plurality, there exists* $\overline{\phi}\in(0,1/2)$ *such that if* $\phi<\overline{\phi}$*, then:*

i. For candidate 1, the benefit of an attack on candidate 2 is always larger than that of an attack on candidate 3:

$$\Gamma_{12}^{SB}\left(n\right)>\Gamma_{13}^{SB}\left(n\right)\qquad \textit{for any } n\in\mathcal{N}$$

Thus, in equilibrium, candidate 1 either attacks candidate 2 or no one.

ii. For candidate 2, the benefit of an attack on candidate 1 is always larger than that of an attack on candidate 3:

$$\Gamma_{21}^{SB}\left(n\right)>\Gamma_{23}^{SB}\left(n\right)\qquad \textit{for any } n\in\mathcal{N}$$

Thus, in equilibrium, candidate 2 either attacks candidate 1 or no one.

iii. For candidate 3, the benefit of an attack on candidate 1 is always larger than that of an attack on candidate 2:

$$\Gamma_{31}^{SB}\left(n\right)>\Gamma_{32}^{SB}\left(n\right)\qquad \textit{for any } n\in\mathcal{N}$$

Thus, in equilibrium, candidate 3 either attacks candidate 1 or no one.

Intuitively, each candidate benefits more by attacking the highest ranked candidate among his opponents. Thus, candidate 1 prefers to attack 2 than 3, while candidates 2 and 3 prefer to attack 1 rather than any other candidate. Next, we compare the benefits of an attack between the front-runner and the runner-up in an environment with two versus three candidates, both under simple (single-ballot) plurality rule. Should we expect more or less attacks when a third candidate is added to the race? The next proposition provides a formal answer.

Proposition 4. *Under simple (single-ballot) plurality, and for any s_1^o, s_2^o and s_3^o with $s_1^o > s_2^o > s_3^o > 0$, there exists $\overline{\phi} \in (0, 1/2)$ such that if $\phi < \overline{\phi}$, then:*

i. For candidate 1, the benefit of an attack on candidate 2 is larger when there are only two candidates:

$$\Delta_{12}(n_1) > \Gamma_{12}^{SB}(m) \quad \text{for any } n_1 \in \{0, 1\} \text{ and } m \in \mathcal{N}$$

Furthermore, if in a three-candidate competition there is an equilibrium in which candidate 1 attacks 2, then in equilibrium candidate 1 must also attack 2 in a two-candidate competition.

ii. For candidate 2, the benefit of an attack on candidate 1 is larger when there are only two candidates:

$$\Delta_{21}(n_2) > \Gamma_{21}^{SB}(m) \quad \text{for any } n_2 \in \{0, 1\} \text{ and } m \in \mathcal{N}$$

Furthermore, if in a three-candidate competition there is an equilibrium in which candidate 2 attacks 1, then in equilibrium candidate 2 must also attack 1 in a two-candidate competition.

Our analysis shows that candidates 1 and 2 are more aggressive towards each other when there are only two candidates in the electoral race ("duopoly"). Intuitively, the presence of a third candidate dilutes the benefit of a campaign attack, given that the resulting gains are now split between two candidates. Thus, in equilibrium, we expect to observe more attacks between the front-runner and the runner-up in a two-candidate competition.

Finally, we can show that the pattern of attacks may change substantially when the voting rule changes from single-ballot to dual-ballot plurality. In particular, candidates 2 and 3 may have an incentive to attack each other under dual-ballot plurality. The next proposition summarizes our main comparative result with this respect.

Proposition 5. *Under dual-ballot plurality, and for any s_2^o and s_3^o with $s_2^o > s_3^o > 0$, there exists $\overline{\phi} \in (0, 1/2)$ and $\overline{s}_1 > 0$ such that if $\phi < \overline{\phi}$ and $s_1^o > \overline{s}_1$ then:*

i. For candidate 2, the benefit of an attack on candidate 3 is larger than that of an attack on candidate 1:

$$\Gamma_{21}^{DB}(n) < \Gamma_{23}^{DB}(n) \qquad \text{for any } n \in \mathcal{N}$$

Thus, in equilibrium, candidate 2 either attacks candidate 3 or no one.

ii. For candidate 3, the benefit of an attack on candidate 2 is larger than that of an attack on candidate 1:

$$\Gamma_{31}^{DB}(n) < \Gamma_{32}^{DB}(n) \qquad \text{for any } n \in \mathcal{N}$$

Thus, in equilibrium, candidate 3 either attacks candidate 2 or no one.

Intuitively, if the initial electoral support of the front-runner, s_1^o, is large enough, then under dual-ballot plurality it pays for candidates 2 and 3 to concentrate their attacking efforts on each other in an attempt to obtain a place in the run-off election. While it is well-known from Duverger's Law that voters have an incentive to behave differently under single-ballot and dual-ballot plurality, our analysis shows that candidates and political parties may also have an incentive to adopt different campaign strategies under these two electoral systems.

We now turn to the data to see if these predictions are consistent with the empirical evidence. We begin by describing the political environment in Brazil and then move to estimations using a novel dataset containing judicial information about Brazil's mayors elections of 2012 and 2016.

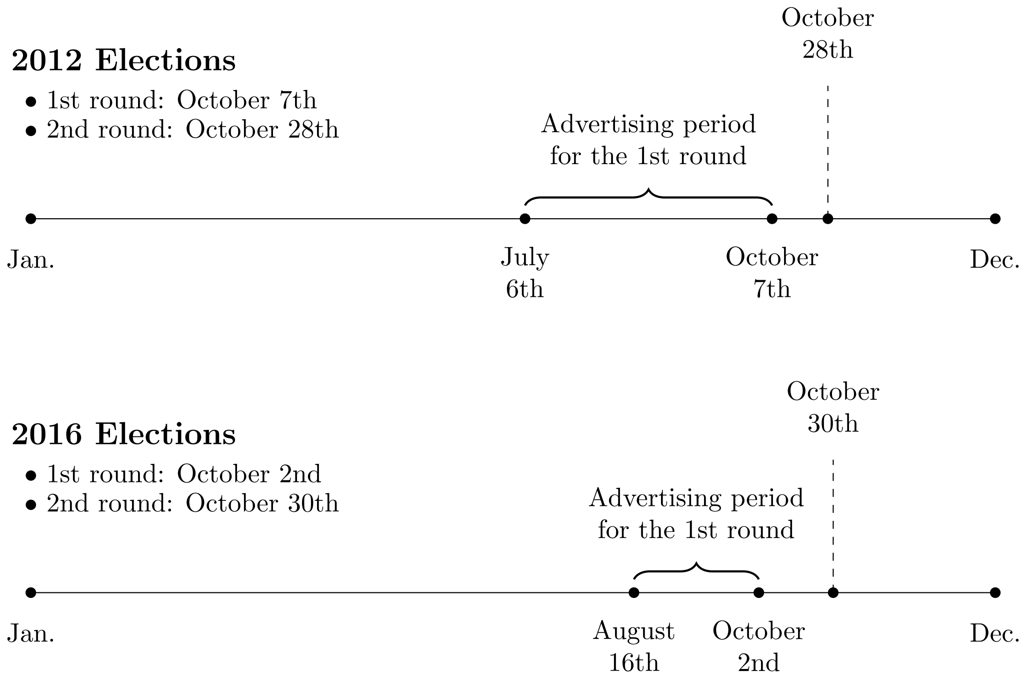


Figura 2.3.1 – Timeline of 2012 and 2016 municipal elections.

2.3 Negative advertising in Brazil

2.3.1 Political environment and advertising in Brazil

At every four years brazilians vote for mayors in more than 5,000 municipalities. Voters must choose among candidates that represent an individual party or a coalition of parties in a multi-party system. Municipalities, however, differ in which voting method they use to elect a single winner: the Constitution states that municipalities with less than $200,000$ voters must adopt a single-ballot plurality system, while municipalities above this threshold may have a second round if no candidate achieves the majority of votes in the first round (runoff system). Nevertheless, first round takes place in the same day in all municipalities and electoral law states the day from which political advertising is allowed. Figure 2.3.1 shows these dates for 2012 and 2016 races.

Despite candidates being proihibited to buy airtime at radio and TV to use it for campaign advertising[2.3.1], there is no restrictions on other medias. Every ad, however, must be in compliance with the electoral law[2.3.2]. The law states that any party or candidate that had been the target of a slanderous, defamatory or untrue advertising has the right to demand the use of the offender's media space to reply the ad content by filing an electoral lawsuit. These lawsuits, called "right of reply"lawsuits, are judged by an electoral court that may require the defendant to publish the plaintiff's reply, to pay for the reply's publishing costs, and to pay a fine that could reach more than $15,000$ BRL in 2016.

In fact, the multi-party system guarantees a natural variability in the number and types of candidates disputing the seat while the discontinuity on electoral rules guarantees a quasi-experimental setting to assess the effect of electoral institutions. By doing so, these institutional features and the law on political advertising provide an environment that is suitable to evaluate how the level of competition, candidates characteristics, and electoral institutions affect the candidates decisions to "go dirty"on their campaigns.

2.3.2 Data description

Our first main source of data is the Brazil's Federal Electoral Court, the *Tribunal Superior Eleitoral* (TSE), from which we collected data on elections results, municipalities and candidates' attributes for all brazilian municipalities in the 2012 and 2016 mayors elections[2.3.3]. We use the final vote share of candidates in the election's first round, gender, and educational background to define a profile for each candidate. We also use the final vote share to define what we call an effective candidate, a candidate who earned at least 15% of the first round election votes Our analysis is robust to specifying different values for this threshold.. Moreover, we compile data on electorate size, electorate education and turnout to assess some of the municipality caractheristics.

[2.3.1] Political advertising time in TV and radio are distributed according to the number of seats in the parlamentary the candidates' parties have.
[2.3.2] Available in portuguese at <http://www.planalto.gov.br/ccivil_03/leis/L9504.htm>.
[2.3.3] <http://www.tse.jus.br/eleicoes/estatisticas/repositorio-de-dados-eleitorais>

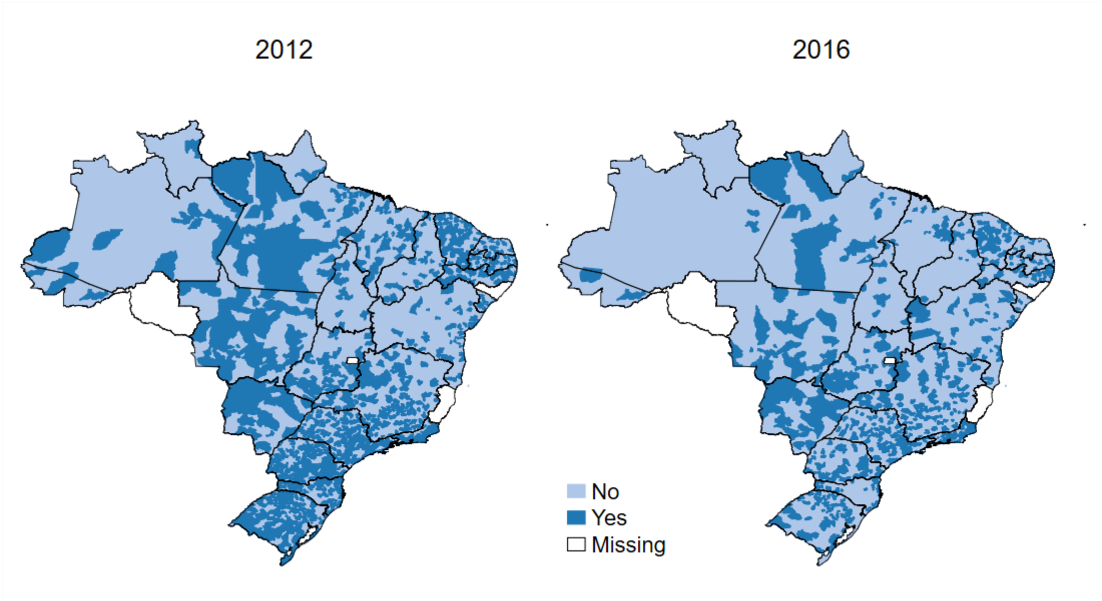


Figura 2.3.2 – Municipalities for which we observe at least one RR case.

To quantify advertising negativeness in these years we use the “right of reply"lawsuits (RR henceforth)[2.3.4]. We collected data on $10,924$ RR cases for over 23 brazilian states (from a total of 26) directly from each of the Brazil’s Regional Electoral Courts, which constitutes our second main source of data. Figure 2.3.2 shows that this data on RR is widespread in the country municipalities for both years.

For each RR we retrieved information on the plaintiffs, defendants, date and municipality in which the litigation took place. For a reduced subsample, we also retrieved dates on the litigation stages and if the judge have considered the defendant guilty. A first look in the data shows that this type of case is handled quickly in the court as it should be, with more than 75% of cases being closed within 13 calendar days.

We measure negativity as an indicator variable that equals one if we identify at least one RR in which the candidate A is suing candidate B in the period of the

[2.3.4] Silveira e Mello (2011) use a similar approach to define negativity, despite using only the number of hits of a search for the expression “direito de resposta"in the Tribunal Superior Eleitoral website as their measure.

first round campaign. Moreover, we define negativity using “at least one RR"and not the absolute number of RR to avoid double counting of the same litigation in the database construction. We also consider all candidates pairs from municipalities of states that we have data but in which we do not observe any RR lawsuit as not engaging in negative advertising. In those cases we replace our indicator variable by 0.

Notice, however, that having a RR involving the candidate pair do not guarantee that the judge have considered the defendant guilty. Aside from the fact that we are able to retrieve only a fraction of the litigations results, to assess the litigation result of each RR is susceptible to measurement error, since the result can be judicially reviewed in an appeal court more than once. Even exploring only the litigations extensive margin, this campaign negativeness measure is somewhat less subjective than evaluating each advertising by its “negative tone": electoral rules in Brazil are defined by law, in which there is a somewhat clear definition of what is considered an ad susceptible to RR and what is not.

Sample and descriptive statistics. We merged the judicial data on RR lawsuits with the TSE dataset using candidates and the mayors' coalitions full names. We leave with a dataset in which we identify the pair of candidates involved on each litigation along with the pair characteristics and the municipality attributes, which constitutes an unique and valuable dataset about negative advertising in politics. Table 2.3.1 shows some characteristics of the data we used in regressions. The final sample used in estimations is at the municipality-pair-year level, with a total of $69,254$ observations in $5,328$ unique municipalities. Panel B of table 2.3.1 shows some statistics on pairs of candidates, including gender, and education. About negativity itself, 6.39% of all observations regarding the two best placed candidates ($1st$ attacking $2nd$ or $2nd$ attacking $1st$) had a RR in 2012 and 2016, showing that this type of lawsuit is indeed relevant in the brazilian electoral process.

Tabela 2.3.1 – Summary statistics

	Mean	Std. Dev.	Min	Max	N
Panel A: Municipality-level					
Electorate composition (2012 and 2016)					
- Number of voters (thousands)	25.80	156.6	0.9	8,886.3	10,462
- Number of voters (% population)	78.04	9.9	30.2	100.0	9,819
- College (% voters)	2.89	2.5	0.1	33.1	10,462
Elections (2012 and 2016)					
- Number of candidates in the municipality	2.74	1.2	0.0	14.0	10,462
- Number of candidates with at least 15% in final vote share	2.14	0.5	0.0	4.0	10,462
- Turnout (% voters)	85.89	5.7	64.4	98.8	10,462
Municipalities characteristics (2010 Census)					
- Income per capita (th BRL)	494.91	244.0	96.3	2,043.7	10,453
- Informality in the labor market (% total employess)	56.42	19.4	10.9	97.0	10,453
- Gini index	0.49	0.1	0.3	0.8	10,453
- Human Development Index	0.66	0.1	0.4	0.9	10,453
Panel B: Pair-level					
Pair demographics (2012 and 2016)					
- Probability of both candidates being men	75.33	43.1	0.0	100.0	69,254
- Probability of both candidates having college degree	31.66	46.5	0.0	100.0	69,254
Negativity on municipalities (2012 and 2016)					
- Probability of having a RR involving $1st - 2nd$ pair	6.39	24.4	0.0	100.0	20,924
- Probability of having a RR involving $1st - 3rd$ pair	2.12	14.4	0.0	100.0	10,114
- Probability of having a RR involving $2nd - 3rd$ pair	1.20	10.9	0.0	100.0	10,114

2.4 Econometric methodology and results

We are now interested in testing wheter the model's predictions hold to the data. Specifically,

H1. **Opponent's support.** Is a candidate, in a single-ballot race with only two candidates, more likely to attack its opponent when the opponent's initial support increases as in proposition 2?

H2. **Intensity of attacks.** Are candidates 1 and 2 more likely to attack each other and candidate 3 more likely to attack candidate 1, under single-ballot plurality in a three-candidate competition, as in proposition 3?

H3. **Competition effect.** Are the two best placed candidates, in a single-ballot race, more likely to attack each other when there are only two candidates in the electoral race as in proposition 4?

H4. **Electoral rules.** Are the second and third best placed candidates more likely to attack each other when the voting rule changes from single-ballot to dual-ballot plurality system as in proposition 5?

By empirically testing these predictions we seek to document novel aspects of negative advertising and candidates decisions on campaign strategies. To do so, we use both a fixed-effect regression methodology and a regression discontinuity design in order to better assess the multidimensional aspects of our dataset. We also present a series of robustness checks for both cases.

When presenting the results we further restricts the estimation sample to municipalities with two or three effective candidates so the model and the dataset become fully comparable. This isn't so restrictive since more than 82% of the municipalities in 2012 and 2016 had at most 3 candidates running for the seat, and 99% had at most 3 candidates with at least 15% of the first round election votes running for the seat.

2.4.1 Fixed-effect regressions

Testing the opponent's support hypothesis (H1). In order to evaluate the predictions of proposition 2 we use a fixed-effect regression following the equation:

$$Y_{ijmt} = \beta_0 + \beta_1 \cdot VOTE_{jmt} + \beta_2 \cdot INCUMBENT_{jmt} + \Gamma.X_{ijmt} + c_m + \epsilon_{ijmt} \quad (2.4.1)$$

where Y_{ijmt} is a dummy variable that equals 1 if there was a lawsuit about negative advertising in which candidate i attacked candidate j in the municipality m, year t. $VOTE_{jmt}$ is the first round vote share of candidate j, and $INCUMBENT_{jmt}$ is a dummy variable that equals 1 if candidate j is an incumbent. X_{ijmt} is a vector of control variables, and c_m is a municipality fixed-effect. Ideally we would like to regress Y_{ijmt} on the initial support of candidate j. Unfortunately, this is an unobserved characteristic of candidates in municipal elections[2.4.1]. We then use

[2.4.1] Results of pre-election polls, which may be seen as the immediate proxy for candidates' initial support, were available only for a few large brazilian municipalities.

incumbency and the final vote share of candidates as regressors, given that both are very likely correlated with the initial support of candidates[2.4.2].

As controls we first include a dummy for the year of 2016 which captures any systematic difference between the 2012 and 2016 elections. Indeed, new rules on campaign expenses took place in september 2015 as part of a campaign finance reform, which led to less money available to candidates during elections (AVIS et al., 2017). By imposing a cap on the amount of money candidates could spend, this reform could have also affected the likelihood of negative advertising. Figure 2.3.2 shows that negative advertising was indeed more frequent in 2012 than in 2016 elections.

The X_{ijmt} vector also contains some characteristics of the candidate pair. We include as regressors a dummy indicating if both candidates are men, a dummy indicating if one candidate is a man and the other is a woman, and a dummy indicating if both candidates have a college degree. Moreover, the electorate size and education may also be predictive of negative advertising. To control for it, we add the number of voters in the municipality, the fraction of the electorate with college degree and the fraction of the electorate that is illiterate. In this specification, we further include as regressors the final vote share of candidate i and an indicator variable that equals one if the candidate i was an incumbent.

Table 2.4.1 present the estimates of equation (2.4.1) for the case of 2 effective candidates. We find that one of the two best placed candidates is more likely to attack its opponent during the first round campaign when the opponent final vote share is higher and when the opponent is an incumbent. Our OLS estimations for β_1 and β_2 are statiscally significant in all specifications, even when controling for municipality fixed effects and candidates' characteristics. Moreover, coefficients are stable across specifications and magnitude is substantial. In column (4), our prefered specification, we find that being an incumbent increases in 1.3 percentage points the probability of being attacked relative to the mean of 6.0%.

[2.4.2]Political economy and political science literatures indeed suggest the existence of an *a priori* incumbency advantage in most elections (PASTINE; PASTINE, 2012). This advantage may be evidence of a positive correlation between incumbency and innitial support. The correlation between innitial support and final vote share is straightforward though.

Tabela 2.4.1 – Candidates incentives to “go dirty": races with 2 effective candidates on single-ballot municipalities

	Dependent variable: lawsuit dummy			
	(1)	(2)	(3)	(4)
Attacked candidate vote share	−0.1064***	0.1700***	0.1495***	0.1764**
	[0.0400]	[0.0419]	[0.0426]	[0.0727]
Attacked candidate is an incumbent	0.0225***	0.0243***	0.0242***	0.0137**
	[0.0052]	[0.0050]	[0.0050]	[0.0069]
Controls	No	Yes	Yes	Yes
State fixed effects	No	No	Yes	No
Municipality fixed effects	No	No	No	Yes
Number of municipalities-year	4,924	4,924	4,924	4,924
Observations	16,480	16,480	16,480	16,480
Adj. R-squared	0.003	0.068	0.079	0.286
Dependent variable mean	0.060	0.060	0.060	0.060

Notes: In brackets, standard errors are clustered at the municipality level. Significance at the 10% level is represented by $*$, at the 5% level by $**$, and at the 1% level by $***$. Controls includes a dummy for 2016, the voters number in the municipality, the fraction of the municipality electorate that is illiterate and with college degree, a dummy indicating if both candidates are men, if both candidates have a college degree, and a dummy indicating if one candidate is a man and the other is a woman.

Results on β_1 and β_2 are consistent with predictions of proposition 2. By assuming that both the final vote share and incumbency are positive correlated to the candidate's initial support, the positive and statistical significant estimates of β_1 and β_2 suggest that candidates become more likely to attack an opponent when the opponent's initial support increases, even when we control for a wide range of municipality and candidates characteristics.

Testing the intensity of attacks hypothesis (H2). First, we present qualitative evidence on the intensity of attacks among the best three placed candidates in a three-candidate competition under single-ballot system. Figure 2.4.1 shows the probabilities of each candidate attacking his opponents according to our dataset

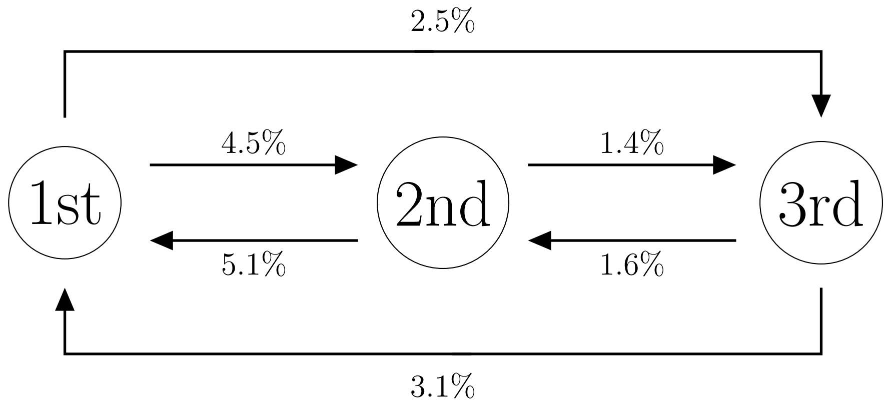


Figura 2.4.1 – Probabilities of attack among candidates.

on right of reply lawsuits. In fact, candidate 1 and 2 are more likely to attack each other and candidate 3 more likely to attack candidate 1.

Second, we use a fixed-effect regression with dummies identifying each pair of candidates as regressors following the equation:

$$\begin{aligned} Y_{ijmt} = \beta_0 &+ \beta_1 \cdot D_{12mt} + \beta_2 \cdot D_{13mt} \\ &+ \beta_3 \cdot D_{21mt} + \beta_4 \cdot D_{23mt} \\ &+ \beta_5 \cdot D_{31mt} + \beta_6 \cdot D_{32mt} + \Gamma . X_{ijmt} + c_m + \epsilon_{ijmt} \end{aligned} \tag{2.4.2}$$

where Y_{ijmt} is the same negativity dummy as before, and D_{12mt} is a dummy variable that equals one if candidate i is the best placed candidate in the race and j is candidate 2. X_{ijmt} is the same vector of control variables of equation (2.4.1), and c_m is a municipality fixed-effect.

Table 2.4.2 shows the estimation results. Column (1) is the regression counterpart of the probabilities diagram in figure 2.4.1. In columns (2), (3) and (4) we add controls, state and municipalities fixed effects to this regression. In all cases, $\hat{\beta}_1 > \hat{\beta}_2$, $\hat{\beta}_3 > \hat{\beta}_4$, and $\hat{\beta}_5 > \hat{\beta}_6$, which is exactly what we should expect given the predictions on the intensity of the attacks of proposition 3.

To evaluate if these differences between coefficients are statistically signifi-

Tabela 2.4.2 – Candidates' intensity of attacks: races with 3 effective candidates on single-ballot municipalities

	Dependent variable: lawsuit dummy			
	(1)	(2)	(3)	(4)
1st attacking 2nd (β_1)	0.0453***	0.0467***	0.0456***	0.0436***
	[0.0053]	[0.0056]	[0.0056]	[0.0061]
1st attacking 3rd (β_2)	0.0251***	0.0266***	0.0254***	0.0235***
	[0.0039]	[0.0043]	[0.0043]	[0.0053]
2nd attacking 1st (β_3)	0.0515***	0.0530***	0.0519***	0.0499***
	[0.0057]	[0.0059]	[0.0059]	[0.0062]
2nd attacking 3rd (β_4)	0.0145***	0.0161***	0.0150***	0.0129***
	[0.0031]	[0.0035]	[0.0035]	[0.0045]
3rd attacking 1st (β_5)	0.0314***	0.0329***	0.0317***	0.0298***
	[0.0044]	[0.0047]	[0.0048]	[0.0055]
3rd attacking 2nd (β_6)	0.0163***	0.0180***	0.0169***	0.0148***
	[0.0032]	[0.0036]	[0.0037]	[0.0047]
H0: $\beta_1 = \beta_2$	0.001	0.001	0.001	0.002
H0: $\beta_3 = \beta_4$	0.000	0.000	0.000	0.000
H0: $\beta_5 = \beta_6$	0.004	0.005	0.005	0.006
Controls	No	Yes	Yes	Yes
State fixed effects	No	No	Yes	No
Municipality fixed effects	No	No	No	Yes
Number of municipalities-year	1,386	1,386	1,386	1,386
Observations	18,714	18,714	18,714	18,714
Adj. R-squared	0.032	0.053	0.038	0.119

Notes: In brackets, standard errors are clustered at the municipality level. Significance at the 10% level is represented by $*$, at the 5% level by $**$, and at the 1% level by $***$. Controls includes a dummy for 2016, the voters number in the municipality, the fraction of the municipality electorate that is illiterate and with college degree, a dummy indicating if both candidates are men, if both candidates have a college degree, and a dummy indicating if one candidate is a man and the other is a woman.

cant, it is also reported in table 2.4.2 the p-value of Wald tests against the null hypotheses of $\beta_1 = \beta_2$, $\beta_3 = \beta_4$, and $\beta_5 = \beta_6$. We reject the null hypotheses in all specifications. Thus, even when controling for state and municipality fixed-effects, estimations of equation (2.4.2) are consistent with the theoretical predictions and the statement that, in single-ballot municipalities, candidate 1 and 2 are more likely to attack each other and candidate 3 more likely to attack candidate 1.

Testing the competition effect hypothesis (H3). In order to evaluate the predictions of proposition 4 we also use a fixed-effect regression but following a slightly different equation:

$$Y_{ijmt} = \alpha_0 + \alpha_1 \cdot DUOPOLY_{mt} + \Gamma.X_{ijmt} + c_m + \epsilon_{ijmt} \qquad (2.4.3)$$

where Y_{ijmt} is the same negativity dummy as before, and $DUOPOLY_{mt}$ is a dummy variable that equals one if municipality m in year t had only two effective candidates. X_{ijmt} is the same vector of control variables of equation (2.4.1), and c_m is a municipality fixed-effect.

Panel A of table 2.4.3 shows that the front-runner is more likely to attack the runner-up when they are the only two effective candidates in the race, and panel B shows that the inverse is also true. These results suggest that more electoral competition may decrease the likelihood of dirty campaigning between the best two placed candidates (GHANDI; IORIO; URBAN, 2016). Furthermore, there is also evidence that the effect on the runner-up is higher than on the front-runner. One more time, coefficients are stable across specifications and magnitude is substantial. In column (4), our prefered specification, we find that being on a duopoly increases the likelihood of the second placed candidate attacking the first placed candidate in 3.6 percentage points relative to the mean of 6.4%.

Results on α_1 are consistent with predictions of proposition 4. The positive and statistical significant estimates of α_1 suggest that the best two placed candidates become more likely to attack each other when there are only two candidates in the electoral race, even when we control for a wide range of municipality and candidates characteristics.

Tabela 2.4.3 – Effects of electoral competition on candidates decision to “go dirty"on single-ballot municipalities

	Dependent variable: lawsuit dummy			
	(1)	(2)	(3)	(4)
Panel A: 1st attacking 2nd				
Duopoly	0.0110**	0.0273***	0.0230***	0.0278**
	[0.0051]	[0.0050]	[0.0051]	[0.0109]
Adj. R-squared	0.000	0.053	0.063	0.163
Panel B: 2nd attacking 1st				
Duopoly	0.0172***	0.0383***	0.0350***	0.0365***
	[0.0056]	[0.0056]	[0.0058]	[0.0129]
Adj. R-squared	0.001	0.067	0.077	0.188
Controls	No	Yes	Yes	Yes
State fixed effects	No	No	Yes	No
Municipality fixed effects	No	No	No	Yes
Number of municipalities-year	5,251	5,251	5,251	5,251
Observations	10,299	10,299	10,299	10,299
Dependent variable mean	0.064	0.064	0.064	0.064

Notes: In brackets, standard errors are clustered at the municipality level. Significance at the 10% level is represented by $*$, at the 5% level by $**$, and at the 1% level by $***$. Controls includes a dummy for 2016, the voters number in the municipality, the fraction of the municipality electorate that is illiterate and with college degree, a dummy indicating if both candidates are men, if both candidates have a college degree, and a dummy indicating if one candidate is a man and the other is a woman.

Discussion. These fixed-effects regressions present evidence in favor of propositions 2, 3 and 4. In spite of the concerns of endogeneity that one may have since in these cases we are not in a quasi-experiment or natural experiment environment, we present the estimates in various specifications with a wide range of controls at the candidate-level and municipality-level. Estimates of our coefficients of interest are always statistically significant and stable across specifications, which gives us more confidence about our theoretical statements.

2.4.2 Regression Discontinuity Design

Testing the electoral rules hypothesis (H4). In order to test how the incentives to engage in dirty campaigning varies under single-ballot and runoff electoral systems we exploit the electoral discontinuity in brazilian municipal elections of $200,000$ voters. Just as before, let Y_{ijmt} be our negativity dummy for candidate pair ij in the municipality m in the year t, and R_{mt} be a dummy variable that equals one if the municipality has more than $200,000$ registered voters, and 0 otherwise. Additionally, let V_{mt} be the number of voters minus the $200,000$ threshold, and X_{ijmt} a vector of covariates[2.4.3]. Therefore, we can implement the RDD through a simple local linear regression of the equation below in which we use only observations that belong to a certain window h around the threshold.

$$Y_{ijmt} = \sum_{k=0}^{N} \lambda_k \cdot (V_{mt})^k + R_{mt} \cdot \sum_{k=0}^{N} \gamma_k \cdot (V_{mt})^k + \Pi.X_{ijmt} + \epsilon_{ijmt} \tag{2.4.4}$$

In all pairs subsamples and for a sufficiently narrow bandwidth h, the local average treatment effect is consistently estimated by $\hat{\gamma_0}$, just as presented in Imbens e Lemieux (2008). To put it simply,

$$\gamma_0 = \lim_{h \to 0} E[Y|voters = 200,000 + h] - \lim_{h \to 0} E[Y|voters = 200,000 - h]$$

and $\hat{\gamma_0} \to \gamma_0$. We are, however, cautious about the choice h and the polynomial order N which better fits the underlying data around the threshold, and present results for different choices of h and N.

Table 2.4.4 summarizes the estimations of γ_0 for different bandwidth (h) sizes, and $N \in \{0, 1\}$. We also present the coefficients estimates with the inclusion

[2.4.3] Despite not playing an important role in the RDD identification hypothesis, covariates can be important to improve the precision of estimates, which may be important when we have few observations around the threshold Imbens e Lemieux (2008).

and without the inclusion of covariates[2.4.4]. Results suggest there is a positive and substantial effect of runoff on the likelihood of $1st - 2nd$ and $2nd - 3rd$ pairs engaging in negative advertising. Panel B shows that the likelihood of negative advertising among $1st - 2nd$ and $2nd - 3rd$ pairs increases, respectively, more than 30 and 10 percentage points relative to an average of only 4.8% and 1.5% in single-ballot municipalities. However, results on the $1st - 3rd$ pair are not as clear. Despite estimated local average treatment effect being negative in almost all specifications, it is statistically significant at the 10% level only in some of them.

Figure 2.4.2 provides the graphical version of the 1st order polynomial estimations. Each point in the scatterplots is the average of our negativity dummy for municipalities in a $8,000$ voters interval. There is a clear positive discontinuity around the threshold in all graphs of panels A and C of figure 2.4.2, with statistical significance at the 10% level on almost all bandwidths choices, just as showed in table 2.4.4. Despite confidence intervals being considerably wide given the reduced number of observations around the threshold, the effect of runoff on negativity is very large and indeed suggest an statistically significant increase of more than 30 and 10 percentage points in the likelihood of the $1st - 2nd$ and $2nd - 3rd$ pairs, respectively, being involved in a right of reply lawsuit.

Results on γ_0 are consistent with proposition 5 and the predictions that the second and third best placed candidates are more likely to attack each other in a dual-ballot system. Additionally, we show evidence of a higher probability of the $1st - 2nd$ pair engaging in negative advertising. More than candidates 2 and 3 concentrating their attacking efforts on each other as proposition 5 may imply, these results suggest all candidates concentrate their attacking efforts on the nearest opponent in an attempt to obtain a place in the run-off election.

McCrary test and robustness checks. In order to advocate in favor of a *causal* relation between electoral rules and negativity in the RDD approach, we need to guarantee that there is no manipulation on the running variable around the threshold (MCCRARY, 2008). Moreover, in our case, that the only discontinuity around the threshold is observed in negativity, i.e., covariates should not "jump"when

[2.4.4]Namely, a dummy for the year of 2016. We control for state fixed effects in all specifications.

Tabela 2.4.4 – Negative advertising and electoral rules - Individual pairs

Specification:	0th order polynomial				1st order polynomial			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bandwidth:	50,000	75,000	100,000	125,000	50,000	75,000	100,000	125,000
Panel A: Estimations without covariates								
Subsample: 1st-2nd	0.346***	0.350***	0.289***	0.291***	0.527**	0.530***	0.434***	0.382***
	(0.128)	(0.095)	(0.076)	(0.067)	(0.239)	(0.194)	(0.145)	(0.127)
	66	106	162	244	66	106	162	244
Subsample: 1st-3rd	−0.207**	−0.137*	−0.038	0.017	−0.130	−0.260*	−0.240**	−0.188*
	(0.103)	(0.077)	(0.064)	(0.054)	(0.201)	(0.153)	(0.121)	(0.101)
	66	106	162	244	66	106	162	244
Subsample: 2nd-3rd	0.122	0.149**	0.056	0.089*	0.253	0.189	0.256**	0.076
	(0.092)	(0.071)	(0.057)	(0.049)	(0.180)	(0.147)	(0.107)	(0.093)
	66	106	162	244	66	106	162	244
Panel B: Estimations with covariates								
Subsample: 1st-2nd	0.426***	0.389***	0.345***	0.311***	0.664***	0.613***	0.464***	0.417***
	(0.128)	(0.093)	(0.074)	(0.067)	(0.227)	(0.187)	(0.139)	(0.127)
	66	106	162	244	66	106	162	244
Subsample: 1st-3rd	−0.142	−0.105	−0.006	0.035	−0.038	−0.212	−0.221*	−0.159
	(0.102)	(0.075)	(0.064)	(0.054)	(0.197)	(0.152)	(0.118)	(0.101)
	66	106	162	244	66	106	162	244
Subsample: 2nd-3rd	0.199**	0.194***	0.106*	0.111**	0.373**	0.274**	0.283***	0.113
	(0.088)	(0.066)	(0.055)	(0.048)	(0.166)	(0.136)	(0.101)	(0.091)
	66	106	162	244	66	106	162	244
Single-ballot averages								
Subsample: 1st-2nd	.048							
Subsample: 1st-3rd	.028							
Subsample: 2nd-3rd	.015							

Notes: The dependent variable is our negativity dummy which equals one if there is a litigation involving the candidate pair. Each table cell consists of the estimated $\hat{\gamma_0}$, its standard error in parenthesis and the size of the sample used in estimation. Single-ballot mean column shows the probability of the dependent variable being equal to one in each subsample for municipality from 0 to $200,000$ registered voters. Significance at the 10% level is represented by $*$, at the 5% level by $**$, and at the 1% level by $***$.

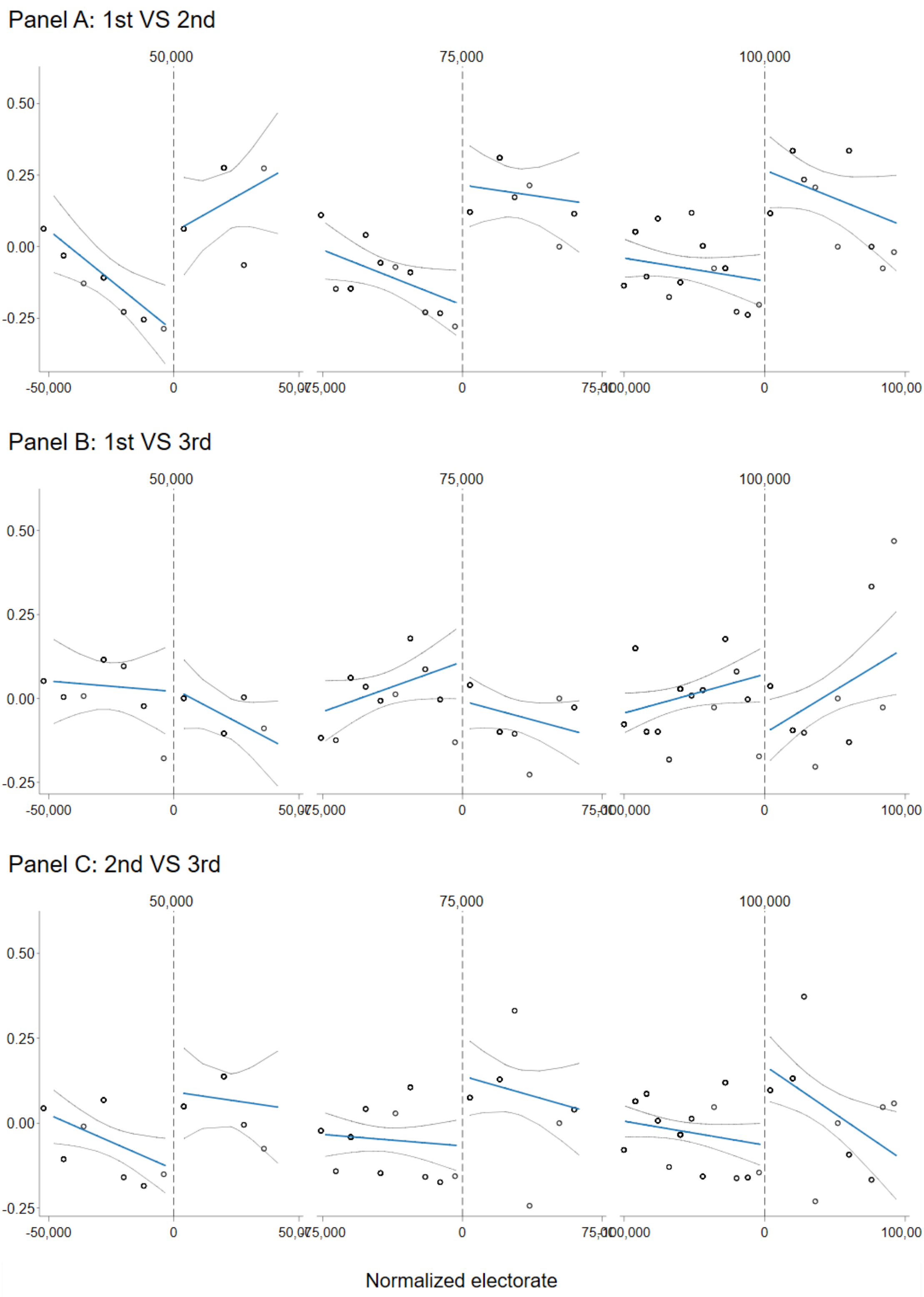


Figura 2.4.2 – Effect of runoff system on negativity between individual pairs on 1st-3rd pair.

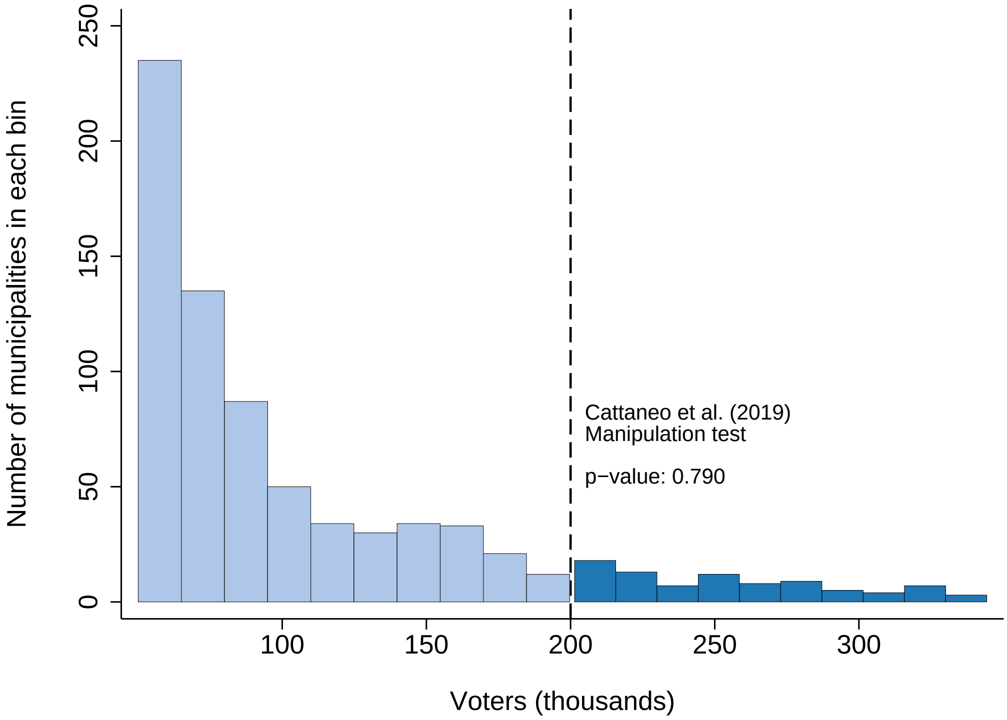


Figura 2.4.3 – Distribution of electorate size.

comparing municipalities at the left to municipalities at the right of the 200, 000 voters threshold.

Figure 2.4.3 suggests that in our database, similar to Fujiwara (2011), there is no sign of strategic manipulation ocurring around the threshold as measured by the Cattaneo, M. e Ma (2019) density test. Furthermore, tables 2.4.5 and 2.4.6 show the results of balance tests of time-invariant and pre-treatment municipality characteristics[2.4.5]: none of the estimations of equation 2.4.4 with these characteristics as outcome variables display a sistematically statistical significant discontinuity around the threshold. These evidence support the validity of our econometric methodology and show that the RDD approach is suitable in this environment.

Even so, could our findings be just the result of chance? As robustness checks,

[2.4.5]Time-invariant characteristics are geographic location (latitude and longitude), long-run year temperature average, and long-run october temperature average. Pre-treatment characteristics refer to income, inequality, poverty, health, employment, and house facilities.

Tabela 2.4.5 – Time-invariant municipality characteristics

Specification:	0th order polynomial				1st order polynomial			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bandwidth:	50,000	75,000	100,000	125,000	50,000	75,000	100,000	125,000
Latitude	0.173	0.074	0.183	0.114	−0.368	−0.071	−0.010	0.245
	(0.174)	(0.172)	(0.154)	(0.147)	(0.358)	(0.341)	(0.290)	(0.277)
	117	189	268	430	117	189	268	430
Longitude	−0.125	−0.061	−0.036	0.067	−0.304	−0.126	0.061	−0.144
	(0.323)	(0.244)	(0.207)	(0.197)	(0.676)	(0.487)	(0.384)	(0.374)
	117	189	268	430	117	189	268	430
Temperature avg. (12 months)	−0.216	−0.256	−0.039	−0.042	−0.195	−0.436	−0.485	−0.270
	(0.217)	(0.186)	(0.165)	(0.169)	(0.445)	(0.369)	(0.313)	(0.323)
	98	145	209	326	98	145	209	326
Temperature avg. (october)	−0.099	−0.221	0.004	−0.023	−0.074	−0.289	−0.412	−0.149
	(0.267)	(0.232)	(0.195)	(0.201)	(0.546)	(0.459)	(0.370)	(0.384)
	98	145	209	326	98	145	209	326

Tabela 2.4.6 – Pre-treatment municipality characteristics (2010 Census)

Specification:	0th order polynomial				1st order polynomial			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bandwidth:	50,000	75,000	100,000	125,000	50,000	75,000	100,000	125,000
Income per capita (th BRL)	15.332	35.073	57.216*	79.899***	55.420	10.982	−22.030	−20.760
	(31.732)	(30.599)	(29.401)	(25.175)	(64.142)	(60.214)	(55.236)	(47.418)
	117	189	268	430	117	189	268	430
Gini index	−0.125	−0.061	−0.036	0.067	−0.304	−0.126	0.061	−0.144
	(0.323)	(0.244)	(0.207)	(0.197)	(0.676)	(0.487)	(0.384)	(0.374)
	117	189	268	430	117	189	268	430
Extreme poverty (% population)	−0.216	−0.256	−0.039	−0.042	−0.195	−0.436	−0.485	−0.270
	(0.217)	(0.186)	(0.165)	(0.169)	(0.445)	(0.369)	(0.313)	(0.323)
	98	145	209	326	98	145	209	326
Life expectancy	−0.099	−0.221	0.004	−0.023	−0.074	−0.289	−0.412	−0.149
	(0.267)	(0.232)	(0.195)	(0.201)	(0.546)	(0.459)	(0.370)	(0.384)
	98	145	209	326	98	145	209	326

we present (i) the distribution of $2,000$ γ_0 estimations at false population thresholds, and (ii) the treatment effect sensitivity to bandwidth choice. In the first case we should expect little evidence of impact on negativity, or at least little evidence of an impact as big as the real one, if we believe in a *causal* relation between negativity and electoral institutions. In order to assess this question, we calculate the kernel density of the estimated coefficients at false population thresholds, ranging from $100,000$ to $300,000$ at a pace of 100, i.e., *cutoff* = 100000,100100,...,299800,299900,300000. Figures 2.4.4 to 2.4.6 reports the results of this exercise for different bandwidth choices. It is clear that the estimated coefficient over the true $200,000$ threshold

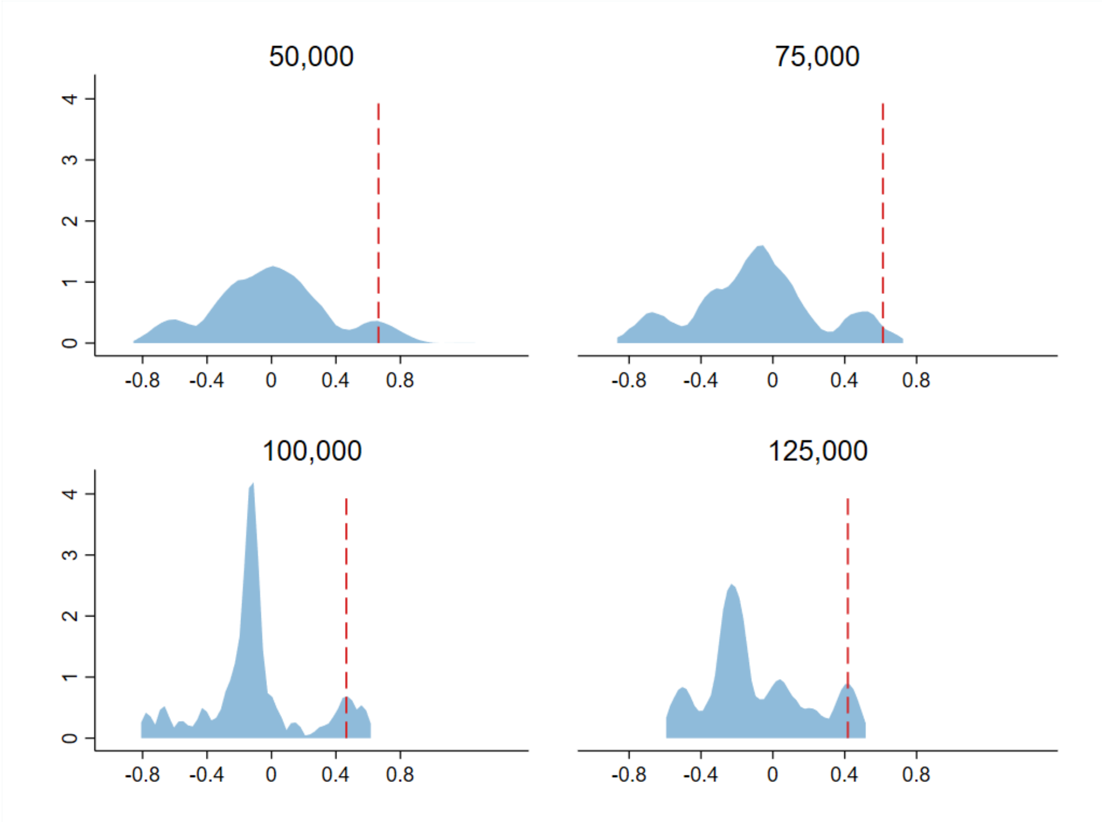


Figura 2.4.4 – Kernel density of ATE estimations for placebo thresholds on 1st-2nd pair.

Notes: The figure depicts the kernel density of all $2,000$ coefficients estimated on placebo cutoffs and the coefficient estimated on the true $200,000$ threshold (red line) for various bandwidth choices.

is bigger than most of placebo estimations for all bandwidths choices in figures 2.4.4, and 2.4.6. However, as the estimations and graphical representations showed there isn't evidence of a robust effect of runoff system on the likelihood of the pair $1st - 3rd$ engaging in negative advertising (figure 2.4.5).

For the second exercise we should expect a higher estimator bias and lower variance as the bandwidth increases, which is result of a bigger sample but with more observations not so close to the threshold. Figures 2.4.7, 2.4.8, and 2.4.9 reports the treatment effect and the 90% confidence interval (CI) over each estimation for some bandwidth choices. As expected, CI is narrower as we choose a larger

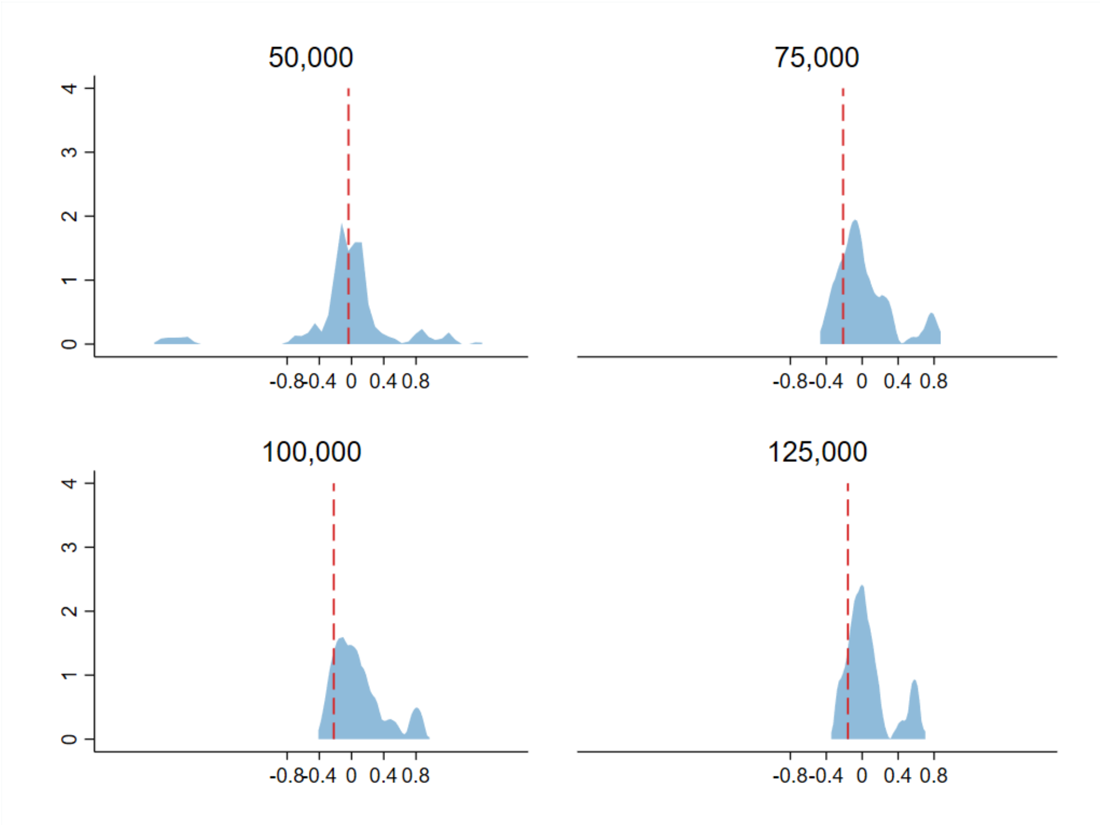


Figura 2.4.5 – Kernel density of ATE estimations for placebo thresholds on 1st-3rd pair.

Notes: The figure depicts the kernel density of all $2,000$ coefficients estimated on placebo cutoffs and the coefficient estimated on the true $200,000$ threshold (red line) for various bandwidth choices.

bandwidth. Moreover, estimated coefficient seems to converge monotonically for a positive value in both specifications of figures 2.4.7 and 2.4.9. This suggest that the variability in estimations for smaller bandwidths may be result of bias being added as we move farther from the threshold and not result of lack of robustness of the *causal* relation. One more time, there isn't evidence of a consistenly significant effect of runoff on the $1st - 3rd$ pair.

Compound treatment. Another threat to the results validity could be a *compound treatment*, in which the same running variable threshold defines more than one policy change. Eggers et al. (2015) show some recent studies that use brazilian

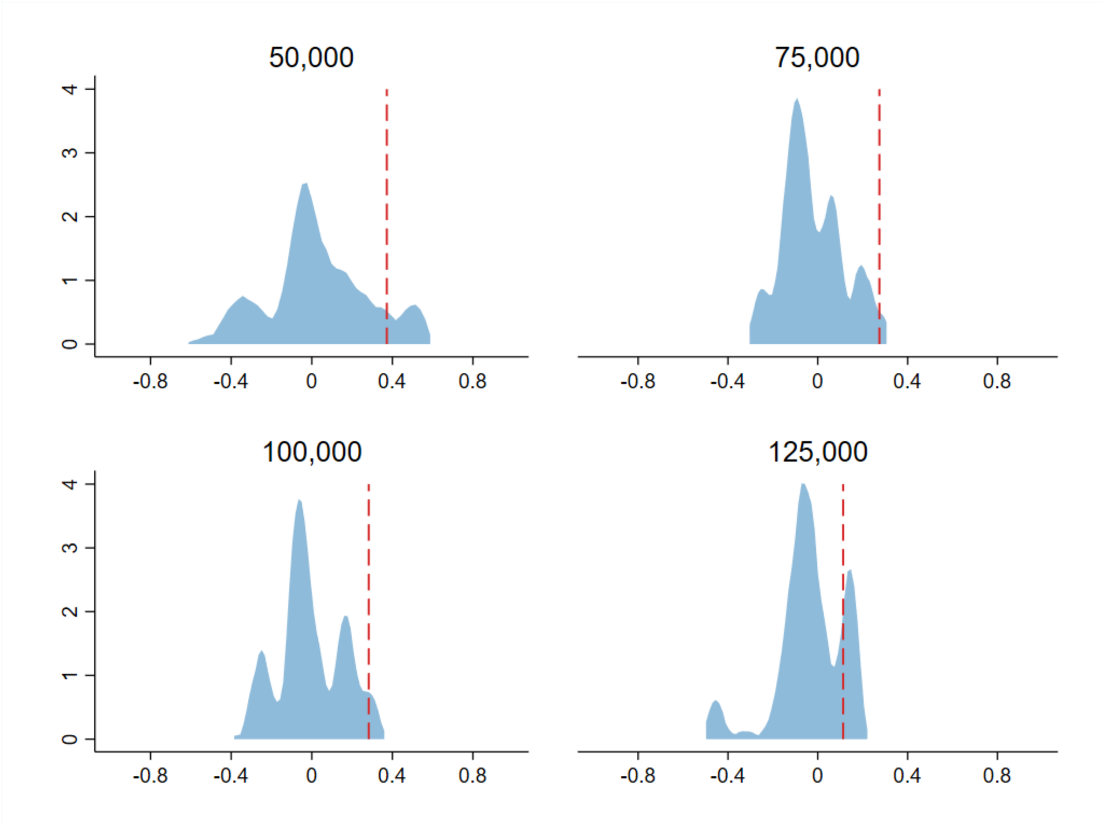


Figura 2.4.6 – Kernel density of ATE estimations for placebo thresholds on 2nd-3rd pair.

Notes: The figure depicts the kernel density of all 2,000 coefficients estimated on placebo cutoffs and the coefficient estimated on the true 200,000 threshold (red line) for various bandwidth choices.

data and population thresholds as the main source of econometric identification. With what could be the closest threshold to ours, Ferraz e Finan (2009) use municipal population thresholds (10,000, 50,000, 100,000, 300,000, and 500,000), introduced by a constitutional amendment in 2000, to assess how monetary incentives impact quality and performance of politicians. This amendment introduced a cap on the maximum salary that local legislators could receive. However, the relation between legislators' salary and campaign of mayors candidates are far from clear. Furthermore, this amendment imposed a population threshold and not a voters threshold, with the latter being the one that we use. Therefore, just as in Fujiwara (2011), our findings are result of only one institutional change: from an

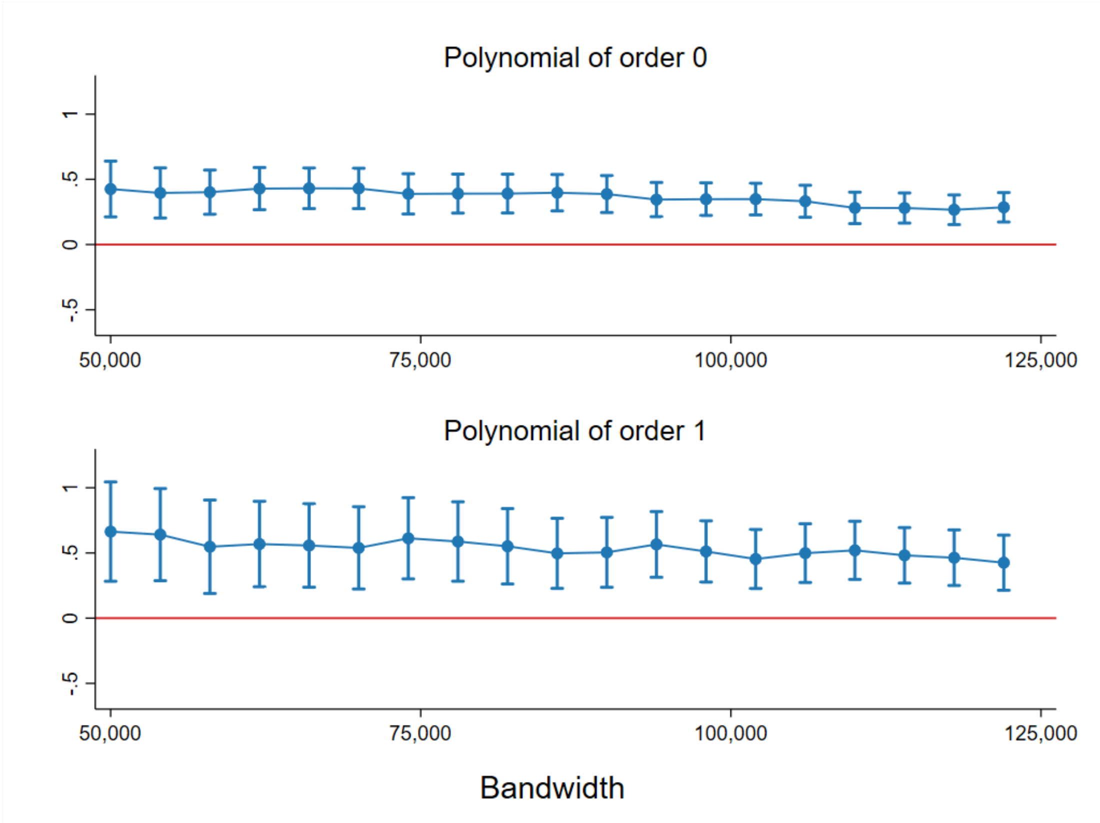


Figura 2.4.7 – Coefficient estimates and 90% CI for the $1st - 2nd$ pair subsample.

electoral single-ballot system to a runoff system.

2.5 Concluding remarks

Dirty campaigning has long been recognized as a threat to democracy and voters participation (ANSOLABEHERE et al., 1994). However, the potential negative impact has increased significantly in recent years with the advent of social media and the fake news phenomen. Understanding the incentives that shape candidates decision on this matter is now more important that it has ever been.

This paper studies both the theory and evidence behind the candidates decision to engage in a more negative campaign. We developed a model of decision on dirty campaigning that fully characterize the equilibrium behavior of candidates

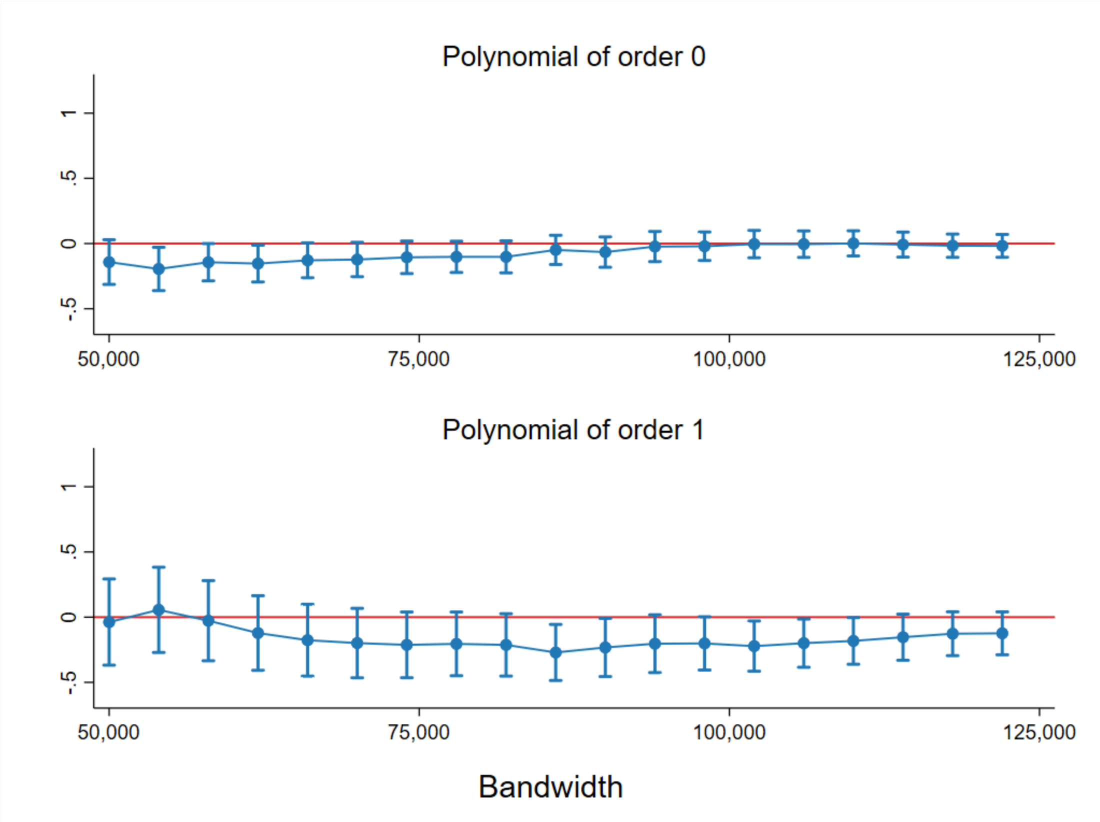


Figura 2.4.8 – Coefficient estimates and 90% CI for the $1st - 3rd$ pair subsample.

in the case of two and three-candidates competition and in both single-ballot and dual-ballot systems. The model predicts that (i) a candidate is more likely to attack an opponent as the opponent's innitial support increases, (ii) under single-ballot system, candidate 1 and 2 are more likely to attack each other while candidate 3 is more likely to attack candidate 1, (iii) the two best placed candidates are more aggresive towards each other when there isn't any other candidate disputing the seat, and (iv) that the second and third best placed candidates are more likely to attack each other when we go from a single-ballot to a dual-ballot plurality system. We show that the characteristics of the race, such as competition and electoral institutions, deeply affect the incentives candidates have to engage in negative advertising regardless candidates preferences and political background.

Using an unique database containing judicial information about Brazil's mayors elections of 2012 and 2016, we tested the model predictions on how the

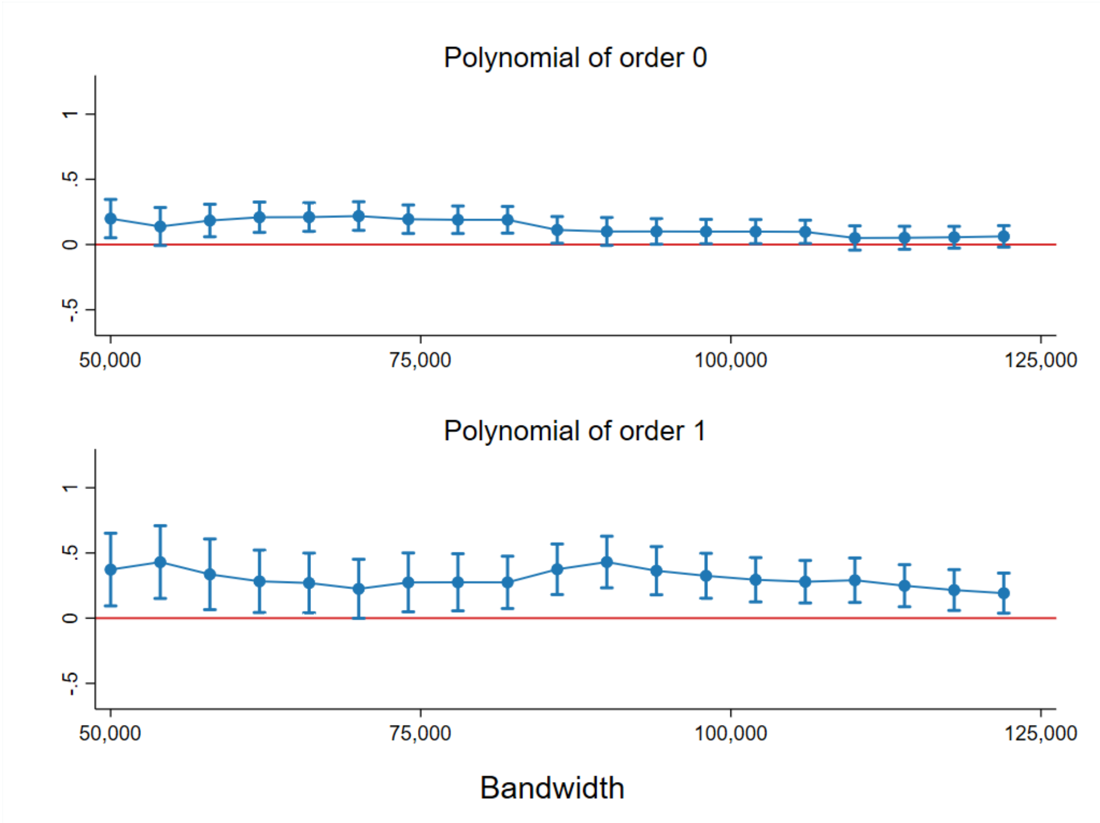


Figura 2.4.9 – Coefficient estimates and 90% CI for the $2nd - 3rd$ pair subsample.

characteristics of the race affect the probability of candidates being part in a “right to reply"lawsuit, a new and objective proxy for campaign tone. Our fixed-effects regressions show that dirty campaigning is more likely to occur in more competitive races and against stronger opponents. Moreover, we show that in single-ballot municipalities, candidate 1 and 2 are more likely to attack each other while candidate 3 is more likely to attack candidate 1. Estimated coefficients are stable across specifications and magnitude is substantial in most cases. Regression discontinuity design estimations show that dirty campaigning between the two best placed candidates ($1st - 2nd$ pair) and between the second and third best placed candidates ($2nd - 3rd$ pair) increases significantly when one moves from a single-ballot to a runoff plurality system. RDD findings are robust to bandwidth choices and other falsification tests. In both fixed-effect and RDD estimations, we present empirical evidence in favor of all model predictions.

By evaluating how the characteristics of the dispute relate to candidates' decisions on their campaign's tone, this paper helps to bring some light on mechanisms through which dirty campaigning shows up. We present a formal representation of the process and document empirical aspects of negative advertising that weren't available until now. The paper contributes to the political economy literature in several ways and may be helpful to policymakers who aim to promote a more policy-oriented political campaign.

3 LABOR COURTS, JOB SEARCH AND EMPLOYMENT: EVIDENCE FROM A LABOR REFORM IN BRAZIL

Com Raphael Corbi, Rafael Ferreira e Renata Narita

3.1 Introduction

For the large majority of households in the world, the labor market is the single most important market. That is especially true for those in developing economies, where the majority of the population derive all of their income from the work they sell. Given the importance of these markets, the enactment of new employment protection regulations (or the reform of old ones) is often preceded by heated debate.

Proponents of stronger regulations often point to inequities in the distribution of surplus between employer and employee and argue that market frictions inevitably reduce worker welfare. In contrast, the opposing side usually cautions that poorly designed labor market regulations can also be detrimental to worker welfare and points to the heterogeneous effects these regulations can have on the well-being of different types of workers, with regular employees usually benefiting at the expense of temporary, unemployed or informal workers.

The extent to which each side is correct is mostly an empirical matter. In this paper, we analyse decision patterns of Brazilian labor judges and document their tendency to decide in favor of workers. Then, we study how judges that are relatively more biased towards workers affect employment, wages and firm financial distress and survival rates. Finally, we interpret our results within a search and matching framework in which workers decide whether to take firms to court, so that we can conduct counterfactual exercises emulating the changes brought by a large labor reform that took place in 2017, in Brazil.

We started by building an original data set of over two million first instance judicial decisions from judges of the largest set of labor courts in Brazil, serving the city of São Paulo and its neighboring municipalities. We did it by extracting information of each labor case (including its outcome) from court publications. Using these decisions, we compute a residualized leave-one-out measure of relative pro-worker judicial bias, to capture the difference between the judge's decision and the average decision pattern in her jurisdiction. This approach is the same as the one in Dahl, Kostol e Mogstad (2014), Dobbie, Goldin e Yang (2018), Bhuller et al. (2019) and Cahuc, Carcillo e Patault (2019). Next, we linked this data to other rich administrative data sets with matched employer-employee records (RAIS) and information on bankruptcy filings. By exploiting the random assignment of cases to judges within the same jurisdiction, we are able to estimate the causal effect of higher labor costs on several labor market and firm performance outcomes.

Our results indicate negative and significant effects of labor regulation costs on employment and wages of new hires. Specifically, the growth rate of employment decreases by 2.1% and the growth rate of the average wage of new hires by 0.8% as we increase the judge pro-worker bias in one standard deviation. We also find evidence that these costs increase the likelihood that firms experience financial distress or go out of business.

Finally, we use the data on labor judges decisions to calibrate an equilibrium search and matching model for the labor market, extended to allow workers take the former employer to court after a match is destroyed and before a new match is formed. We then perform counterfactual exercises by partially shifting legal fees payments to employees, in line with a large Labor Reform that came into force in 2017. Initial results indicate an increase in employment and a decrease in the relative number of lawsuits among the employed, in equilibrium.

The article is related to several works in a at least two branches of the literature. The first of these branches is the literature on employment protection legislation and its impact on economic outcomes. There is a well documented cross-country correlation between restrictive labor market regulation and economic efficiency (BOTERO et al., 2004). However, cross-country studies usually present identification challenges. Other works have exploited some regional variation in the timing of the adoption of labor regulation within the same country to ascertain its impacts. Autor, III e Schwab (2006), Autor, Kerr e Kugler (2007) take advantage of variations in the timing and the extent of the adoption of new wrongful-discharge laws across U.S. states to assess their impact on wages, employment and productivity. Sapkal (2016) exploits differences in the intensity of labor inspections and regional variations in labor regulations in India on firm hiring practices. Almeida e Carneiro (2012), Almeida e Carneiro (2009) also look at the intensity of labor inspection to study its effect on informality and firm size in Brazil. Our paper complements this literature by focusing on labor courts, another element of enforcement of labor regulations, and one with particular importance in Brazil and in other developing and European countries. Bamieh (2017) and Cahuc, Carcillo e Patault (2019) are two recent works on the same topic, focusing on European countries (Italy and France, respectively).

The second branch is the now extensive literature on how parties in a lawsuit are affected by judge characteristics. For example, there is evidence that ethnicity bias (GAZAL-AYAL; SULITZEANU-KENAN, 2010; SHAYO; ZUSSMAN, 2011; DEPEW; EREN; MOCAN, 2017; ARNOLD; DOBBIE; YANG, 2018), gender bias (KNEPPER, 2018) and the judge's experience on the bench (IVERSON et al., 2018) all influence judicial decisions. Documenting how these types of variables influence outcomes is interesting in itself. Ideally, justice should be served in consistent and systematic manners to all parties. The judge's characteristics, being unrelated to the merits of each case, should play no part in its outcome. Therefore, measuring the variation in judicial decision patterns in cases, on average, similar to one another gives a good sense of how large judicial uncertainty is in a given legal system. In our setting, we find considerable variation in the decision patterns of Brazilian labor judges. Compared to one of the 10% least pro-worker judges, being assigned to one of the 10% most pro-worker judges increases the expected amount granted by the judge from 5.5 to 13.1 monthly wages. It also increases the likelihood that the complaint will be accepted by 21 percentage points, from 61% to 82%. As a comparison, for french labor appellate courts, Cahuc, Carcillo e Patault (2019) finds a smaller variation, with only a five percentage point difference in this likelihood between the median-biased judge and one of the 10% most pro-worker judges.

Additionally, many works in this literature are also interested in outcomes outside of the judicial process. The usual approach in these cases is instrumenting the judicial decision in a particular case with the average judicial decision made by the same judge in prior cases. This method is particularly suited to research questions like ours. Employment protection institutions tend to apply to an entire country or region. Because we only observe these environments either with or without any of such institutions, we cannot compare outcome values under both circumstances. The differences in characteristics of judges randomly assigned to each case provide the exogenous variation we need to properly estimate the effects of increased labor costs. There are several other works using the same methods to study, for example, the impact of pre-trial detentions on criminal and job market outcomes (DOBBIE; GOLDIN; YANG, 2018), the effect of different bankruptcy procedures on asset allocation (BERNSTEIN; COLONNELLI; IVERSON, 2019), and the effect of patents on inovation (GALASSO; SCHANKERMAN, 2015).

We proceed as follows. The next section presents the institutional setting of our empirical exercise, outlining how labor justice is administered in Brazil and describing a recent labor reform that significantly change labor legal environment. The next section also presents the details of our dataset and some descriptive statistics. Section 3 describes the empirical strategy, how the instrument was built, and the reduced-form evidence on the effect of judge pro-worker bias on firm survival and firm outcomes. In sections 4 we present the model, its main objects, and discusses the early counterfactual results and the impact of the 2017 Labor Reform. We conclude in section 5.

3.2 Institutional Framework and Data

When an employee decides to sue a current or former employer in Brazil, her case will be heard in an specialized branch of the Federal Judiciary called *Justiça do Trabalho*. This branch is divided into 24 regional jurisdictions, called regions (or *regiões*), that in most cases coincide with the geographic territory of Brazilian states.[3.2.1] Each of these regions has its own set of first instance courts (grouped together into courthouses) and its own appellate court. Together these courts form a *Tribunal Regional do Trabalho (TRT).*

Labor courthouses usually handle cases coming from one or more municipalities, so that a case originated in a certain municipality might have to be filed in a courthouse located in a neighboring town. Once filed, the case is then randomly assigned to one the first instance labor courts (or *Varas do Trabalho*) in the courthouse. Each of these courts consists of a regular judge (*juiz titular*) and a substitute judge. The position of substitute judge is an entry-level[3.2.2] position for recently admitted judges, who are responsible for replacing regular judges when they are on leave.[3.2.3] With time, as regular judge positions become vacant, a substitute judge with long enough tenure and good performance reviews tends to be promoted to regular judge and assigned to one *Vara do Trabalho.* As her career progresses, she might be elevated to a position of appellate judge, reviewing appeals from cases coming from lower level courts.

[3.2.1]Most states have their own region of the *Justiça do Trabalho*, the exceptions being the states of Acre and Rondônia (14th Region), Amazonas and Roraima (11th Region), Tocantins and the Federal District (10th Region) and Pará and Amapá (8th Region). All of these regions contain at least one state that is too small to merit its own jurisdiction. The state of São Paulo is the only one with two regions: the 2nd Region, for the state capital and its neighboring municipalities; and the 15th region, for the state's remaining municipalities.

[3.2.2]Judges are admitted into the profession through a competitive civil service selection process that includes written and oral exams and the evaluation of academic and professional credentials. Labor judges are paid wages that put them above the 99th percentile in the distribution of labor income in Brazil.

[3.2.3]A substitute judge might be appointed to one or more temporary posts in a set of contiguous jurisdictions called judicial circumscription or sub-region. Therefore, while a more senior judge is most likely in charge of only one *Vara do Trabalho* (in one jurisdiction), a substitute judge often rotates between different courts, in different jurisdictions within the same sub-region. Positions of substitute labor judges are filled according to the ranking in the entrance exam, with each judge being able to choose an unfilled position that has not been previously chosen by a higher-ranking individual.

Tabela 3.2.1 – First instance courts in jurisdictions with two or more courts in the 2^{nd} Region

Sub-region	**Courthouse**	**Jurisdiction**	**Courts**
Capital	Capital (Central)	Downtown and the North and West parts of São Paulo	90
	Capital (Zona Sul)	South part of São Paulo	20
	Capital (Zona Leste)	East part of São Paulo	14
Guarulhos	Guarulhos	Guarulhos	13
	Mogi das Cruzes	Mogi das Cruzes, Biritiba-Mirim, Guararema and Salesópolis	4
	Suzano	Suzano	2
ABC	São Bernardo do Campo	São Bernardo do Campo	8
	Santo André	Santo André	5
	Diadema	Diadema	4
	Mauá	Mauá	3
	São Caetano do Sul	São Caetano do Sul	3
Santos	Santos	Santos	7
	Cubatão	Cubatão	3
	Guarujá	Guarujá and Bertioga	3
	Praia Grande	Praia Grande	2
	São Vicente	São Vicente	2
Osasco	Osasco	Osasco	6
	Barueri	Barueri	5
	Carapicuíba	Carapicuíba	2
	Cotia	Cotia	2
	Franco da Rocha	Franco da Rocha, Francisco Morato and Mairiporã	2
	Itapecerica da Serra	Itapecerica da Serra, Embu Guaçu, Juquitiba and São Lourenço da Serra	2
	Itaquaquecetuba	Itaquaquecetuba	2
	Santana de Parnaíba	Santana de Parnaíba and Pirapora do Bom Jesus	2
	Taboão da Serra	Taboão da Serra	2

Courthouses with jurisdiction over the municipalities of Arujá, Caieiras, Cajamar, Embu das Artes, Ferraz de Vasconcelos, Itapevi, Jandira, Poá, Ribeirão Pires, Rio Grande da Serra and Santa Isabel have only one court.

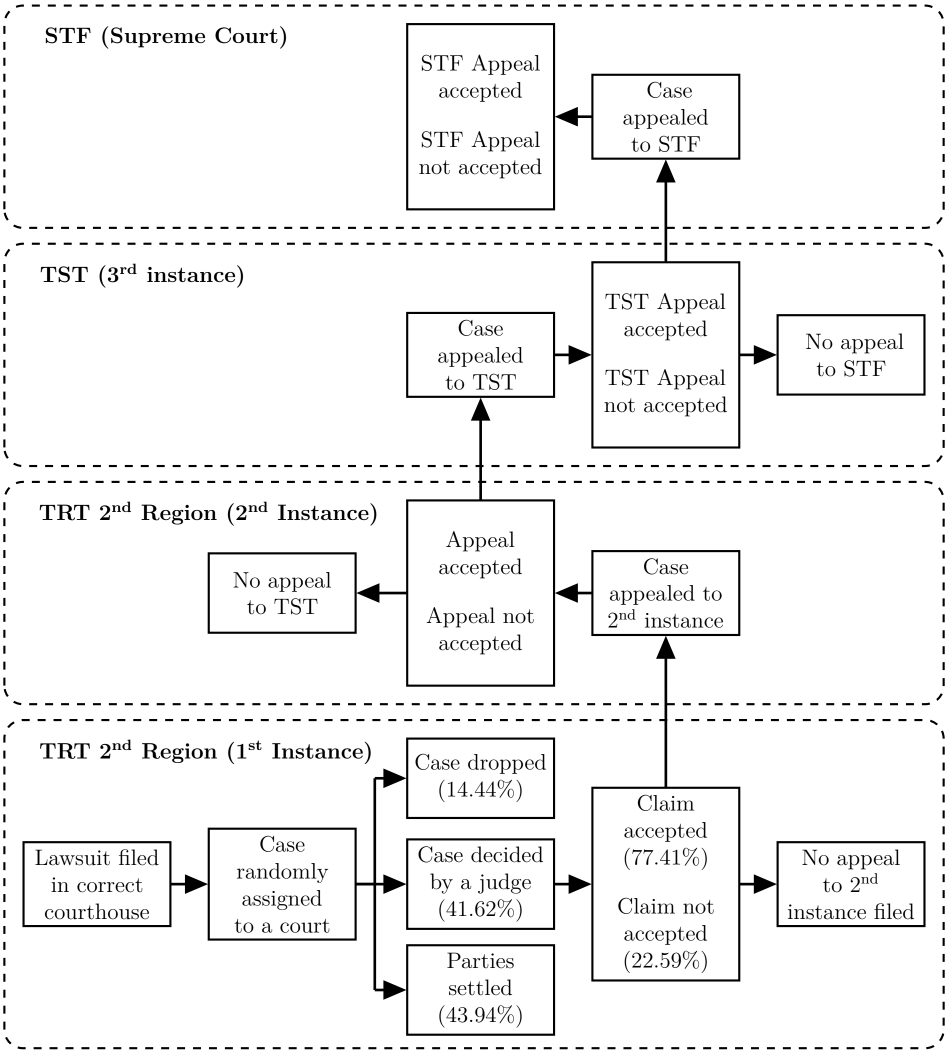
In this paper, we focus on decisions made by first instance judges from the TRT for the 2nd Region (henceforth TRT2), the largest one in the country, with nearly $500,000$ new cases filed per year. It has jurisdiction over the capital of São Paulo state and 45 of its neighboring municipalities, with a total of 217 labor courts. Table 3.2.1 shows how these courts are grouped into jurisdictions, courthouses and sub-regions. This is relevant mostly because our identification strategy takes advantage of the random assignment of cases to courts, which takes place at the courthouse level.

The court then schedules an initial hearing, in which the judge inquires if a deal can be reached between the plaintiff worker and the defendant firm. If an agreement is reached and approved by the judge, the firm commits to the agreed upon payment schedule and the suit is terminated. Otherwise, the judge may choose to continue with the hearing, calling parties and witnesses to testify and lawyers to present their cases. She may also choose to do so in a later date, by scheduling a new hearing if key witnesses are absent, important evidence needs to be produced or the appraisal of an expert is required. After all witnesses and parties have been heard, the judge then once again asks if an agreement can be reached. If not, she may either make her ruling immediately or add the case to the queue of cases waiting for a decision in a later date.

When the judge finally comes to a conclusion, her ruling can be of three types: *procedente*, when all of the worker's requests are granted; *improcedente*, when none of the requests are granted; or *parcialmente procedente*, when the plaintiff's demands are only partially met. After that, any party that is dissatisfied with the judge's ruling may choose to take the case to the next higher level court. Appeals can take the case to even higher courts, such as the highest appellate courts for labor matters, the *Tribunal Superior do Trabalho (TST)* or even the *Supremo Tribunal Federal (STF)*, Brazil's Supreme Court.[3.2.4] A judicial ruling is only enforced when it is final and can no longer be appealed.

[3.2.4] Decisions from labor judges in Brazil are often appealed. First instance TRT2 judges are called to review 16.6% of their own decisions and have 71% of their decisions reviewed by second instance appellate judges. In turn, TRT2 appellate judges review 23.6% of their own cases and have 39% of their cases moved up to the TST.

Figura 3.2.1 – Stages of a typical labor lawsuit in Brazilian courts.



Percentages indicate conditional probabilities, so that percentages from nodes with the same parent add to 100%.

3.2.1 Labor Laws and Labor Judges

Consistent with its Civil Law[3.2.5] origin, Brazilian Labor Law extensively regulates employment relationships. The core of Brazilian Labor Law consists of a decree from 1943 called Consolidation of Labor Laws, or *Consolidação das Leis do Trabalho (CLT)*, that has more than 900 provisions and unified several previous labor laws that had been enacted in the 1930's.[3.2.6]

Another important legislation on this matter is the country's Constitution, adopted in 1988. Brazil's lawmakers decided to add to the Constitution several articles on workers rights and benefits, such as overtime pay, length of work days and work weeks, Christmas and vacation bonuses, maternity leave and minimum wages. The added benefits constituted a major and hard to reverse[3.2.7] increase in labor costs.

Together the 1988 Constitution and the CLT created a legal framework with strict rules for what a employment relationship should look like. They detail which rights workers have and which employment practices are lawful. Up until recently, for example, intermittent work schedules or freelance work contracts were outlawed and companies that had workers in any of these types of employment often faced high risk of litigation. If sued, they would be most likely ordered to compensate the worker for the entire duration of the contract, as if she was a full time regular employee. Any mutually agreed and more flexible arrangement between employer and employee had little hope of prevailing over labor laws.

In addition to being highly restrictive, Brazilian Labor Law is also notoriously and overtly pro-worker[3.2.8] in the way legal proceedings are done, differing

[3.2.5]For a review of the differences between Civil Law and Common Law legal origins and their effect on labor regulation, see Botero et al. (2004).

[3.2.6]See Gonzaga, Maloney e Mizala (2003)

[3.2.7]As in most countries, constitutional changes in Brazil are difficult to achieve. Constitutional amendments have to be approved in two rounds of vote in both the lower and the upper houses of Congress, by three-fifths of representatives and senators.

[3.2.8]One of the tenets of Brazilian Labor Law is the "Principle of Protection", which indicates that Labor Law should correct or soften the inherent imbalances in labor contracts. This guideline is often interpreted as a suggestion for judges to decide, whenever in doubt, in favor of the worker; and to always enforce the company policy or legal article that is most favorable to the worker, whenever they believe that more than one rule may apply. See Delgado (2012, p. 193).

significantly from other fields of Law. Historically, Labor Procedural Law, which refers to the rules by which the labor courts hear the cases, has provided several incentives for workers to litigate, such as low requirements for filing a suit, generous exemptions from payment of court fees and rare punishment for frivolous lawsuits.

For example, up until recently, workers could have their lawyers file a generic complaint with a long list of demands and ask for an arbitrary amount of money in compensation from a former employer. Even if the complaint contained unreasonable and groundless demands, plaintiffs were unlikely to be punished. Their lawyers could wait until the defendant had presented its case and then, conditional on having seen the opposing side's defense, decide on whether or not to drop the case, with no consequence to the lawyers or their client. Lawyers could also ask for their clients to be exempt from paying court fees, expert witnesses' fees or attorney's fees and the judge would most likely grant the request, so workers often faced no direct costs from litigation. This environment provided great incentives for workers and their lawyers to bring lawsuits against former employers. For each demand added to the initial complaint, the worker could end up having her claim accepted by the judge, in which case she would receive a compensation from her former employer; or could have the request denied by the judge, without incurring in any costs.

Another source of litigation risk comes from the judge assigned to the case. From the start, the law grants judges a lot of freedom to determine the pace and the specifics of the legal process.[3.2.9] For example, they decide whether or not to dismiss the suit, after the initial complaint;[3.2.10] determine, in many cases, how many hearings are necessary;[3.2.11] decide which witnesses and experts need to be heard; determine which party, if any, should pay for court and court-related fees;[3.2.12] and decide whether or not additional physical evidence[3.2.13] or documents should be presented to court. A labor judge can also nullify settlements or apply

[3.2.9] According to article 765 of CLT, the labor judge "has ample freedom in the guidance of the case and should ensure its rapid progress". It also says judges can "order any judicial proceeding necessary to the clarification of cases."

[3.2.10] Art. 840, §1 of the CLT.

[3.2.11] Art. 813, §2 of the CLT.

[3.2.12] Art. 790, §4 of the CLT.

[3.2.13] Art. 852-D of the CLT.

pressure to parties unwilling to accept or negotiate a settlement when she believes they should.

In fact, this broad procedural freedom granted to judges goes even further. In labor lawsuits, oral testimonies from parties and witnesses (usually presenting conflicting narratives) play a major role. During hearings, a labor judge can question parties, witnesses and experts directly[3.2.14] and later base her decision entirely on the testimonials she finds the most convincing, even when formal documents contradict these testimonials.This increases the chances that the judges' implicit and explicit biases play an important role in the outcome of the case. And in fact there is some evidence to that effect. Pinheiro (2003) ran a survey with 741 Brazilian judges, from 11 different states and found that 45.1% of respondent judges considered that, in labor cases, decisions are often or very often based more on the judge's political views than on the "strict reading of the law." Only for cases involving privatisation and regulation of public services did a higher percentage of respondents say the same. The same survey also found that 45.8% of labor judges believe that "the seeking of social justice sometimes justifies decisions that violate contracts."

When considered together, Brazil's biased judges and its rigid and pro-worker labor laws create large incentives for litigation[3.2.15], impose high costs on firms, and place many restrictions on the demand for labor. In fact, the 2009 World Bank Enterprise Survey with business owners and top managers from 1802 Brazilian firms found that 63.2% of firms identify labor regulations as a major constraint, compared to only 16.7% of firms of all of Latin America and the Caribbean, and 11.2% of firms from all of the 143 surveyed countries. Of all of these countries, Brazil was the one with the highest share of firms mentioning labor regulations as a major issue, with Argentina in a distant second with 49%.

[3.2.14] Art. 820 of the CLT.

[3.2.15] According to the *National Justice Council* (CNJ), 26.4% of all new cases filed in 2016 in the Labor, State, and Federal Justice were labor cases.

3.2.2 2017 Labor Reform

In 2017, Brazilian Labor Law went through its most significant changes in decades. The main objectives of these reforms were lowering expected labor costs and increasing employment, by addressing some of the issues mentioned in the previous section. In March 2017, Federal Law 13,429 allowed firms to outsource work that was previously not permitted by law to be outsourced. Until then, firms were only allowed to outsource secondary work, unrelated to the firm's core activities. For example, schools were allowed to outsource their security and cleaning, but not their teachers. After the reform, any work activity can be outsourced, without the firm being subject to legal action in labor courts.

In July 2017, a larger reform altered several articles in the Brazilian Constitution and in the CLT. These changes were expected to reduce labor costs by three channels: i) by increasing the set of employment practices deemed legal; ii) by removing from procedural labor law its most noticeable incentives for excessive and groundless litigation, and to manipulate the system; and iii) by reducing the discretion of judges to decide in accordance to their views when these are in conflict with the law. Table 3.2.2 lists some of the main changes brought by the July 2017 labor reform.

Rows 1 through 10 in Table 3.2.2 describe changes in rights and responsibilities of employers and employees. These are related mostly to increases in the set of lawful labor practices and to the loosening of the rules on how workers can exert their rights. Importantly, these reforms made legal more flexible employment practices that were indispensable for the viability of some economic activities. They also addressed a few of the main sources of labor complaints, most often related to alleged unpaid overtime, disagreements about severance payments, failure to give proper notice of termination of contract and unpaid unused vacation days.[3.2.17]

Rows 11 through 17 in Table 3.2.2 describe changes in procedural labor law. Taken together, these changes affected both expected gains and losses from labor litigation. By making legal previously illegal employment practices, some of the

[3.2.17] See Tribunal Superior do Trabalho (2019).

Tabela 3.2.2 – Some of the changes brought by the July 2017 Labor Reform in Brazil

	Topic	How it was before the reform	How it came to be
1	Paid vacation	For every 12 months of work, firms had to provide 30 consecutive days of paid vacation to employees.	Vacation days can be split into three during the year, as long as each vacation period is not shorter than 5 days and one of these periods is not shorter than 14 days.
2	Termination of contract	When an employee is fired, she is entitled to make a withdraw from her account in a social insurance fund called FGTS.[3.2.16] The firing firm pays the worker a fine of 40% of the balance in the worker's FGTS account. Workers who quit their jobs were not allowed to withdraw these funds nor were paid the fine.	If employer and employee reach an agreement to terminate the contract, the worker can withdraw 80% of her FGTS funds and is entitled to a fine of 20% of the balance in her FGTS account.
3	Lunch break	During work shifts of 6 hours or more, the worker had to take a break of at least one hour for lunch.	Employer and employee may agree to a shorter break of 30 minutes.
4	Freelance or Intermittent work	Both of these types of work were forbidden. Courts often ruled that workers performing freelance or intermittent work were entitled to the same rights as full-time regular employees.	Freelance work is explicitly allowed. Employers may hire workers on demand and by the hour, as long as the worker is given a three-day notice. Firms may also impose fines on absent workers.
5	Part-time work	Allowed, up to 25 hours per week. Overtime was forbidden for these types of workers.	Employers may hire part-time workers for up to 30 hours a week, without overtime; or for no more than 26 weekly hours, with overtime limited to 6 hours per week.
6	Work from home	There was no legal basis for working from home.	Work from home is allowed. Employer must pay for the related costs.
7	Overtime	Employees that worked overtime right after the end of their regular shifts had to rest for 15 minutes before starting the extra work hours.	Overtime can begin right at the end of the regular work shift.
8	Commuting	Firms providing transportation for their workers in hard-to-access locations or in places where there was no public transportation available had to pay them for the commuting time.	Worker starts to get paid only when her work effectively begins.
9	Individual agreements	Individual agreements cannot outweigh collective agreements.	High income workers can make individual agreements with their employers that take precedence over collective agreements negotiated by their union.

Tabela 3.2.2 – Some of the changes brought by the July 2017 Labor Reform in Brazil (cont.)

	Topic	How it was before the reform	How it came to be
10	Unions	Unions had to approve mass layoffs and collective agreements regarding banked hours. Union dues were mandatory and workers had to pay the equivalent of one day of labor income to the union.	Employers do not need the Union's approval for mass layoffs and collective agreements on banked hours can be made directly between employer and employees. Union dues are optional.
11	Frivolous and groundless litigation	There was no legal basis in labor law for imposing penalties for bringing a frivolous lawsuit.	Plaintiffs bringing frivolous and groundless lawsuits can be required to pay a fine (between 1% and 10% of the amount being asked). They can also be required to compensate the defendant for damages and legal expenses.
12	Out-of-court settlements	There was no legal basis for out-of-court agreements. Even when such agreements were later reviewed and approved by a judge, the worker could still bring a case against the firm in the future.	Out-of-court agreements that are reviewed and approved by a judge are legally binding. Any future labor complaints cannot include issues that were settled in such an agreement.
13	Court fees	Parties seeking exemption from payment of court fees did not need to prove they were eligible for the benefit.	Parties seeking exemption from payment of court fees now have to demonstrate that they do not have enough income nor wealth to pay for court fees.
14	Expert's fees	Losing side was exempt from paying the court-appointed expert's fee when they were considered too poor to pay for court fees.	Even if the losing side is considered too poor to pay for court fees, they still have to pay the expert's fees.
15	Attorney's fees	Contrary to what happens in most civil cases, the losing party in labor cases did not had to pay the winning side's attorney's fees.	The losing party now pays for the winning side's attorney's fees, even when the losing party is considered too poor to pay for court fees.
16	Withdrawal of complaint	Plaintiffs could drop the case, even after the opposing side had presented its defense.	After the opposing party has presented its defense, the plaintiff is required to get the consent of the opposing side to drop the case.
17	Complaint requirements	Labor complaints had only to contain a brief description of the reasons for the claim and the plaintiff's demands, and were only dismissed if considered to be incomprehensible to the extent that it would harm the opposing party's ability to defend itself. Complaints often contained vague or generic claims, with unrealistic or arbitrarily defined demands.	Labor complaints must follow guidelines similar to those of civil complaints. Besides a brief description of the reasons for the claim, the complaint must now explicitly mention all of the plaintiff's demands and specify the precise amount being asked for each demand. Fail to do so may lead to the complaint being dismissed without prejudice.

claims that would be easily accepted prior to the reform now have a harder time being granted by a judge. Additionally, with the establishment of higher standards for granting exemption from payment of court and court-related fees, for many workers filing a complaint is no longer a costless endeavor. Even more so in frivolous or groundless cases, with the the real possibility of being ordered to pay a fine and having to pay attorney's fee and damages for the opposing side. With the reduced expected payoff of labor litigation, the number of lawsuits experienced a sharp drop right after the reform came into effect. Figure 3.2.2 shows the number of labor complaints filed per year, both in the 2nd Region and in Brazil. In both cases, the reform reverted an upward trend, resulting in a drop of over 33% and 36%, respectively.

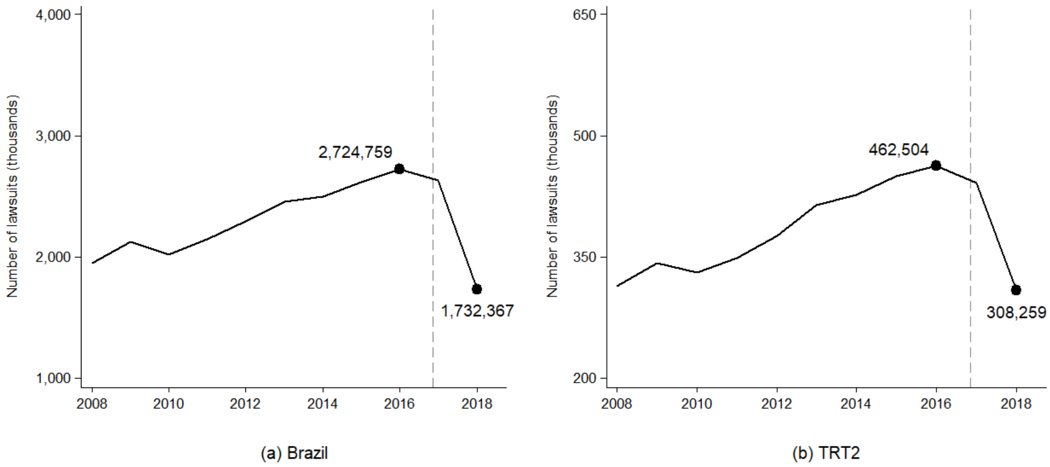


Figura 3.2.2 – Number of labor lawsuits filled per year in Brazil

3.2.3 Dataset Construction and Descriptives Statistics

Our empirical analysis is built upon labor court decisions over the period 2008-2013 published in the *Diário Oficial Eletrônico* (DOE)[3.2.18], the oficial register of by *Tribunal Regional do Trabalho da 2ª Regiao* (TRT2). This data includes

[3.2.18] Available in <<https://aplicacoes1.trtsp.jus.br/ConsultaDOE/doe/completo.jsp>>. Starting in 2014, the DOE was phased out in favor of *Diário Eletrônico da Justiça do Trabalho* (DEJT).

the name of all parties involved, (plaintiff, defendant, lawyers, and judge), court number, filing date, decision date, categorical ruling (pro-worker or pro-firm) and the value of the claim[3.2.19].

The second main source of data is *Relação Anual de Informações Sociais* (RAIS), a yearly matched employer-employee administrative dataset that comprises information on the universe of Brazilian labor market contracts. This dataset includes workers' information on workers such as age, gender, education level, race, occupation and wage, as well as firms characteristics such as sector, establishment size and location for firms.

Finally, the third source of data was gathered from the *Tribunal de Justiça de São Paulo* (TJSP), which stores legal information about the judicial cases - in particular, the bankruptcy requests - filed in the state of São Paulo. We collected the information of the $7,715$ bankruptcy requests filed between 2000 and 2015. This includes the full names of firms that went bankrupt, their tax identifier, and the date of bankruptcy.

Sample. Using workers, judges and firm full names, we are able to merge the first two sources of data, and then, with the firm tax identifier, merge it with the bankruptcy dataset. Table 3.2.3 describes the several steps leading to our final sample of labor litigation cases. In Panel A we report the total number of usable observations. We are able to find $1,758,007$ cases out of a total of over 2 million cases filed in a TRT2 labor court over the 2008-2013 period[3.2.20].

[3.2.19]We have built an algorithm, which relies basically in the massive use of the Python's package *Pandas* and regular expressions, that converted the DOE files in html format and from which we gather lawsuits unique identifiers. We have reached all the informations from available lawsuits by searching these identifiers in the TRT2 website.

[3.2.20]According to *CNJ*, $2,126,892$ new cases were filed in the TRT2 labor courts over the 2008-2013 period. The $1,758,007$ number of cases that we have found in the DOE is the number of cases that were filed in those years and that have not been inactive in the same period.

Tabela 3.2.3 – Database description

	Lawsuit's filing year						
	2008	2009	2010	2011	2012	2013	Total
Panel A: Lawsuits that we have found in the labor justice journal							
Total number of lawsuits according to *CNJ*	314,445	342,771	330,364	348,885	376,497	413,930	2,126,892
Lawsuits in the DOE (A)	258,123	281,623	297,461	298,936	324,760	297,104	1,758,007
Proportion of the total number of lawsuits as reported by *CNJ*	0.821	0.822	0.900	0.857	0.863	0.718	0.827
Lawsuits in which firms, universities or unions are suing workers (B)	11,948	11,175	15,308	17,021	16,831	12,707	84,990
Lawsuits that do not have the parties identification (C)	9,886	10,450	26,089	17,569	19,770	24,595	108,359
Panel B: Matching step							
Lawsuits in the DOE (D=A-B-C)	236,289	259,998	256,064	264,346	288,159	259,802	1,564,658
Parties for which we found correspondence in RAIS (% lawsuits)							
Defendants	0.64	0.66	0.68	0.68	0.69	0.70	0.68
Plaintiffs	0.78	0.79	0.80	0.81	0.81	0.81	0.80
Judges	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Defendants and plaintiffs	0.51	0.53	0.55	0.56	0.57	0.58	0.55
Defendants and judges	0.64	0.66	0.67	0.68	0.69	0.70	0.67
Plaintiffs and judges	0.78	0.79	0.80	0.81	0.81	0.81	0.80
Defendants, plaintiffs and judges	0.51	0.53	0.55	0.56	0.57	0.57	0.55
Panel C: Restrictions on the dataset							
Lawsuits in which we haven't found firm and judge in RAIS (E)	84,586	88,356	83,330	84,223	90,150	79,104	509,749
Lawsuits that do not have a clear decision (F)	20,697	22,425	23,071	23,502	26,462	24,119	140,276
Lawsuits in which a settlement was reached (G)	64,763	75,281	80,383	83,254	91,531	84,580	479,792
Lawsuits from judges that had judged only 1 case in the sample (H)	2	1	0	1	1	5	10
Lawsuits from firms with more than 1 case in the sample (I)	60,581	68,759	64,787	68,918	75,243	67,305	405,593
Lawsuits that have been judged after 2015 (J)	44	47	61	98	208	539	997
Lawsuits from inactive or non-private sector firms (K)	3,049	2,397	1,977	1,821	1,830	1,518	12,592
Lawsuits filed in small courts or that had been reviewed in the Appeal Court (L)	1,124	1,138	1,099	1,167	1,313	929	6,770
Final database (M=D-E-F-G-H-I-J-K-L)	1,443	1,594	1,356	1,362	1,421	1,703	8,879

We focus our attention on cases whose plaintiff is not identified as a firm, union, church, university, or a municipality. Observations in which the names of the parties were not available were also droped. Our baseline sample further restricts the dataset to lawsuits that we were able to found the judge and the firm in RAIS, and to lawsuits in which the parties did not reach a settlement prior to the judge's decision since we do not have information on how far a settlement was from the initial claim. We also kept only the private sector[3.2.21] firms that went to court a single time and only the judges that appeared more than once in the 2008-2013 period in order to avoid collective actions and outlier judges. Finally, we select only the lawsuits that have been judged in big courts[3.2.22] no later than 2015, that have not been reviewed in the Appeal Court, and against firms that were active in the lawsuit's filing year or the year before. Table 3.2.3 shows how all these drops led us to our final sample of $8,879$ labor lawsuits.

Descriptive statistics. Table 3.2.4 reports some summary statistics for our final estimation sample. The worker's claim is (partially or totally) accepted by the judge in 72% of the cases, and the average amount of compensation granted is equivalent to 9.2 months of the worker's last reported monthly wage. Besides, the first instance result is contested in the appeal court in 46% of the cases.

Regarding plaintiffs, around 18% of the workers that are suing firms have a college degree and almost none are reported as illiterate. These workers earned, on average, $1,590$ BRL in the last reported job, which is almost 3 times the average minimum wage of 539 BRL in the 2008-2013 period. Firms employ on average almost 17 workers, but 50% of them employ up to 5 workers. In spite of being very small, the median firm is over 8 years old.

3.3 Identification and reduced form estimations

In this section, we discuss our empirical strategy and the reduced-form evidence. We begin by discussing the challenges on econometric identification and

[3.2.21] With the exception of firms in the agricultural, domestical services, or extractive industry sector.
[3.2.22] Courts in which more than $1,000$ lawsuits were filed each year.

Tabela 3.2.4 – Summary statistics

	Mean	Std. Dev.	Median	Min	Max	N
Lawsuit characteristics						
- Claim accepted	0.72	0.45	1.00	0.0	1.0	8,879
- Compensation in worker's monthly wage	9.17	16.42	4.14	0.0	284.5	3,786
- Result contested in the appeal court	0.46	0.50	0.00	0.0	1.0	8,879
Worker characteristics						
- Illiterate worker	0.00	0.04	0.00	0.0	1.0	3,690
- Worker with college degree	0.18	0.39	0.00	0.0	1.0	3,690
- Worker's last monthly wage (th BRL)	1.59	2.58	0.97	0.0	54.8	4,718
Firm characteristics						
- Firm size (workers)	17.46	93.16	5.00	0.0	4,118.0	5,905
- Firms with less than 10 workers	0.46	0.50	0.00	0.0	1.0	5,905
- Firm age in years	12.16	10.83	8.71	0.4	63.4	3,869
- Firm sector: manufacturing	0.12	0.33	0.00	0.0	1.0	8,706
- Firm sector: retail	0.39	0.49	0.00	0.0	1.0	8,706
- Firm sector: food and hotel	0.09	0.29	0.00	0.0	1.0	8,706

how we have built our instrument. We then present our main findings on the effect of judge toughness on firm outcomes.

3.3.1 Empirical strategy

We are interested in identifying the causal effect of labor lawsuits on firm outcomes, such as new hires, employment, average wage of employees, and survival. In order to do so, we begin with the regression model

$$Y_{ijcT} = \beta_0 + \beta_1 \cdot I_{ijct} + X'_{ijct} \cdot \gamma + \epsilon_{ijcT} \tag{3.3.1}$$

where Y_{ijcT} is the outcome of the firm sued in case i, assigned to judge j in court c, measured $T - t > 0$ years after the judge decision. I_{ijct} is an indicator variable that equals one if the judge accepted the worker's claim and considered the firm guilty in period t, and X_{ijct} is a vector of controls variables that includes court $\times$ year fixed effects and other case-level charactheristics. However, OLS estimates of this equation are likely to be biased as I_{ijct} is correlated with case-characteristics that are unobserved, such as the quality of the lawyers in both sides of the lawsuit. We address this concern by exploiting that, conditional on year and court fixed effects, cases are randomly assigned to judges. Moreover, judges vary in how tough with the firms they are. We then use this judge toughness as an instrument to I_{ijct} in order to assess the causal effects of labor lawsuits. Our baseline empirical model is given by:

$$I_{ijct} = \alpha_0 + \alpha_1 \cdot z_{ijct} + X'_{ijct} \cdot \phi + v_{ijct} \tag{3.3.2}$$

$$Y_{ijcT} = \beta_0 + \beta_1 \cdot I_{ijct} + X'_{ijct} \cdot \gamma + \epsilon_{ijcT} \tag{3.3.3}$$

where z_{ijct} is the judge j toughness measure for case i in the court c and year t. We estimate β_1, our coefficient of interest, using two-stage least squares (2SLS) with equation (3.3.2) as the first stage and equation (3.3.3) as the second stage.

Instrumental variable construction. We construct our z_{ijct} instrument using a residualized, leave-one-out judge toughness measure (DAHL; KOSTOL; MOGSTAD, 2014; DOBBIE; GOLDIN; YANG, 2018; BHULLER et al., 2019; CAHUC; CARCILLO; PATAULT, 2019). More formally, to compute the judge bias as in Cahuc, Carcillo e Patault (2019), we can estimate:

$$I_{ijct} = \eta_{ct} + z_{ijct} \tag{3.3.4}$$

where $E[z_{ijct}|\eta_{ct}]$ is assumed to be zero. This equation means that the lawsuit outcome I_{ijct} is assumed to be result of a random term plus a term that is common to all cases judged in the same court c and year t, which captures some regional particularities and trends. This common term, the court $\times$ year fixed effect, is defined by

$$\eta_{ct} = E[I_{ijct}|\eta_{ct}] \tag{3.3.5}$$

which can be estimated as

$$\hat{\eta}_{ct} = \frac{1}{n_{ct}} \sum_{i \in (c,t)} I_{ijct} \tag{3.3.6}$$

where n_{ct} is the number of cases judged in court c at year t, and $i \in (c, t)$ represents all the cases judged in this same court and year. Therefore, the estimator of the judge fixed effect, conditional on the court × year fixed effect is given by

$$\hat{z}_j = \frac{1}{n_j} \sum_{i \in j} \hat{z}_{ijct} \tag{3.3.7}$$

where $i \in j$ stands for all the cases judged by judge j. However, in order to avoid reflection problems when we analyze the correlation between judge j toughness and the case i outcome, we need to estimate the toughness of the assigned judge as the toughness of the judge in all of her other cases, except for the concerned one. Thus, the toughness of judge j is measured by the leave-one-out mean of case i, which makes this measure both judge and case specific. The toughness measure of judge j for case i is finally defined as:

$$\hat{z}_{ijct} = \frac{1}{n_j - 1} \left(\left(\sum_{i \in j} \hat{z}_{ijct} \right) - \hat{z}_{ijct} \right) \tag{3.3.8}$$

First stage. Figure 3.3.1 shows the distribution of our judge toughness measure as defined in equation 3.3.8. Figure 3.3.2 presents the distribution of an alternative measure that is build using the amount granted by the judge in workers' monthly wages as the case outcome Y_{ijct} in equation 3.3.4. In both cases there is considerable variability in how tough judges are, showing that regardless the dimension in which we are measuring the judge toughness prosecuted firms are exposed to sistematically different judges. Being assigned to one of the 10% most pro-worker judges as compared to the 10% most pro-firm judges increases the probability of the claim being accepted from 61% to 82%, and the amount granted by the judge from 5.5 to 13.1 months of wage.

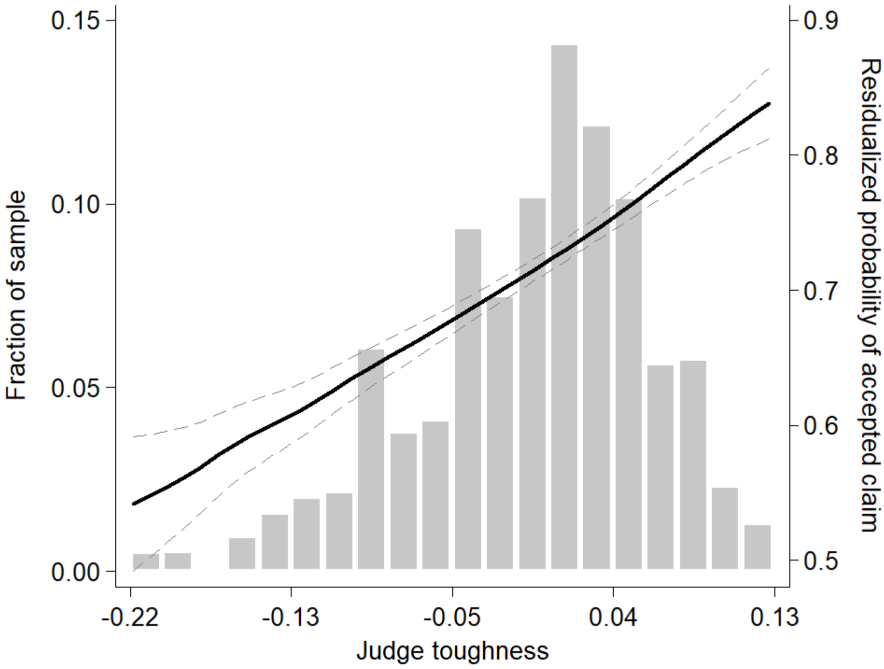


Figura 3.3.1 – Judge toughness measure with respect to the acceptance of the claim

Notes: The histogram shows the density of judge toughness along the left-axis (top and bottom 1% excluded). The solid line is the local polynomial fit of the residualized lawsuit outcome explained by the judge toughness. Dashed lines represent 90% confidence intervals.

Figures 3.3.1 and 3.3.2 also show in the right-axis a graphical and flexible representation of our first-stage estimations in which we plot the local polynomial fit of the residualized lawsuit outcome explained by the judge toughness. In both cases the lawsuit outcome is positively correlated with our measure of judge toughness and almost linear. Tables 3.3.1 and 3.3.2 show the first-stage estimations from equation 3.3.2. Our coefficient of interest is stable across specifications and always highly significant. Consistent with the graphical version, these estimates show that the instrument is highly predictive of lawsuit outcomes. In the specification with the full set of controls, an increase of one standard deviation in the judge toughness measure means a 5.5% higher probability of the claim being accepted and a 4.5 months of wage higher compensation.

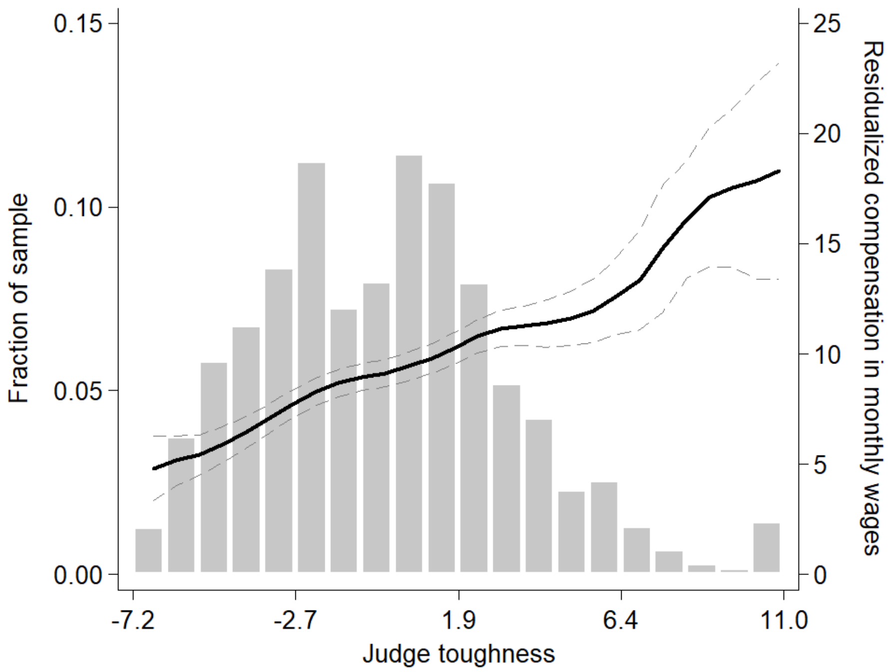


Figura 3.3.2 – Judge toughness measure with respect to the worker's compensation

Notes: The histogram shows the density of judge toughness along the left-axis (top and bottom 1% excluded). The solid line is the local polynomial fit of the residualized lawsuit outcome explained by the judge toughness. Dashed lines represent 90% confidence intervals.

In order to assess if cases are indeed randomly assigned to judges and assess the validity of our instrument, we look whether our judge toughness measure is correlated with case-characteristics. The first column of tables 3.3.3 and 3.3.4 show the regression of the lawsuit outcome on the worker's characteristics and firm's characteristics. We control for court × year fixed effects and standard errors are clustered at the judge level. We show that the firm probability of losing the case decreases as the firm become older and if the firm is sued by workers with college degree. The worker's education is also highly predictive of the amount granted by the judge. The second column of both tables, however, shows that regardless the instrument there is evidence that different judges are assigned to very similar cases. In spite of some individual significant coefficients they are all considerable smaller than in column 1 and the p-value of the joint F-test is equal to 0.689 in table 3.3.3

Tabela 3.3.1 – Effect of judge toughness on worker's claim being accepted

	Dependent variable: Claim accepted				
	(1)	(2)	(3)	(4)	(5)
Judge toughness	0.8539***	0.8548***	0.7740***	0.9978***	0.8115***
	[0.0811]	[0.0810]	[0.1378]	[0.1205]	[0.1992]
Court × time fixed effects	No	Yes	Yes	Yes	Yes
Plaintiff characteristics	No	No	Yes	No	Yes
Defendant characteristics	No	No	No	Yes	Yes
Observations	8,879	8,879	2,811	3,869	1,343
Adjusted R2	0.016	0.017	0.014	0.033	0.033
IV mean	-0.003	-0.003	-0.003	-0.003	-0.003
IV std dev	0.068	0.068	0.068	0.068	0.068

Notes: In brackets, standard errors are clustered at the judge level. Significance at the 10% level is represented by $*$, at the 5% level by $**$, and at the 1% level by $***$. Court × year fixed effects are used. The plaintiff characteristics included in column (3) and (5) are dummy variables that account for the worker's education and the value of the last reported monthly wage. Defendant characteristics in columns (4) and (5) include firm age in years and dummies for the firm sector.

and equal to 0.504 in table 3.3.4.

Tabela 3.3.2 – Effect of judge toughness on the compensation for the claim

	Dependent variable: Compensation as worker's last monthly wage				
	(1)	(2)	(3)	(4)	(5)
Judge toughness	0.6308***	0.6387***	0.6353***	0.6952***	0.6894***
	[0.0711]	[0.0712]	[0.1094]	[0.1376]	[0.2214]
Court × time fixed effects	No	Yes	Yes	Yes	Yes
Plaintiff characteristics	No	No	Yes	No	Yes
Defendant characteristics	No	No	No	Yes	Yes
Observations	3,786	3,786	2,264	1,919	1,124
Adjusted R2	0.021	0.023	0.040	0.033	0.048
IV mean	-0.083	-0.083	-0.083	-0.083	-0.083
IV std dev	3.758	3.758	3.758	3.758	3.758

Notes: In brackets, standard errors are clustered at the judge level. Significance at the 10% level is represented by $*$, at the 5% level by $**$, and at the 1% level by $***$. Court × year fixed effects are used. The plaintiff characteristics included in column (3) and (5) are dummy variables that account for the worker's education and the value of the last reported monthly wage. Defendant characteristics in columns (4) and (5) include firm age in years and dummies for the firm sector.

Tabela 3.3.3 – Randomization test for the judge toughness measure with respect to the acceptance of the claim

	Claim accepted	Judge toughness
	(1)	(2)
Worker characteristics		
- Worker with college degree	−0.0635*	0.0079*
	[0.0369]	[0.0044]
- Worker's last monthly wage (th BRL)	0.0023	0.0002
	[0.0061]	[0.0008]
Firm characteristics		
- Firms with less than 10 workers	0.0380	−0.0015
	[0.0254]	[0.0034]
- Firm age in years	−0.0045***	0.0001
	[0.0011]	[0.0002]
- Firm sector: manufacturing	−0.0322	0.0019
	[0.0400]	[0.0054]
- Firm sector: retail	0.0559**	0.0037
	[0.0280]	[0.0037]
- Firm sector: food and hotel	0.0639	0.0048
	[0.0423]	[0.0065]
Joint F-test	0.000	0.689
Observations	1,343	1,343

Notes: In brackets, standard errors are clustered at the judge level. Significance at the 10% level is represented by $*$, at the 5% level by $**$, and at the 1% level by $***$. Court $\times$ year fixed effects are used.

Tabela 3.3.4 – Randomization test for the judge toughness measure with respect to the worker's compensation

	Compensation as worker's last monthly wage	Judge toughness
	(1)	(2)
Worker characteristics		
- Worker with college degree	−2.6084***	0.2901
	[0.6755]	[0.3590]
- Worker's last monthly wage (th BRL)	−0.7206***	0.0665
	[0.1317]	[0.0480]
Firm characteristics		
- Firms with less than 10 workers	0.3742	−0.0336
	[0.9202]	[0.2151]
- Firm age in years	−0.0237	0.0050
	[0.0444]	[0.0133]
- Firm sector: manufacturing	0.1048	−0.1781
	[1.2963]	[0.3250]
- Firm sector: retail	2.5232**	−0.0243
	[1.0121]	[0.2417]
- Firm sector: food and hotel	1.0463	−0.2975
	[1.0782]	[0.3883]
Joint F-test	0.000	0.504
Observations	1,124	1,124

Notes: In brackets, standard errors are clustered at the judge level. Significance at the 10% level is represented by $*$, at the 5% level by $**$, and at the 1% level by $***$. Court $\times$ year fixed effects are used.

3.3.2 Results on firm outcomes

Table 3.3.5 presents the results of reduced form estimations in which we regress the firm outcomes on both judge toughness measures. Panel A shows that the growth rate of employment decreases by 2.1% and the growth rate of the average wage of new hires by 0.8% as we increase the judge toughness with respect to the acceptance of the claim in one standard deviation. There is also evidence of an increase in the probability of the firm becoming inactive and in the probability of the firm going bankrupt after the judge's decision. On the other hand, panel B shows that an increase in the judge toughness measure with respect to the worker's compensation seems to impact only the growth rate of the average wage of new hires and the probability of the firm going bankrupt.

Tabela 3.3.5 – Effect of judge toughness on firm's outcomes - Reduced form regression

	Growth rate between $t-1$ and $t+1$					
	Employment	Avg wage	Avg wage of entrants	Active in $[t, t+1]$	Liquidation after t	Liquidation before t
Panel A: Instrument 1						
Judge toughness	−0.3073**	−0.0041	−0.1112**	−0.1260*	0.0114***	0.0043
	[0.1476]	[0.0535]	[0.0543]	[0.0707]	[0.0043]	[0.0041]
Observations	8,706	6,884	5,196	8,706	8,706	8,706
IV mean	-0.003	-0.003	-0.003	-0.003	-0.003	-0.003
IV std dev	0.068	0.068	0.068	0.068	0.068	0.068
Panel B: Instrument 2						
Judge toughness	−0.0001	−0.0003	−0.0029*	0.0008	0.0003*	0.0001
	[0.0045]	[0.0016]	[0.0017]	[0.0020]	[0.0001]	[0.0001]
Observations	3,739	3,045	2,367	3,739	3,739	3,739
IV mean	-0.083	-0.083	-0.083	-0.083	-0.083	-0.083
IV std dev	3.758	3.758	3.758	3.758	3.758	3.758
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Sector fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: In brackets, standard errors are clustered at the judge level. Significance at the 10% level is represented by $*$, at the 5% level by $**$, and at the 1% level by $***$.

These results can be viewed as the firm's response to an increase in labor costs and in labor regulations. Conditional on having been sued, a firm that is randomly assigned to a judge that is more pro-worker than the average faces an exogenous increase in the expected labor costs and reacts by reducing both the growth rate of its employment level and the growth rate of the average wage of new hires. Moreover, the increase in total costs implies an exogenous reduction in profits, which leads to a higher probability of the firm becoming inactive and going bankrupt after the judge's decision.

However, these reduced form results alone do not fully address the question of how firms react to changes in expected labor costs and labor regulations such as the ones introduced by the Brazil's Labor Reform of 2017. In order to understand the mechanisms behind table 3.3.5 results, we now build a search-matching model in which dismissed workers have the option to sue the firm. We also perform counterfactual analysis to evaluate the impact that the Labor Reform had on the brazilian firms outcomes.

3.4 An equilibrium search and matching model with labor courts

Our model is based on a search and matching framework in continuous time where workers are risk neutral, infinitely lived and discount the future at rate r. Labor market frictions are characterized by a standard matching function. When a meeting between a firm and an unemployed worker takes place, they accept keeping the match if and only if the net surplus of the match is higher than zero. Upon meeting and accepting the match, a match-specific productivity x is drawn from a stationary distribution $G(x)$ and then workers and firms engage in Nash bargaining. A match can be destroyed at an exogenous rate δ.

We extend this standard environment by allowing workers who are laid off to decide whether to file a lawsuit against the firm or not, namely, a plaintiff or non-plaintiff unemployed. This decision is taken soon after the worker gets fired and before any new match takes place. The idea is that the plaintiff unemployed receives a positive compensation that depends on the job productivity as well as may obtain an idiosyncratic gain attributed to a more favorable case. On the other hand, the plaintiff joins a blacklist from which firms hire at a lower rate. We assume that the worker's litigation history resets every time she is admitted to a job such that no history is carried forward.

The details of this search and matching process are better described as follows.

3.4.1 Workers

While working, individuals enjoy the wage that depends on the job match productivity, x, and may receive a termination shock at the rate δ in which case they immediately decide whether to file a lawsuit against the firm or not. In case of dispute, workers (plaintiff) obtain the value of unemployment U_1 in addition to an expected gain $A(x)$ that vary with the worker's productivity x. They also receive an i.i.d. draw of an idiosyncratic "case quality"shock which we assume is separate from the value of unemployment. ϵ has a Type-I Extreme Value distribution with mean zero and scale parameter equal to σ. When workers decide not to file a lawsuit against the firm (non-plaintiff), they receive the value of unemployment U_2.

For the plaintiff unemployed, the contact rates (λ_i) are lower as employers observe whether the job seeker has filed a lawsuit against the previous employer, such that they fill vacancies with these workers at a lower intensity, i.e. $\lambda_1 < \lambda_2$. The flow value of an unemployed in the situation i=1,2 is

$$rU_i = b + \lambda_i \int \int \max\left\{W(w(x), \epsilon) - U_i, 0\right\} dH(\epsilon) dG(x) \tag{3.4.1}$$

where b is the flow utility while unemployed.

The flow value of a job is

$$rW(w(x), \epsilon) = w(x; \epsilon) + \delta \left[\max\left\{U_1 + A(x) + \sigma\epsilon, U_2\right\} - W(w(x), \epsilon)\right] \tag{3.4.2}$$

where the wage is a function of productivity x and the idiosyncratic shock ϵ that affects the worker's outside option when bargaining with the employer.

3.4.2 Firms

Firms post vacancies. The flow value of an open vacancy is

$$rV = -c + \zeta_1 \int \int_{\epsilon_x^*}^{\infty} \max\left\{J(x,\epsilon) - V, 0\right\} dH(\epsilon) dG(x) + \\ \zeta_2 \int \int_{-\infty}^{\epsilon_x^*} \max\left\{J(x,\epsilon) - V, 0\right\} dH(\epsilon) dG(x) \quad (3.4.3)$$

where c is the per period vacancy cost, ζ_1 is the rate at which employers meet the plaintiff unemployed, and ϵ_x^* is the level of idiosyncratic gain that determines the decision of the unemployed to file a lawsuit against the firm, conditional on the match productivity. Since firms have a preference for unemployed who did not file a lawsuit against the previous firm, they hire such workers at a higher intensity $\zeta_2 = \psi\zeta_1$, where $\psi > 1$ and is assumed exogenous.[3.4.1]

The value of a filled job depends on the profit per period $x - w(x;\epsilon)$ and the expected termination occurring at rate δ, in which case the firm may be sued with probability $\Phi(x)$.[3.4.2] In this case, they are expected to pay a fixed cost due to hiring defense attorneys K in addition to $A(x)$ (the transfer to the unemployed who filed a lawsuit). The flow value of a filled job is

$$rJ(x,\epsilon) = x - w(x;\epsilon) + \delta\left\{V - \Phi(x)\left[K + A(x)\right] - J(x,\epsilon)\right\} \quad (3.4.4)$$

[3.4.1] An interesting extension for future work would allow for directed search, with firms choosing a selection rule for workers type 2 over type 1 as well as wages.

[3.4.2] The idea here is that the litigation history of the worker affects his/her wage at the time of hiring but then resets after hiring. Therefore if a firm hires a "high ϵ"worker (a plaintiff), this does not imply that the worker is going become a plaintiff for sure again in case of firing.

3.4.3 Equilibrium

In this section we solve for endogenous objects, namely, the contact rates, the reservation shock that determines the unemployed's choice between filing a lawsuit or not, wages and reservation productivity.

3.4.3.1 Matching

We assume that the meeting between employers and employees are governed by a matching function. As it is standard in the literature we assume a Cobb-Douglas function, such that the number of meetings is given by

$$m = (u_1 + u_2)^{1-\eta} v^{\eta}$$

where $u_1 + u_2$ is the total number of unemployed in the market, respectively, plaintiff and non-plaintiff, and v the number of vacancies.[3.4.3] Given the market tightness $\theta = \frac{v}{u_1+u_2}$, the firms hire the unemployed type 1 (plaintiff) and type 2 (non-plaintiff) at the following rates

$$\zeta_1 = \frac{m}{v} = \theta^{\eta-1}$$
$$\zeta_2 = \frac{\psi m}{v} = \psi\theta^{\eta-1}$$

Likewise, since vacancies are more visible to non-plaintiff unemployed, these individuals also meet firms more frequently. The contact rates for the unemployed λ_i $(i = 1, 2)$ are

[3.4.3] The elasticity of the matching function with respect to vacancies, η, is typically between 0.3-0.5 (Petrongolo and Pissarides, 2001).

$$\lambda_1 = \frac{m}{u_1}\frac{u_1}{u_1+u_2} = \theta^\eta$$
$$\lambda_2 = \frac{\psi m}{u_2}\frac{u_2}{u_1+u_2} = \psi\theta^\eta$$

3.4.3.2 The unemployed's choice

Individuals who enter the labor market search as a non-plaintiff unemployed. But previously employed workers who lost their jobs direct their search as plaintiff or non-plaintiff unemployed, situations that are denoted by 1 and 2, respectively. Since taking the firm to court implies a lower contact rate from a next future employer but entitles the worker for some positive compensation as well as an idiosyncratic gain, the worker who just lost a job decides based on the following maximization:

$$\max\{U_1 + A(x) + \sigma\epsilon, U_2\} \tag{3.4.5}$$

Given that ϵ is drawn from an i.i.d. Type-I Extreme Value distribution with mean zero and scale parameter equal to σ, the probability that an unemployed individual with productivity x takes a case to the labor court is

$$\Pr\{\epsilon > \epsilon_x^* \mid x\} = \Phi(x) = \frac{\exp(\frac{U_1+A(x)}{\sigma})}{\exp(\frac{U_1+A(x)}{\sigma}) + \exp(\frac{U_2}{\sigma})}$$

where ϵ_x^* solves $U_1 + A(x) + \sigma\epsilon_x^* = U_2$, given x.

3.4.3.3 Wage determination

After drawing a shock ϵ and deciding whether to become a plaintiff or non-plaintiff against the previous employer, she searches for jobs.

When firms and workers meet, they draw a match-specific productivity $x \sim G(x)$ and they start the bargaining process to determine the wage and to decide if accepting the match or not. The firm observes the unemployed status (plaintiff or non-plaintiff) and the realized shock ϵ. Thus, for each x and ϵ, the wage is obtained through Nash bargaining as follows,[3.4.4]

$$\max_{w} \{W(w(x), \epsilon) - \max\{U_1 + A(x) + \sigma\epsilon, U_2\}\}^{\beta} \{J(x, \epsilon) - V\}^{1-\beta}$$

where β is the labor share, taken as exogenous. In equilibrium with free entry of vacancies, $V = 0$. From the first order condition, we arrive at the wage function is

$$w(x; \epsilon) = (1 - \beta) r \max\{U_1 + A(x) + \sigma\epsilon, U_2\} + \beta(x - \delta\Phi(x)[K + A(x)]) \quad (3.4.6)$$

The wage in equilibrium is an average between the flow value of their outside option given by $\max\{U_1 + A(x) + \sigma\epsilon, U_2\}$ and workers' productivity x minus the cost of firms with a labor lawsuit which diminishes wages as firms anticipate that they will have to face them in the future.

This wage function also shows that when ϵ is lower than the cutoff value, i.e. when becoming a plaintiff is not profitable for the unemployed, then the wage is always increasing in productivity. However, for any ϵ higher than the cutoff value, the unemployed benefits from $A(x)$ which may increase or decrease in with labor productivity such that wages do not need to always increase in productivity. This could be the case for example if labor judges are biased towards low skilled workers in which case the probability of winning a lawsuit decreases on productivity.

3.4.3.4 Reservation productivity

As in standard search-matching models, accepting a job match depends on a reservation productivity that makes the workers and firms indifferent between an

[3.4.4] Note that since wages are only negotiated at the time of hiring, breaking off the bargaining does not bring any lawsuit benefit to the worker or cost to the firm.

agreement and keep searching. Since the bargaining takes place only at the hiring, this is defined by the following zero-surplus condition, for any ϵ,

$$W(w(x_\epsilon^*), \epsilon) - \max\{U_1 + A(x_\epsilon^*) + \sigma\epsilon, U_2\} + J(x_\epsilon^*, \epsilon) - V = 0$$

We then use this restriction, together with $V = 0$ and the value functions (2) and (4) to obtain:

$$x_\epsilon^* = r \max\{U_1 + A(x_\epsilon^*) + \sigma\epsilon, U_2\} + \delta\Phi(x_\epsilon^*)\left[K + A(x_\epsilon^*)\right] \tag{3.4.7}$$

If we assume $A(x)$ is monotone, a unique fixed point solution is obtained for any given shock ϵ. This equation makes it clear that a decrease in K or $A(x)$ shifts the reservation productivity down implying more matches. Our model suggests that a labor reform that reduces the firm's costs and the worker's benefits has the potential to increase employment and reduce the fraction of plaintiff among the unemployed.

3.4.4 Simulations

The model is calibrated using our data from TRT2 labor lawsuits in the 2008-2013 period. We then use it to evaluate the effects of the changes introduced by the 2017 Labor Reform. As explained in section 2, by shifting lawsuit costs to the losing party and imposing more restrictive rules on how detailed a worker's claim should be, the 2017 Labor Reform reduced the firm's expected costs and the worker's expected benefits with the lawsuit, which implies a lower K and $A(X)$, respectively. The first counterfactuals policy simulations points to the direction of an increase in employment and a decrease in the relative number of lawsuits among the unemployed. Therefore, both the reduced form evidence and the conterfactual analysis imply an increase in employment level in response to a reduction in labor costs and labor regulations. The complete set of counterfactuals results, tables and figures should be available in the next version of this draft.

3.5 Conclusion

This paper evaluates how labor regulations, enforced by labor judges in lawsuits decisions, affect firm survival and firm outcomes. We begin by presenting reduced form evidence on the effect that labor judges pro-worker bias has on some firm outcomes. Using firm data merged with a new dataset on labor lawsuits from the largest brazilian municipality and its neighboring municipalities, we show that the employment growth rate of a prosecuted firm that was randomly assigned to a more pro-worker judge decreases. There is also a negative effect on the growth rate of the average wage of new hires, and on the probability of the firm remaining active after the decision. We then interpret these reduced form results within a new search-matching model in which workers who are laid off decide whether to file a lawsuit against the firm or not. More than that, we use the model to evaluate the effects of the labor regulations changes introduced by the Brazil's 2017 Labor Reform. Similar to the reduced form evidence, the first counterfactuals simulations show that by reducing labor regulations and the expected costs with labor lawsuits, the Labor Reform may have indeed increased the employment level of the economy.

REFERENCES

ACEMOGLU, D.; ANGRIST, J. How large are the social returns to education? evidence from compulsory schooling laws. In: BERNANKE, B.; ROGOFF, K. (Ed.). *NBER Macroeconomic Annual*. [S.l.]: NBER, 2000. p. 9–59.

AHSAN, H.; HAQUE, M. E. Threshold effects of human capital: schooling and economic growth. *Economics Letters*, Vol. 156, p. pp. 48–52, 2017.

ALLCOTT, H.; GENTZKOW, M. Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, Vol. 31, n. No. 2, p. pp. 211–236, 2017.

ALMEIDA, R.; CARNEIRO, P. Enforcement of labor regulation and firm size. *Journal of Comparative Economics*, Elsevier, v. 37, n. 1, p. 28–46, 2009.

ALMEIDA, R.; CARNEIRO, P. Enforcement of labor regulation and informality. *American Economic Journal: Applied Economics*, v. 4, n. 3, p. 64–89, 2012.

ANSOLABEHERE, S. et al. Does attack advertising demobilize the electorate? *The American Political Science Review*, Vol. 88, n. No. 4, p. pp. 829–838, December 1994.

ARNOLD, D.; DOBBIE, W.; YANG, C. S. Racial bias in bail decisions. *The Quarterly Journal of Economics*, Oxford University Press, v. 133, n. 4, p. 1885–1932, 2018.

AUTOR, D. H.; III, J. J. D.; SCHWAB, S. J. The cost of wrongful-discharge laws. *The Review of Economics and Statistics*, Vol. 88, n. No. 2, p. pp. 211–231, 2006.

AUTOR, D. H.; KERR, W. R.; KUGLER, A. D. Does employment protection reduce productivity? evidence from us states. *The Economic Journal*, v. 117, n. 521, p. F189–F217, 2007. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0297.2007.02055.x>.

AVIS, E. et al. *Money and politics: the effects of campaign spending limits on political competition and incumbency advantage*. [S.l.], 2017.

BAMIEH, O. Firing costs, employment and misallocation: evidence from randomly assigned judges. 2017.

BARON, D. P. Electoral competition with informed and uninformed voters. *The American Political Science Review*, Vol. 88, n. No. 1, p. pp. 33–47, 1994.

BARRO, R.; LEE, J. A new data set of educational attainment in the world, 1950-2010. *Journal of Development Economics*, Vol. 104, p. pp. 184–198, 2013.

BARRO, R. J. Economic growth in a cross section of countries. *The Quarterly Journal of Economics*, Vol. 106, n. No. 2, p. pp. 407–443, May 1991.

BENHABIB, J.; SPIEGEL, M. M. The role of human capital in economic development: evidence from aggregate cross-country data. *Journal of Monetary Economics*, Vol. 34, n. No. 2, p. pp. 143–173, 1994.

BERNSTEIN, S.; COLONNELLI, E.; IVERSON, B. Asset allocation in bankruptcy. *The Journal of Finance*, v. 74, n. 1, p. 5–53, 2019. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/jofi.12740>.

BHULLER, M. et al. Incarceration, recidivism and employment. *Journal of Political Economy*, 2019.

BILS, M.; KLENOW, P. J. Does schooling cause growth? *The American Economic Review*, Vol. 90, n. No. 5, p. pp. 1160–1183, 2000.

BOTERO, J. C. et al. The regulation of labor. *The Quarterly Journal of Economics*, MIT Press, v. 119, n. 4, p. 1339–1382, 2004.

CAHUC, P.; CARCILLO, S.; PATAULT, B. Are pro-worker judges detrimental to firm survival and employment? 2019.

CASELLI, F. Accounting for cross-country income differences. In: AGHION, P.; DURLAUF, S. N. (Ed.). *Handbook of Economic Growth*. [S.l.]: Elsevier, 2005. Vol. 1A, cap. 9, p. pp. 680–738.

CATTANEO, M. D.; M., J.; MA, X. Simple local polynomial density estimators. *Journal of the American Statistical Association*, Vol. 0, n. No. 0, p. pp. 1–7, 2019.

CHEETY, R.; FRIEDMAN, J. N.; ROCKOFF, J. E. Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood. *The American Economic Review*, Vol. 104, n. No. 9, p. pp. 2633–2679, September 2014.

CHEN, K. Sabotage in promotion tournaments. *Journal of Law, Economics and Organization*, Vol. 19, n. No. 1, p. pp. 119–140, 2003.

COHEN, D.; SOTO, M. Growth and human capital: good data, good results. *Journal of Economic Growth*, Vol. 12, n. No. 1, p. pp. 51–76, 2007.

DAHL, G. B.; KOSTOL, A. R.; MOGSTAD, M. Family welfare cultures. *The Quarterly Journal of Economics*, Vol. 129, n. No. 4, p. pp. 1711–1752, 2014.

DELGADO, M. G. *Curso de Direito do Trabalho*. 11th. ed. São Paulo: LTR, 2012.

DELGADO, M. S.; HENDERSON, D. J.; PARMETER, C. F. Does education matter for economic growth? *Oxford Bulletin of Economics and Statistics*, Vol. 76, n. No. 3, p. pp. 334–359, 2014.

DEPEW, B.; EREN, O.; MOCAN, N. Judges, juveniles, and in-group bias. *The Journal of Law and Economics*, v. 60, n. 2, p. 209–239, 2017. Disponível em: <https://doi.org/10.1086/693822>.

DESPOSATO, S. *Going negative in comparative perspective: electoral rules and campaign strategies*. [S.l.], 2007.

DOBBIE, W.; GOLDIN, J.; YANG, C. S. The effects of pretrial detention on conviction, future crime, and employment: evidence from randomly assigned judges. *American Economic Review*, Vol. 108, n. No. 2, p. pp. 201–240, 2018.

DUFLO, E. Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *The American Economic Review*, Vol. 91, n. No. 4, p. pp. 795–813, September 2001.

EGGERS, A. C. et al. *Regression Discontinuity Designs Based on Population Thresholds: Pitfalls and Solutions*. [S.l.], 2015. Disponível em: <https://ideas.repec.org/p/diw/diwwpp/dp1503.html>.

FERRAZ, C.; FINAN, F. *Motivating Politicians: The Impacts of Monetary Incentives on Quality and Performance*. [S.l.], 2009. (Working Paper Series, 14906).

FIGLIO, D. N. Teacher salaries and teacher quality. *Economics Letters*, Vol. 55, n. No. 2, p. pp. 267–271, 1997.

FUENTE, A. de la; DOMENECH, R. Human capital in growth regressions: how much difference does data quality make? *Journal of the European Economic Association*, Vol. 4, n. No. 1, p. pp. 1–36, 2006.

FUJIWARA, T. A regression discontinuity test of strategic voting and duverger's law. *Quarterly Journal of Political Science*, Vol. 6, n. No. 3-4, p. pp. 197–233, 2011.

GALASSO, A.; SCHANKERMAN, M. Patents and cumulative innovation: Causal evidence from the courts. *The Quarterly Journal of Economics*, MIT Press, v. 130, n. 1, p. 317–369, 2015.

GALOR, O.; MOAV, O. Ability-biased technological transition, wage inequality, and economic growth. *The Quarterly Journal of Economics*, Vol. 115, n. No. 2, p. pp. 469–497, May 2000.

GALOR, O.; MOAV, O. From physical to human capital accumulation: Inequality and the process of development. *The Review of Economic Studies*, Vol. 71, n. No. 4, p. pp. 1001–1026, October 2004.

GAZAL-AYAL, O.; SULITZEANU-KENAN, R. Let my people go: Ethnic in-group bias in judicial decisions—evidence from a randomized natural experiment. *Journal of Empirical Legal Studies*, v. 7, n. 3, p. 403–428, 2010. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1740-1461.2010.01183.x>.

GHANDI, A.; IORIO, D.; URBAN, C. Negative advertising and political competition. *The Journal of Law, Economics and Organization*, Vol. 32, n. No. 3, p. pp. 433–477, August 2016.

GILPIN, G.; KAGANOVICH, M. The quantity and quality of teachers: Dynamics of the trade-off. *Journal of Public Economics*, Vol. 96, n. No. 4, p. pp. 417–429, April 2012.

GONZAGA, G.; MALONEY, W. F.; MIZALA, A. Labor turnover and labor legislation in Brazil [with comments]. *Economía*, JSTOR, v. 4, n. 1, p. 165–222, 2003.

GROSSMAN, G. M.; HELPMAN, E. Electoral competition and special interest politics. *The Review of Economic Studies*, Vol. 63, n. No. 2, p. pp. 265–286, 1996.

HALL, R. E.; JONES, C. I. Why do some countries produce so much more output per worker than others? *The Quarterly Journal of Economics*, Vol. 114, n. No. 1, p. pp. 83–116, 1999.

HANUSHEK, E. A. Economic growth in developing countries: The role of human capital. *Economics of Education Review*, Vol. 37, p. pp. 204–212, December 2013.

HANUSHEK, E. A.; PACE, R. R. Who chooses to teach (and why)? *Economics of Education Review*, Vol. 14, n. No. 2, p. pp. 101.–117, 1995.

HANUSHEK, E. A.; WOESSMANN, L. Schooling, educational achievement, and the latin american growth puzzle. *Journal of Development Economics*, Vol. 99, n. No. 2, p. pp. 497–512, November 2012.

HARRINGTON, J.; HESS, G. A spatial theory of positive and negative campaigning. *Games and Economic Behavior*, Vol. 17, n. No. 2, p. pp. 209–229, 1996.

IMBENS, G. W.; LEMIEUX, T. Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, Vol. 142, n. No. 2, p. pp. 615–635, fev. 2008.

IVERSON, B. C. et al. Learning by doing: Evidence from bankruptcy judges. *Available at SSRN 3084318*, 2018.

JACOBSON, G. C. How do campaigns matter? *Annual Review of Political Science*, Vol. 18, p. pp. 31–47, 2015.

JAIMOVICH, N.; REBELO, S. Nonlinear effects of taxation on growth. *Journal of Political Economy*, Vol. 125, n. No. 1, p. pp. 265–291, 2017.

KALAITZIDAKIS, P. et al. Measures of human capital and nonlinearities in economic growth. *Journal of Economic Growth*, Vol. 6, n. No. 3, p. pp. 229–254, 2001.

KNEPPER, M. When the shadow is the substance: Judge gender and the outcomes of workplace sex discrimination cases. *Journal of Labor Economics*, v. 36, n. 3, p. 623–664, 2018.

LAU, R. R.; ROVNER, I. B. Negative campaigning. *Annual Review of Political Science*, Vol. 12, p. pp. 285–306, 2009.

LAU, R. R. et al. The effects of negative political advertisements: a meta-analytic assessment. *American Political Science Review*, Vol. 93, n. No. 4, p. pp. 851–875, 1999.

LAU, R. R.; SIGELMAN, L.; ROVNER, I. B. The effects of negative political campaigns: a meta-analytic reassessment. *Journal of Politics*, Wiley Online Library, Vol. 69, n. No. 4, p. pp. 1176–1209, 2007.

LAZEAR, E. Pay equality and industrial politics. *Journal of Political Economy*, Vol. 97, n. No. 3, p. pp. 561–580, 1989.

LEE, S. Y. T.; SESHADRI, A. On the intergenerational transmission of economic status. *Journal of Political Economy*, Vol. 127, n. No. 2, p. pp. 855–921, 2019.

MANKIW, N. G.; ROMER, D.; WEIL, D. N. A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, Vol. 107, n. No. 2, p. pp. 407–437, May 1992.

MANUELLI, R. E.; SESHADRI, A. Human capital and the wealth of nations. *American Economic Review*, Vol. 104, n. No. 9, p. pp. 2736–2762, 2014.

MCCRARY, J. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, Vol. 142, n. No. 2, p. pp. 698–714, February 2008.

MCFADDEN, D. L. Conditional logit analysis of qualitative choice behavior. In: ZAREMBKA, P. (Ed.). *Frontier in Econometrics*. [S.l.]: Academic Press: New York, 1974. p. 105–142.

NAGLER, M.; PIOPIUNIK, M.; WEST, M. R. *Weak Markets, Strong Teachers: Recession at Career Start and Teacher Effectiveness*. [S.l.], 2015. Disponível em: <http://www.nber.org/papers/w21393>.

PASTINE, I.; PASTINE, T. Incumbency advantage and political campaign spending limits. *Journal of Public Economics*, Vol. 96, n. No. 1-2, p. pp. 20–32, 2012.

PERSSON, T.; TABELLINI, G. Is inequality harmful for growth? *The American Economic Review*, Vol. 84, n. No. 3, p. pp. 600–621, 1994.

PINHEIRO, A. C. *Judiciário, Reformas e Economia: A Visão dos Magistrados*. [S.l.], 2003.

PODGURSKY, M.; MONROE, R.; WATSON, D. The academic quality of public schools teachers: an analysis of entry and exit behavior. *Economics of Education Review*, Vol. 23, n. No. 5, p. pp. 507–518, 2004.

PRITCHETT, L. Where has all the education gone? *The World Bank Economic Review*, Vol. 15, n. No. 3, p. pp. 367–391, 2001.

PSACHAROPOULOS, G. Returns to investment in education: a global update. *World Development*, v. 22, n. 9, p. 1325–1343, September 1994.

PSACHAROPOULOS, G.; VALENZUELA, J.; ARENDS, M. Teacher salaries in latin america: a review. *Economics of Education Review*, Vol. 15, n. No. 4, p. pp. 401–406, 1996.

RIVKIN, S. G.; HANUSHEK, E. A.; KAIN, J. F. Teachers, schools, and academic achievement. *Econometrica*, Vol. 73, n. No. 2, p. pp. 417–458, March 2005.

ROCKOFF, J. E. The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, Vol. 94, n. No. 2, p. pp. 247–252, May 2004.

ROGERS, M. L. Directly unproductive schooling: how country characteristics affect the impact of schooling on growth. *European Economic Review*, Vol. 52, n. No. 2, p. pp. 356–385, 2008.

SAPKAL, R. S. Labour law, enforcement and the rise of temporary contract workers: empirical evidence from india's organised manufacturing sector. *European Journal of Law and Economics*, Springer, v. 42, n. 1, p. 157–182, 2016.

SCHOELLMAN, T. Education quality and development accounting. *The Review of Economic Studies*, Vol. 79, n. No. 1, p. pp. 388–417, January 2012.

SCHUNDELN, M.; PLAYFORTH, J. Private versus social returns to human capital: education and economic growth in india. *European Economic Review*, Vol. 66, p. pp. 266–283, 2014.

SHAYO, M.; ZUSSMAN, A. Judicial ingroup bias in the shadow of terrorism. *The Quarterly Journal of Economics*, Vol. 126, n. No. 3, p. pp. 1447–1484, 2011.

SILVEIRA, B. S.; MELLO, J. M. P. D. Campaign advertising and election outcomes: Quasi-natural experiment evidence from gubernatorial elections in brazil. *Review of Economics Studies*, Vol. 78, n. No. 2, p. pp. 590–612, February 2011.

SKAPERDAS, S.; GROFMAN, B. Modeling negative advertising. *American Political Science Review*, Vol. 89, n. No. 1, p. pp. 49–61, 1995.

TAMURA, R. Teachers, growth, and convergence. *Journal of Political Economy*, Vol. 109, n. No. 5, p. pp. 1021–1059, October 2001.

TEMPLE, J. R. W. Generalizations that aren't? evidence on education and growth. *European Economic Review*, Vol. 45, n. No. 4-6, p. pp. 905–918, 2001.

Tribunal Superior do Trabalho. *Relatório Geral da Justiça do Trabalho*. [S.l.], 2019.

WALTER, A. S. Negative campaigning in western europe: similar or different? *Political Studies*, Vol. 62, n. No. 1, p. pp. 42–60, 2014.

WALTER, A. S.; BRUG, W. Van der; PRAAG, P. V. When the stakes are high: party competition and negative campaigning. *Comparative Political Studies*, Vol. 47, n. No. 4, p. pp. 550–573, 2014.