Three essays about Brazilian public polices: broadband internet and political outcomes; effects of Pronatec; and returns to technical education

Thiago Mendes Rosa

08/07/2021

Three essays about Brazilian public polices: broadband internet and political outcomes; effects of Pronatec; and returns to technical education

Thiago Mendes Rosa

Tese apresentada ao Departamento de Economia da Universidade de Brasília como requisito parcial para a obtenção do título de Doutor em Economia – Área de Concentração: Setor Público.
Orientador: Rafael Terra de Menezes.

Acknowledgement

This work was only possible with the support of many people.

Firstly, I am thankful for my family support, notably my mother, Edelair, and my aunt, Ivone, who always encouraged education in my life.

I would like to thank all my professors who helped to pave my way until here, particularly Flávio.

I am especially grateful for the opportunity to have a couple of first quality supervisors, Rafael and Ana Carolina, whose expertise made this study feasible.

I would also like to thank Bruno and Luiz Rubens, for the partnership in research, co-authorship of the third essay and for the valuable friendship in work and life.

I very thankful for all support, love, and patience of Najla, my partner in life who always believed in me.

Lastly, I want to thank all Brazilian people, the truly providers of this unique opportunity.

Resumo

Essa tese contém três ensaios independentes sobre políticas públicas no Brasil.

No primeiro ensaio, investigamos a relação entre a velocidade de internet banda larga fixa e os resultados eleitorais no Brasil (2008, 2010 e 2012). Utilizando uma estratégia de identificação robusta, um RDD aplicado à implementação do programa Backhaul, exploramos saltos na velocidade da internet de acordo com o tamanho da população dos municípios. Os resultados indicam que não há relação entre a velocidade da internet e variáveis eleitorais – participação, percentual de votos brancos ou nulos, percentual de votos em partidos de esquerda, percentual de votos ou orçamento de partidos pequenos ou candidatos jovens. Esses resultados divergem das relações positivas/negativas reportadas anteriormente, geralmente aplicados a democracias com *background* institucional distinto daquele observado no Brasil, o que sugere que tal relação entre internet e resultados eleitorais pode não se aplicar em todos os contextos.

No segundo ensaio, avaliamos o impacto de um programa de treinamento profissional de curta duração (Pronatec) na empregabilidade e nos salários, no curto, médio e longo prazo, bem como "efeitos de transbordamento". Aproveitando as regras do programa, comparamos participantes com não-participantes por motivos involuntários (como excesso de procura por matrícula ou cancelamento de turma). Nossa base quase-experimental é associada a um rico registro administrativo, o qual contém as variáveis de resposta no mercado formal de trabalho, cobrindo o período de 2010 a 2019 e mais de quatro milhões de pessoas. Os resultados, para um desemprego de curta duração, indicam efeitos positivos logo após o término do programa (seis meses), duradouro até o sétimo ano (em geral, uma razão de chances entre 16% e 20% maior de estar empregado), com padrão similar para os salários. O percentual de concluintes em um mercado de trabalho específico impacta negativamente as chances de emprego e salário no curto prazo, sugerindo efeitos de transbordamento do Pronatec. Além disso, as estimativas indicam uma importante heterogeneidade, com efeitos menores para as mulheres, efeitos maiores para os jovens e nenhum efeito para cursos auto demandados. Finalmente, o Pronatec parece ser altamente econômico, com um retorno de R$ 1,89 para cada R$ 1,00 gasto.

No terceiro ensaio, estimamos o retorno da educação técnica no Brasil entre 2007 e 2018. Utilizando um rico registro administrativo do mercado de trabalho formal (RAIS), construímos um painel com todos os trabalhadores formais de 2007, seguindo-os até 2018. A identificação dos trabalhadores foi realizada com base no Catálogo Nacional de Ocupações Técnicas. Adicionalmente, utilizamos as duas rodadas da Pesquisa Nacional por Amostra de

Domicílios (PNAD) sobre educação técnica (2007 e 2014). Os resultados sugerem um prêmio salarial positivo e significante – entre 21,3% e 24,9% para os trabalhadores em ocupação técnica, controlando para todas as variáveis observáveis disponíveis (escolaridade, idade, tamanho da firma, setor econômico e tempo no emprego). Mesmo restringindo nossa amostra para jovens trabalhadores (18 anos) com ensino médio no início do período – perfil escolar necessário para uma ocupação técnica – o resultado se mantém, mas em uma magnitude menor (entre 5,8% e 7,8%). Contudo, os dados *cross-section* sugerem a inexistência de prêmio salarial para a geração jovem de 2014 e, portanto, coortes podem ser afetados de maneira distinta por esse tipo de educação. Considerando-se os benefícios encontrados, a educação técnica parece ser custo-efetiva caso demande até R$ 8.595,10 por mês (a uma taxa de desconto de 6% ao ano).

Palavras-chave: Internet, Eleições, Pronatec, Emprego, Salários, Educação Técnica.

Abstract

This thesis has three independent essays about regarding public policies in Brazil.

In the first essay, we investigate the relationship between broadband internet velocity and election outcomes in Brazil (2008, 2010 and 2012). Using a robust identification strategy, a RDD applied to the roll out of Backhaul program, we explore jumps in internet velocity according to population size as identification strategy. Results indicate no relationship between internet speed and political outcomes – turnout, blank and null percentage votes, left parties vote share, small party or young candidate vote share and campaign budget. Our findings diverge from some positive/negative results reported before, usually applied to democracies with institutional backgrounds distinct of the one observed in Brazil, suggesting that the relationship between internet and electoral outcomes does not apply everywhere.

In the second essay, we evaluate the impact of a Brazilian short training program (Pronatec) on employment and wages, in short, mid, and long runs, as well spillover effects. Taking advantage of program's rules, we compare participants with non-participants due to involuntary reasons (like over subscription or cancellation of classes). Our quasi-experimental database is linked to a rich administrative data to collect outcomes in the formal labor market, covering the 2010 to 2019 period and almost 4 million people. Results, for a short unemployment spell, indicate positive effects right after the training conclusion (six months) lasting until the seventh year (between 16% and 20% higher odds to be employed than control in most periods), with similar pattern observed for wages. The share of concluders in a specific formal labor market negatively impacts the employment chances and wages in the short run, suggesting spillover effects due to Pronatec. Also, estimates indicate important heterogeneity, with lower results for women, stronger effects for youth and no effects for self-demand training. Finally, the Program seems to be highly economical, returning R$ 1.89 for each R$ 1 spent.

In the third essay, we estimate the return to technical education in Brazil between 2007 and 2018. A rich administrative record of formal labor market (RAIS) allows us to construct a panel of all formal workers in 2007, following them until 2018. Identification of workers was based on National Catalog of Technical Occupations. Also, we use the two waves (2007 and 2014) of National Household Survey (PNAD), about technical education, to further investigate. Results suggests a positive and significant wage premium – between 21.3% and 24.9% in favor of workers in technical occupation/with technical education, controlling for other available observable variables (schooling, age, firm size, economic sector and job spell). Even restricting our sample to young workers (18 years-old) having high school diploma

in the beginning of period – profile required for a technical occupation – the positive and significant wage premium remains, but in a lower magnitude (between 5.8% and 7.8%). However, cross-section data suggest no wage premium for the young generation of 2014 and, hence, cohorts are affected distinctly by technical education. Taking into consideration the benefits found, technical education seems to be cost-effective if it costs up to R$ 8595.10 monthly (at a 6% annual interest rate).

Keywords: Internet, Elections, Pronatec, Employment, Wages, Technical Education.

Contents

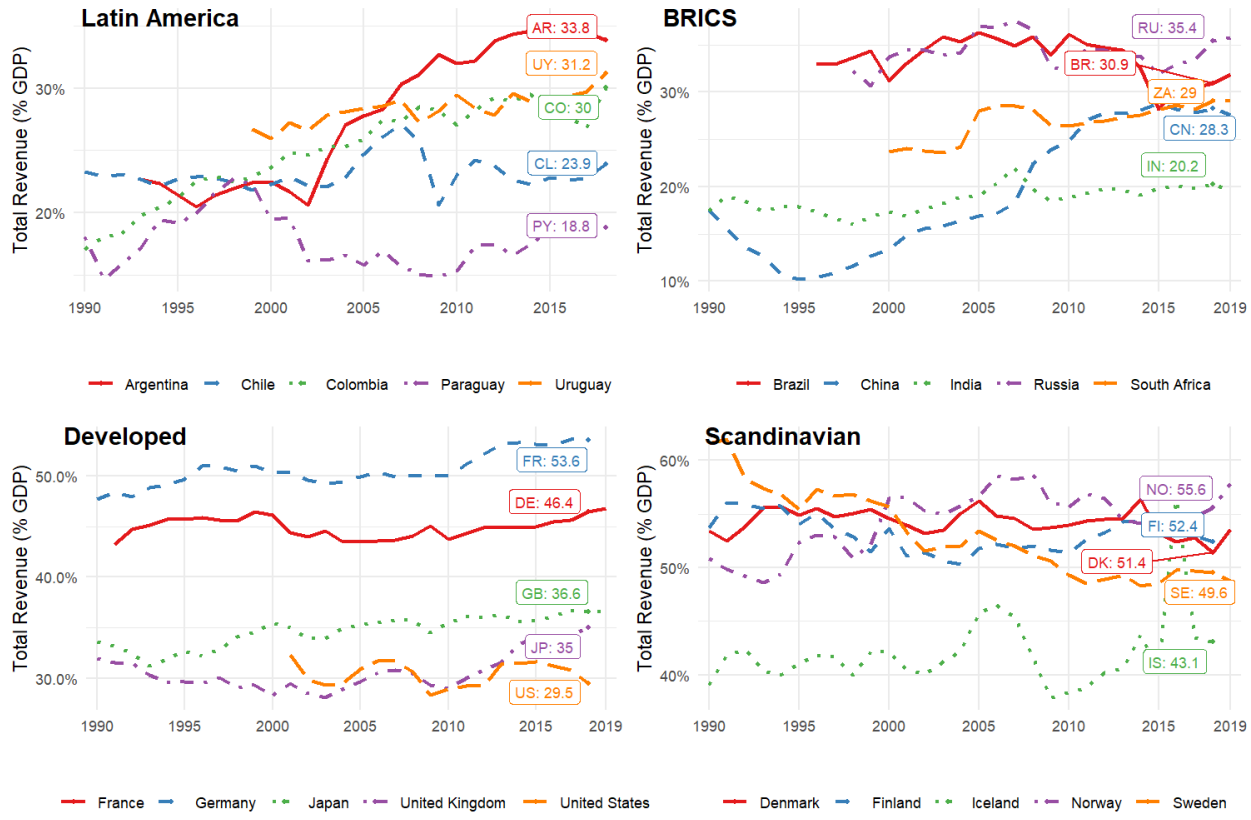
List of Figures

List of Tables

1 Introduction

Brazil has a constitution[1] that puts in the public sector responsibility for several essential services, like universal access to health, education, housing and public security (BOUERI; ROCHA; RODOPOULOS, 2015). To guarantee these services, government undertakes several programs, which requires a revenue compatible to all these activities. Since 1996, in average, the ratio between total government revenue and GDP was over 30%, a figure higher than peers in Latin America (Latam), all the BRIC[2] countries (but Russia), some developed countries (like Japan, United States and United Kingdom), only lower than Scandinavian countries (Figure 1).

Figure 1: Total government revenue over GDP, 1990-2019, selected group of countries



Source: IMF/World Revenue Longitudinal Data (WoRLD)

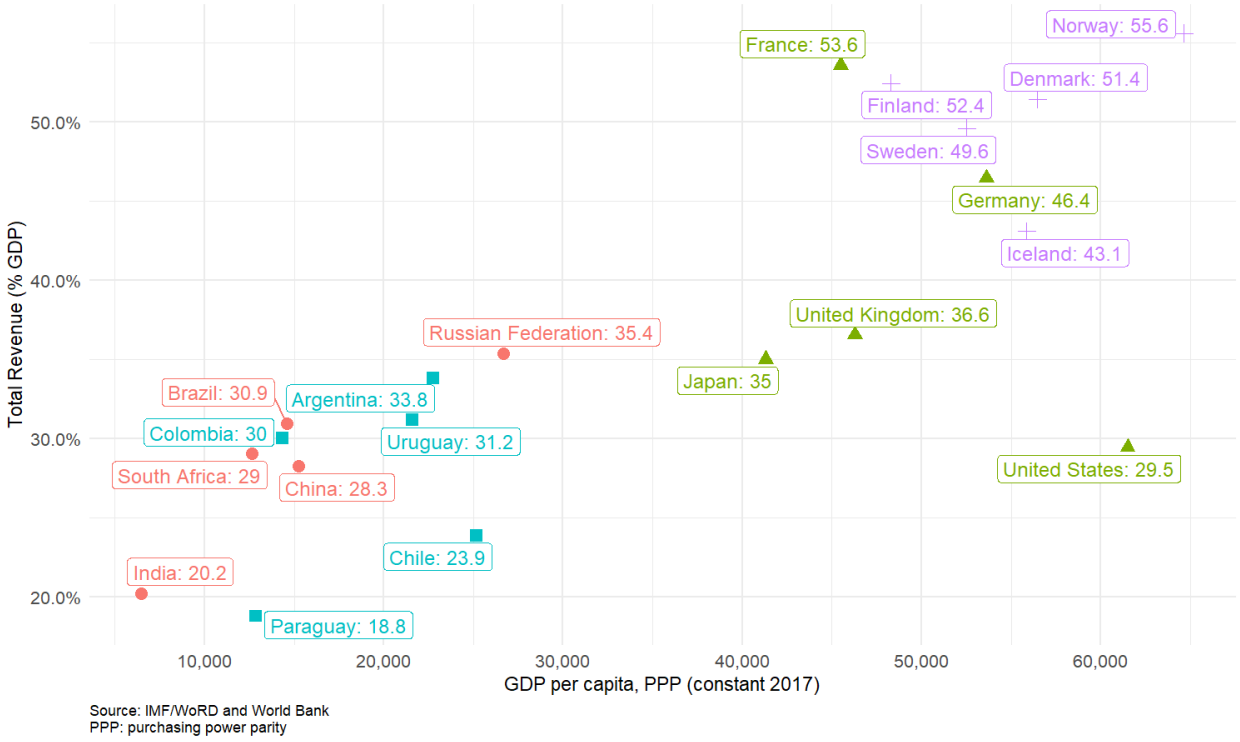
In 2018, Brazil had 30.9% of its GDP as total government revenue, a figure comparable to its neighbors, Argentina, Uruguay and Colombia, but with lower level of GDP *per capita*

[1] Available in: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm

[2] BRICS is the group of main emerging countries – Brazil, Russia, India, China and South Africa. Together, they represent about 42% of world population.

at PPP (purchasing power parity) than the first two[3]. It is also compatible with China, South Africa and United States, while this last one has about four times the Brazil GDP *per capita* (Figure 2).

Figure 2: Total government revenue over GDP by GDP per capita in PPP, 2018, selected group of countries



In sum, with its 1988 constitution, Brazilian government represents nearly one third of its GDP, a relatively large figure for its development level. To fulfill one of the principles of the public administration, the efficiency, program evaluation should be mandatory, even more considering the size of the public sector. Nonetheless, this is not a solid reality in the country.

In the world, there is a growing movement in favor of public policies with evaluation embodied in their designs, including randomized controlled trials (RCT) (RUBIN, 1974) whenever it is feasible. For example, according to ANGRIST (2004), education programs in USA demand scientific base relied upon rigorous design with RCT if possible. However, in Brazil, this culture is quite incipient. According to MENEZES-FILHO (2012), the first studies of this kind have just begun in this century in Brazil and is not settled in Brazilian programs. So, most studies must rely on quasi-experimental or non-experimental techniques

[3]In 2014, the Brazilian GDP growth was only 0.5%, and in the next two years, it felt -3.3% and -3.5%. The 1.3% growth in 2017 and 2018, may explain the recent trend.

to assess the effects of public programs, which is also our case here. In this thesis, we study three different topics, with three independent essays, relying on secondary data bases.

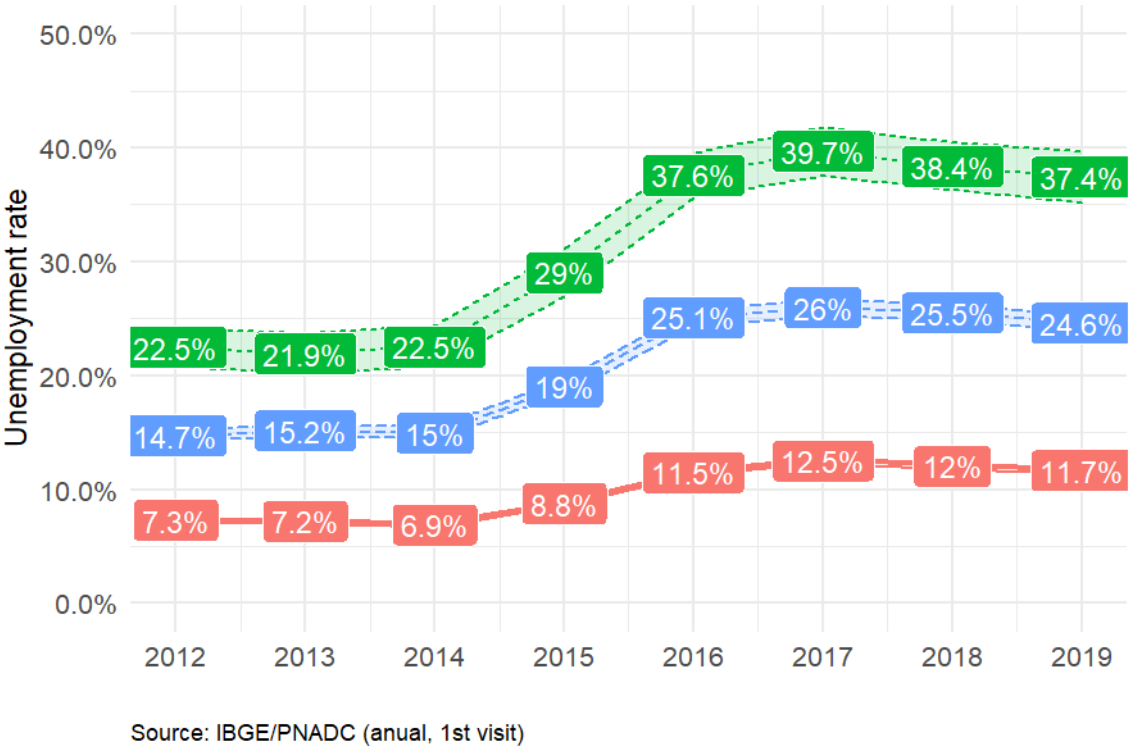
In the first essay, we study the impacts of the Backhaul Program – provision of broadband internet connection in municipalities – on political outcomes. The relationship between media and information acquisition has changed in the last century, from newspaper to radio, television and then internet (BIMBER, 2003). In the beginning of the 21st century, the internet broadband, with high connection velocities, brought possibilities of streaming and the emergency of new social medias. This new scenario made flourish studies trying to investigating the impacts of internet on political outcomes (BIMBER, 1998; CAMPANTE; DURANTE; SOBBRIO, 2017; CZERNICH, 2012; FALCK; GOLD; HEBLICH, 2014; GAVAZZA; NARDOTTO; VALLETTI, 2019; JABER, 2013; LARSON, 2004; POY; SCHÜLLER, 2016). Taking advantage of velocity jumps according to population size, we use a Regression Discontinuity Design (RDD) with multiple cutoffs (CATTANEO; TITIUNIK; VAZQUEZ-BARE, 2018), to assess its relationship with political outcomes (turnout, blank or null percentage votes, left parties vote share, the number of candidates, a small party and young candidates vote share and campaign budget) in 2008, 2010 and 2012 elections. This period covers a national and two local suffrages, and results suggests no relationship between broadband velocity and the outcomes, which diverges from previous studies. It might be the case that institutional background has a role for the emergence of significant results.

In the second essay, we study the impacts of the National Program for Access to Technical Education and Employment (Pronatec), particularly the *Bolsa-formação* arm, a qualification initiative to improve job placement and overall quality of education in Brazil. In short, it is an Active Labor Market Program (ALMP), which is largely investigated in the literature (CARD; KLUVE; WEBER, 2010, 2018; CRÉPON; VAN DEN BERG, 2016; HECKMAN; LALONDE; SMITH, 1999; KLUVE, 2010; VOOREN et al., 2019). We take a step further also analyzing the spillover effects of the program, following FERRACCI; JOLIVET; BERG (2014). The application rules of Pronatec offers a quasi-experimental data base (due to oversubscription or class cancellation), that, joined with administrative registers, allow us to assess the job path in formal labor market of participants and non-participants. Results suggests a positive wage effect of the program (between 16% and 20%), with negative spillover effects in the short run. When the costs are taken into consideration, the program seems to be highly economic, returning R$ 1.89 for each R$ 1 spent.

In the third essay, we study the wage premium of technical education. The youth unemployment (18 to 24 years) rate was almost twice of the general population in 2014 (PNAD), and about three times for people aged 18. Figure 3 shows the unemployment rate

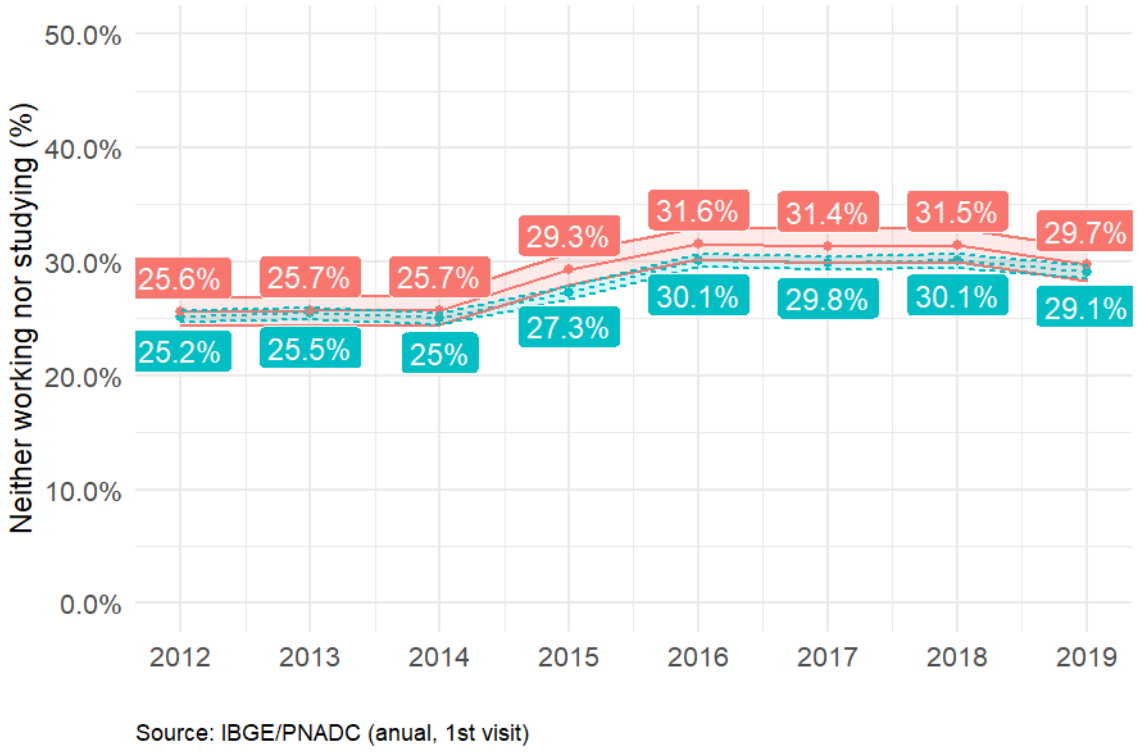
from 2012 to 2019 for all workers and young.

Figure 3: Unemployment rate for all workers, young aged 18-24 and young aged 18, 2012 to 2019, Brazil



We see that, independently of age, situation has worsened in recent years. It was more unfavorable for 18 years-old generation, widening unemployment rate gap. If young people are not studying too, there is an additional problem, once chances to find a job are not increasing through qualification. In Brazil, this figure was around 1/4 of young in 2014, surpassing 30% by 2016. The situation is similar to all young aged 18 to 24 or just with 18 years-old (Figure 4).

Figure 4: Neither working nor studying, young aged 18-24 and young aged 18, 2012 to 2019, Brazil



The school-to-work transition presents many challenges, with several initiatives hoping to alleviate these figures, where technical education is presented as an alternative by offering job skills associated with formal education (FRIGOTTO, 2005). On the other hand, technical education is pointed out with a low rate of social return due to its high unitary costs (PSACHAROPOULOS, 1994), besides not being a popular choice by families to their children when compared to the "safer path" of formal education (BIAVASCHI et al., 2012; KAHYARARA; TEAL, 2008). Using the RAIS data base, from 2007 to 2018, we estimate returns to technical education, identifying all workers that in the initial year had a technical occupation according to National Catalog of Technical Courses. Also, we use two waves of PNAD (2007 and 2014) about technical education to validate our estimates with a different source of information. Our results suggest a wage premium between 21.3% and 24.9% in favor of technical workers, robust to matching data or an alternative random effect model (within-between following BELL; FAIRBROTHER; JONES (2019)). However, when we restrict our sample to young aged 18 in 2007, effects are considerably lower (between 5.8% and 7.8%). The cross-section estimates point to same direction for all workers, but in lower magnitude (around 14% in 2007 and 12% in 2014), with no significant effect for young in

2014. Returns estimates here are more than enough to cover all the ideal costs estimated for technical education, according to ARAÚJO et al. (2016) figures based on the National Plan for Education (with a discount rate up to 18%).

The next three chapters presents each essay, while the last offers the conclusions summarizing our findings.

ESSAYS

2 Broadband connection and election in Brazil: what is role of the internet?

2.1 Introduction

The aim of this paper is to assess the impact of broadband internet velocity on political outcomes: turnout, vote share, types of votes and campaign budget.

The way how people get informed about politics has changed dramatically over the years (BIMBER, 2003). If in XIX century newspaper was the main source of information, in the beginning of XX's radio took its place, surpassed by television in the middle of the same century. Today, a new type of media seems to be taking the lead: the internet.

Although the world wide web is an around 30 years-old technology, broadband connection is an even recent event (MADDUX; JOHNSON, 1997). Internet velocity, capable of streaming videos, became popular just in the XXI century. Social networks, like Facebook, YouTube and Twitter are relatively infant phenomenon[4], becoming popular globally only in the late of 2000's. Mobile broadband connections, thanks to 3G technology (followed by 4G) and massification of smartphones[5], helped internet to reach a greater number of users. New social medias, like WhatsApp, Instagram and Telegram are now everyday tools[6], with popularity increasing in an exponential fashion, being common even for business. A new wave, with 5G technology and the "internet of things" is coming to continue the revolution begun in the end of past century, with connections speed and quality increasing every day, with new possibilities of business.

Thus, information dissemination gained range and speed, reaching more people, almost instantaneously, nearly in any part of the world. Geographic barriers were broken, and the amount of information are vast. Before these new possibilities, a question arises: how this new scenario affects social interaction? Furthermore, how do people are doing politics in this new environment? If people have tools to be more informed, do they increase their participation in elections? Could vote preferences change with introduction of this new technology? Or, on contrary, have these new possibilities of entertainment deviate people from political discussion? Is it possible that broadband internet did not change politics at all?

These questions are not easy to be answered for several reasons. Availability of internet is not random and characteristics like income, schooling and geographic conditions may determine if an internet service provider will be accessible for individuals (CAMPANTE; DURANTE; SOBBRIO, 2017; FALCK; GOLD; HEBLICH, 2014; MINER, 2015). Also,

[4]Facebook was launched in 2004, YouTube in 2005 and Twitter in 2006.
[5]The first Iphone was launched in 2007.
[6]WhatsApp was launched in 2009, Instagram in 2010 and Telegram in 2013.

institutional and political backgrounds may possibly influence internet-political relationship.

Relationship between internet and politics is not a novel issue, being focus of study in several fields (BIMBER, 1998; CAMPANTE; DURANTE; SOBBRIO, 2017; CZERNICH, 2012; FALCK; GOLD; HEBLICH, 2014; GAVAZZA; NARDOTTO; VALLETTI, 2019; JABER, 2013; LARSON, 2004; POY; SCHÜLLER, 2016). This paper aims to contribute with this literature, studying internet velocity impacts on politics in Brazil. Focusing on a single country, we have best tools to control for possible confounders, and, taking advantage of a specific rule for broadband roll out, where number of inhabitants determines the internet velocity of municipalities (the backhaul program), we have a robust identification strategy to deal with internet velocity endogeneity. Following CATTANEO; TITIUNIK; VAZQUEZ-BARE (2018), we use multiple cumulative cutoffs design to estimate effects.

In the most recent Brazilian presidential election, internet had a major role in result. In 2018, Mr. Bolsonaro, with a little fraction of financial resources used by his opponents[7] and with only eight seconds of national advertising time in television[8], managed to go to the second round of presidential elections, with 46.03% of votes, and won elections with a 10 p.p. margin difference. According to Brazilian newspapers, the strength of Mr. Bolsonaro in the social medias was capital to his victory[9], which makes Brazil an interesting case of study regarding the relationship between internet and elections.

We go back a little in time and study if the beginning of broadband internet played a role in election outcomes. Results suggests, in general, that broadband internet speed is not related to political outcome in Brazil. It seems that internet velocity did not influence turnout, blank or null percentage votes, left parties vote share, the number of candidates, a small party and young candidates vote share and campaign budget, in 2008, 2010 and 2012 elections, which covers a national and two local suffrage. The offices considered (president, mayor, deputies, or local legislators) did not make difference in results. These finds are different from previous results reported in the literature, meaning that institutional background or local idiosyncrasies may play an important role in studies relating politics and internet. Positive and negative relationship are reported for Germany, Italy and United Kingdom

[7]While Mr. Bolsonaro expended R$ 2.46 million in his campaign, the second place, Mr. Haddad, expended R$ 37.5 million, a figure 15 times higher. Complete figures are available in http://divulgacandcontas.tse.jus.br/.

[8]https://agenciabrasil.ebc.com.br/politica/noticia/2018-08/tse-apresenta-tempos-de-radio-e-tv-de-presidenciaveis

[9]https://www.correiobraziliense.com.br/app/noticia/politica/2018/10/28/interna_politica,715584/bolsonaro-fez-das-redes-sociais-o-caminho-certo-para-uma-provavel-vito.shtml, https://g1.globo.com/politica/blog/cristiana-lobo/post/2018/12/31/redes-sociais-mudam-completamente-a-relacao-dos-eleitores-com-seus-representantes.ghtml, https://noticias.uol.com.br/politica/eleicoes/2018/noticias/2018/10/09/como-midias-sociais-e-orcamentos-enxutos-derrubaram-cinco-mitos-eleitorais.htm

(CAMPANTE; DURANTE; SOBBRIO, 2017; FALCK; GOLD; HEBLICH, 2014; GAVAZZA; NARDOTTO; VALLETTI, 2019), all of them with distinct political institutional background compared with Brazilian's, as well developing level.

This paper is organized as follows: the first section presents the theoretical framework linking internet to political outcomes, while the next reports the previous findings regarding its application. The third section reviews the Brazilian institutional political background, followed by the section with empirical strategy, databases and descriptive statistics. The fifth section presents our results, with a final discussion in the last section.

2.2 Theoretical framework

There are some theories looking to explain why people vote (ALDRICH, 1993; DOWNS, 1957; FEREJOHN; FIORINA, 1974; RIKER; ORDESHOOK, 1968; UHLANER, 1989). One approach is to treat as a microeconomic problem in the following way. In elections, individuals' problem is to choose the best candidate(s) according to their preferences. But there is an asymmetry of information: there are many candidates (not considering uncontested elections), and voters are not fully informed about their abilities. Acquire information about them is costly, since they must spend resources to consume information (e.g., from television, radio, newspaper, internet or another people), that may include money and time. Show up to cast the ballot also requires resources (transportation and time, for example). More accurate decision requires more information, which demands more resources, i.e., is more costly. So, it can be viewed as a maximization problem from the microeconomics point of view, which can be solved by equalizing marginal costs and benefits. Benefits can be viewed as the policies the most preferred candidate will conduct, a civil duty or being party of the democratic process (ALI; LIN, 2013).

This problem changes over time with entrance of new technologies (GENTZKOW, 2006). For example, when radio, television and internet were not available, there were fewer options to people get informed about candidates. Also, there were available less leisure alternatives. With emergence of radio, then television and, finally, internet, costs and substitution effects may have changed. A first natural question is: did these new technologies affect the decision of voters? For newspaper, GERBER; KARLAN; BERGAN (2009), GENTZKOW; SHAPIRO; SINKINSON (2011) and DRAGO; NANNICINI; SOBBRIO (2014) report effects on elections participation. According to STRÖMBERG (2004) and HORACIO; MONTEIRO (2014), radio affects people perception about politics, while DELLAVIGNA; KAPLAN (2007), ENIKOLOPOV; PETROVA; ZHURAVSKAYA (2011), DURANTE; KNIGHT (2012), GENTZKOW (2006) and OBERHOLZER-GEE; WALDFOGEL (2009) show the impact of

television (through news) on elections results.

In Brazilian case, FERRAZ; FINAN (2008) show that, by making publicly available audit of expenditures and transferred funds, results in elections has changed, in 2004, for incumbents. Results were stronger in municipalities with local radio stations (and, so, not only broadcasting national or regional news). According to the authors, more informed electorate, in association with local media, had an important role in local elections.

How about the internet? Relationship between internet and politics has been investigated since the end of 1990's (BIMBER, 1998). The effect on information acquisition may be ambiguous depending on the hypothesis used: if internet makes available new possibilities of entertainment, people may substitute the time spent learning about politics with new types of leisure; on the other hand, if internet bring to people new sources of political information and channels of discussion, people may be pushed toward politics. Finally, the cost and the time necessary to find candidates information or to find new possibilities of entertainment may have changed relative prices. Once someone has access to internet, it is possible to consume a variety of information with, in general, no additional cost (except time). The same is valid to leisure. A last possibility is that technology is the only thing substituted to consume information and leisure, making no difference in resources allocation at all[10].

Changes may also take time to happen. Many types of media on internet depends on broadband connection (like video streaming), only available to the large public in the beginning of the XXI century. Moreover, all content we have today were not available with the launch of the internet. The same was true for television, where the diversity of programs and shows existing today took time to be developed and aired. Emergency of new technologies and its spread also affects relative prices both for information and leisure over time.

While newspaper, radio and TV content production are more restricted and with barrier entries, internet have opened doors to virtually anyone (connected) produce information and media, interact with people and organize groups of common interest, everything at a lower cost and time. Thus, it is likely to exist a shift both in the demand and supply of information and entertainment with internet arrival. It can potentially alter the manner of how politics are made, since, with internet, politicians can reach more people, quickly and at lower costs when compared with other medias.

One situation this new scenario brings is the social media consumption of "fake news"[11]

[10]If there is no, or little, consumption of political information with an older technology, it might be the case that, even with a new technology, there is no preference for this type of information, resulting in no, or limited, shifting in its demand.

[11]Fake news is a popular term to define, in general, the spread of misleading or false information like if it

and its possible impact on elections. In the problem treated here, misleading information may have a market that deviate people from optimal choice (see ALLCOTT; GENTZKOW, 2017 for a theoretical framework). Media capture by politicians put an additional flavor to this discussion (BESLEY; PRAT, 2006), where internet could break other types of media control or enhance an existing control.

With this framework in mind, we analyze previous research in the field to collect results and identification strategies, pointing resemblances and contrasts between them. Common outcomes between internet and politics relationship are voting turnout, election results, public polices and politician's accountability.

2.3 Literature

Sources from where people consume information and leisure are not exogenous. For example, if television or internet is expensive, only people with enough income can have access. If this kind of people have preferences regarding candidates, then there is a bias if relationship between internet and politic outcomes is treated as unconditional. The same is valid for another characteristics, like race, schooling, age or housing location.

Due to endogeneity of internet supply and demand, geographical characteristics (e.g., landscape or rainfall) or previous telecommunication infrastructure are common strategies to instrumentalize internet in order to link it to political outcomes. CAMPANTE; DURANTE; SOBBRIO (2017) study the impact of broadband diffusion on political participation for municipalities of Italy between 1996 and 2013 with this strategy. MINER (2015) take similar path for Malaysia, CZERNICH (2012) and FALCK; GOLD; HEBLICH (2014) for Germany, GAVAZZA; NARDOTTO; VALLETTI (2019) for UK, JABER (2013) for USA and MENEZES (2015) for Brazil. With slightly different approach, LELKES; SOOD; IYENGAR (2017) explore variation in state laws related to internet infrastructure to study influence of this technology on polarization in USA, while POY; SCHÜLLER (2016) use similar strategy to analyze broadband effects on turnout and vote share in rural and sparse areas in Italy.

For Italy, CAMPANTE; DURANTE; SOBBRIO (2017) report a negative effect on turnout in elections following high speed internet implantation (2008), changing its direction for later elections (2013). An interesting result in Italian case is that internet affected ideological groups distinctly, according to vote share results, paving the way for organization of new political groups in online platforms. POY; SCHÜLLER (2016) echoes these results, linking high speed internet (ADSL2+) to increases in turnout in 2008 and 2013 Italian elections, as

were real. See LAZER et al. (2018) for a brief discussion.

well transitory increases in vote share of some parties (center-left and right-fringe).

In Malaysian case, MINER (2015) reports important effects of internet on 2008 election results (vote share of opposition parties), but not on turnout and limited effects on turnover. Although the identification strategy is similar, the political background for the Malaysian case is different from the Italian.

A negative effect of internet on turnout is reported by FALCK; GOLD; HEBLICH (2014) for Germany. The mechanism is related to an increase in leisure consumption that crowds out television entertainment, since internet can be viewed as a substitute of this kind of consumption[12]. The impact reported is heterogeneous: west Germany was affected, while in east Germany no effect was observed. Effects on vote shares were not observed in neither place. On the other hand, CZERNICH (2012) found positive effects on participation in German 2002-2005 election.

GAVAZZA; NARDOTTO; VALLETTI (2019) report for UK negative effects of internet on turnout in 2006-2010 elections, with stronger results for less-educated and younger voters. Furthermore, incumbents seem to take advantage, diminishing election competitiveness. Taking a step further, the UK study suggests effects on public policies, lowering public expenses and taxes in areas with higher internet access (with similar heterogeneity effects reported for turnout).

In Brazilian case, MENEZES (2015) shows the association between internet and increasing in vote share of small candidates in 2010 elections, but no relationship with turnout nor with blank votes. This is an important result once the winner of last Brazilian presidential election (2018) had extremely limited advertisement time on radio and television in the first round.

For USA, LELKES; SOOD; IYENGAR (2017) bring light to mechanisms underlying the effects of internet on political outcomes. States with less restrictive laws (and more likely to have broadband coverage) induce people to be exposed to partisan information and be more extreme in partisan preferences. This mechanism is compatible with JABER (2013) results, who reports a positive impact on turnout, donations to political campaigns and democrats vote share in 2008 presidential elections. In an early study, with weaker identification strategy, TOLBERT; MCNEAL (2003) suggests that, in 1996 and 2000 presidential election, individuals with internet and online elections news reading are more likely to vote.

It is important to notice that countries have distinct political regimes, which could potentially affect results. MINARD; LANDRIAULT (2015) bring this to discussion analyzing

[12]If we consider that people have a fixed amount of time to enjoy leisure activities, internet enters as a new option to compete with television, potentially reducing the time spent with the latter.

how maturity of democracy regimes in Asia responds to internet availability. Immature regimes seem to be more affected by internet than solid democracies according to 2006 cross-country analysis. Hence, the between countries analysis suggests that there are institutional factors playing action on internet-politics relationship, which puts caution to external validity of single country studies.

To sum up, it is clear that there are different results for different countries (even inside the same country), with possible changing effects over time. Also, most studies are concentrated in 2000-decade elections, focusing on the begging of the broadband internet. Few studies report results for elections held in 2010 decade, when smartphone revolution and social media gained strength. Even more, there are no studies about the effects of mobile broadband and smartphones on elections.

In this paper we will address fixed line broadband roll out, studying the Brazilian case, one of the largest democracies in the world. As pointed before, peculiarities of each country seem to be determinant for results, which demands closer analysis of the political system to compare our results with those presented before.

2.4 Brazilian political institutional background

Brazil is a Federal Republic, with three layers of government: central (or Federal), states and municipalities[13] (see SOUZA (2005) for a discussion about the federalism in Brazil). It is a young presidential democracy[14], with bicameral legislative system (Chamber of Deputies and Senate, the National Congress), holding election every four years. President is elected by direct vote since 1989 in national elections, as well national congress, state governors and state assemblies (1994 onward). Local elections, for municipal mayors and local legislators are also held every four years, since 1996[15]. While mayors, senators and the president are elected in a majoritarian system, all the other candidates are elected by proportional representation,

[13]Brazil was under Portugal's control from 1500 to 1822, when its independence was declared and Dom Pedro, son of Portugal's king, became the ruler.

[14]Brazilian Republic, proclaimed in 1889, was initially ruled by military and then by *São Paulo*, *Rio de Janeiro* and *Minas Gerais* oligarchies alternating power until 1930. In 1930, *Getúlio Vargas* took the power until 1945, and, after that, the country had free elections until 1960.

[15]Brazilian dictatorship begun in 1964 and ended in 1985, with general election in 1986, except for president (elected indirectly in the previous year). Before 1985, all other elections (except for president) had direct vote, but under military rules. In 1988, Brazil promulgated a new constitution and, in 1989, elected the president by direct vote again, after 29 years. In 1990, there were elections for state governors, state assemblies and national congress. In 1992, municipal mayors and local assembly members were elected. By 1994 onward, national elections (president, state governors, state assembly and national congress) happens every four years, while local elections (municipal mayor and municipal assembly) happen every four years, since 1996. Thus, Brazil has elections every two years since 1994.

where voters choose first a party and then a candidate[16]. Also, parties, until 2018, could create coalition[17] to run in proportional elections, while in majoritarian elections, coalitions are (and still) permitted. With this system, in 2018, 35 parties ran in the elections. Table 1 presents all parties and the number of candidates in each election from 2000 to 2018.

Considering the large number of parties in Brazil, to make the vote share analysis manageable, we organize parties as left, center or right orientation, based on POWER; ZUCCO JR (2012) party index[18]. Table 2 presents this organization.

The party index has some aggregation of parties as "others", so another classification criterion was necessary. Parties web pages were consulted to analyze their history and beliefs to designate parties to the groups. This methodology may arise questions if some parties labeled as right, are actually centrists. To avoid this issue, we focus on left parties vote shares in results section, since their classification are more direct and mostly based on the party index.

The relatively large number of parties in Brazilian makes both elections and politics complex processes (BOULDING; BROWN, 2015; PETTERSSON-LIDBOM, 2008). In order to help understanding this process, Table 3 shows the winners, by party, in the last five elections for the national congress, while Table 4 shows the same information for state governors and municipalities[19].

Table 3 shows that no party had more than 20% of the deputies. The workers party (PT), who won four out of five last presidential elections (2002 to 2014), did not have most of the congress in any year (at least, without building a block of parties). In the Senate, the party with more seats had around 1/4 of the house until 2010, with more competitive elections since then, specially in 2018. So, to run the country, the president needs to build alliances, otherwise it is unlikely to pass its bills. The number of parties necessary to form at least 60% of the deputies[20] has risen from at least four to eight, including both situation and opposition parties. This means that the necessary number of parties required to rule is even higher. In the senate, the number has risen from three to seven.

[16]There is the option to vote only for a party.

[17]Altered by the Constitutional Amendment 45, available in http://www.planalto.gov.br/ccivil_03/constituicao/Emendas/Emc/emc97.htm

[18]The authors construct a party index based on legislative surveys from 1990 to 2009, taking into consideration the ideological position of congress members in their activities.

[19]Since 1988, Brazil has 26 states and the Federal District. In 2018, there were 5,568 municipalities, with two districts, the Federal capital and the district of *Fernando de Noronha*, in *Pernambuco*. National Congress has 513 Federal Deputies and 81 Senators

[20]Bills that alter constitution require at least 308 votes of deputies and 49 votes of senators, i.e. 60% of the National Congress, in two rounds.

Table 1: Parties and number of candidates in Brazilian elections, 2000-2018

Party	2000	2002	2004	2006	2008	2010	2012	2014	2016	2018
NOVO									137	390
PAN	1,382	393	3,030	465						
PC do B	1,978	181	4,499	338	7,119	757	12,200	752	11,173	745
PCB	171	19	453	87	666	73	395	97	203	60
PCO	60	112	313	95	31	9	11	23	45	43
PDT	24,465	996	22,272	1,189	22,334	909	25,317	900	23,960	853
PEN/PATRIOTA								780	9,366	1,077
PFL/DEM	42,481	803	32,644	848	25,346	703	21,139	537	20,023	622
PGT	1,465	504								
PMB									4,082	386
PMDB/MDB	49,231	1,112	40,331	1,192	39,377	1,085	42,266	1,089	40,754	977
PMN	4,901	325	6,538	565	6,034	606	7,142	466	6,733	655
PPB/PP	33,177	805	27,613	622	24,837	743	28,086	664	25,775	691
PPL							1,896	388	3,265	522
PPS/CIDADANIA	19,388	823	21,159	937	15,748	724	16,698	556	15,408	593
PR/PL	19,551	920	25,101	701	19,757	636	20,913	689	20,791	658
PRB/REPUBLICANOS				80	8,610	516	12,764	651	16,526	793
PRN/PTC	1,194	164	4,893	449	4,669	741	7,109	649	8,058	724
PRONA	1,284	357	2,595	460						
PROS								395	10,093	1,000
PRP	4,607	313	6,053	489	5,048	492	7,564	762	7,853	864
PRTB	2,689	409	4,182	363	3,767	479	5,928	584	5,954	847
PSB	15,599	1,139	16,649	1,035	19,612	999	24,588	1,147	24,786	841
PSC	8,221	568	8,803	690	10,843	770	15,202	822	15,148	791
PSD	10,308	363					22,414	596	27,066	629
PSDB	38,131	964	33,810	1,040	30,675	976	33,254	951	32,843	829
PSDC/DC	3,532	285	6,141	477	4,910	338	6,857	627	6,923	655
PSL	5,198	340	6,522	408	6,131	661	9,349	678	9,662	1,334
PSN/PHS	2,863	338	5,756	568	5,160	515	7,805	808	10,926	894
PSOL				535	2,679	771	4,262	1,056	4,479	1,170
PST	4,639	471								
PSTU	347	210	546	102	213	108	272	234	258	133
PT	25,829	1,495	36,600	1,144	31,765	1,243	40,022	1,190	21,842	1,126
PT do B/AVANTE	3,103	389	4,503	492	4,472	516	7,155	653	6,842	942
PTB	30,268	863	25,919	802	22,833	894	23,819	813	20,653	584
PTN/PODEMOS	1,919	223	4,353	299	4,473	444	6,775	528	8,952	844
PV	6,021	609	10,887	970	13,322	1,154	17,464	918	15,619	811
REDE									3,435	762
SD/SOLIDARIEDADE								465	13,711	723
Total candidates	364,002	16,493	362,165	17,442	340,431	17,862	428,666	21,468	453,344	25,568
Total parties	30	30	27	29	27	27	29	32	35	35

Source: TSE

Obs.: Parties with changes in their names are considered as a unique party.

Competitiveness observed in National Congress elections is also present in subnational suffrage. Table 4 shows that most dominant party had around 20% of mayors and 25% of governors, not necessary the same party in each election, adding another layer of complexity in the Brazilian Federal System (see CANCELA; GEYS (2016) for a discussion about coordination in multilevel elections in Brazil).

Table 2: Party classification according to orientation (left, center or right)

Left	Center	Right
PC do B	NOVO	PEN/PATRIOTA
PCO	PAN	PFL/DEM
PSB	PCB	PMN
PSOL	PDT	PPB/PP
PSTU	PGT	PRN/PTC
PT	PMB	PRTB
	PMDB/MDB	PSDC/DC
	PPL	PSL
	PPS/CIDADANIA	
	PR/PL	
	PRB/REPUBLICANOS	
	PRONA	
	PROS	
	PRP	
	PSC	
	PSD	
	PSDB	
	PSN/PHS	
	PST	
	PT do B/AVANTE	
	PTB	
	PTN/PODEMOS	
	PV	
	REDE	
	SD/SOLIDARIEDADE	

Obs.: Division of parties based on quantiles of party index (0.25, 0.75, 1)

Obs.2: Parties out of party index were allocated based on party description available on their internet page.

Until 2017, parties had the Party Fund[21] and private donations (since 2016, companies are not allowed to donate for elections[22]), which includes own resources from candidates. There are maximum values allowed to be expended by candidates in campaign, stipulated each year by the Supreme Electoral Court (TSE - *Supremo Tribunal Eleitoral*). The Party Fund is distributed by the following rule: 5% is equally given to registered parties[23] and 95% according to votes won in the last deputies' elections. In 2017, with prohibition of companies

[21]The *Fundo Partidário*, created by the Law 9,096/1995, available in http://www.tse.jus.br/legislacao/codigo-eleitoral/lei-dos-partidos-politicos/lei-dos-partidos-politicos-lei-nb0-9.096-de-19-de-setembro-de-1995.

[22]According to a Supreme Court (STF) decision, ADI 4,650/2015 and the Law 13,165/2015. Available in http://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=10329542 and http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2015/Lei/L13165.htm

[23]In 2017, the Constitution Amend 33 created the rules to access the Party Fund, so called *Cláusula de Barreira* (Barrier Clause). In 2019, the party must havr 1.5% of valid votes for deputies in 2018, distributed at least in 1/3 of the states and with at least 1% of the votes in each one, or had nine deputies in at least 1/3 of the states. This rule will be more rigid in 2023, with 2% of the votes or 11 deputies, under the same rules. In 2027, the figures will be 2.5% (and at least 1.5% in 1/3 of the states) or 13 deputies and in 2031, they will be 3% (and at least 2% in 1/3 of the state) or 15 deputies.

Table 3: Distribution of winners, by party, in National Congress, 2002-2018

	2002		2006		2010		2014		2018	
Party	Deputy	Senator	Deputy	Senator	Deputy	Senator	Deputy	Senator	Deputy	Senator
NOVO	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.6	0.0
PAN	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PC do B	2.3	0.0	2.5	3.7	2.9	1.9	2.0	0.0	1.8	0.0
PDT	4.1	7.4	4.7	3.7	5.3	3.7	3.9	14.8	5.5	3.7
PEN/PATRIOTA	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.0	1.0	0.0
PFL/DEM	16.4	25.9	12.7	22.2	8.4	3.7	4.1	11.1	5.7	7.4
PMB	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PMDB/MDB	14.8	16.7	17.4	14.8	15.2	25.9	12.7	18.5	6.6	13.0
PMN	0.2	0.0	0.6	0.0	0.8	1.9	0.6	0.0	0.6	0.0
PPB/PP	9.4	0.0	8.0	3.7	8.6	7.4	7.4	3.7	7.2	9.3
PPL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0
PPS/CIDADANIA	2.9	1.9	4.3	3.7	2.3	1.9	2.0	0.0	1.6	3.7
PR/PL	5.1	3.7	4.5	3.7	8.0	7.4	6.6	3.7	6.4	1.9
PRB/REPUBLICANOS	0.0	0.0	0.2	0.0	1.6	1.9	4.1	0.0	5.8	1.9
PRN/PTC	0.0	0.0	0.6	0.0	0.2	0.0	0.4	0.0	0.4	0.0
PRONA	1.2	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PROS	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	1.6	1.9
PRP	0.0	0.0	0.0	0.0	0.4	0.0	0.6	0.0	0.8	1.9
PRTB	0.0	0.0	0.0	3.7	0.4	0.0	0.2	0.0	0.0	0.0
PSB	4.3	5.6	5.3	3.7	6.8	7.4	6.6	11.1	6.2	3.7
PSC	0.2	0.0	1.8	0.0	3.3	1.9	2.5	0.0	1.6	1.9
PSD	0.8	1.9	0.0	0.0	0.0	0.0	7.0	7.4	6.6	7.4
PSDB	13.7	14.8	12.9	18.5	10.5	11.1	10.5	14.8	5.7	7.4
PSDC/DC	0.2	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.2	0.0
PSL	0.2	0.0	0.0	0.0	0.2	0.0	0.2	0.0	10.1	7.4
PSN/PHS	0.0	0.0	0.4	0.0	0.4	0.0	1.0	0.0	1.2	3.7
PSOL	0.0	0.0	0.6	0.0	0.6	1.9	1.0	0.0	2.0	0.0
PST	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PT	17.7	18.5	16.2	7.4	16.8	20.4	13.4	7.4	10.9	7.4
PT do B/AVANTE	0.0	0.0	0.2	0.0	0.6	0.0	0.2	0.0	1.4	0.0
PTB	5.1	3.7	4.3	11.1	4.3	1.9	4.9	7.4	2.0	3.7
PTN/PODEMOS	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.0	2.1	1.9
PV	1.0	0.0	2.5	0.0	2.5	0.0	1.6	0.0	0.8	0.0
REDE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	9.3
SD/SOLIDARIEDADE	0.0	0.0	0.0	0.0	0.0	0.0	2.9	0.0	2.5	1.9

Source: TSE

Obs.: Parties with changes in their names are considered as a unique party.

donation, another fund was created, the Electoral Fund[24], which is distributed with following rule: 2% equally between registered parties; 35% to parties with, at least, one deputy; 48% according to deputy's proportion; and 15% according to senators' proportion. For example, in 2018, the Fund Party was R$ 888.7 million (US$ 175 million), while the Electoral Fund was R$ 1.7 billion (US$ 334.1 million).

Another important aspect of Brazilian suffrage regards campaign advertisement. There

[24]Called *Fundo Especial de Financiamento de Campanha*, it was created by the Laws 13,487/2017 and 13,488/2017. Available in http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2017/Lei/L13487.htm and http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2017/lei/L13488.htm.

Table 4: Distribution of winners by party in local executive elections, 2002-2018

	2000/2002		2004/2006		2008/2010		2012/2014		2016/2018	
Party	Mayor	Gov.	Mayor	Gov.	Mayor	Gov.	Mayor	Gov.	Mayor	Gov.
NOVO	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.7
PAN	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PC do B	0.0	0.0	0.2	0.0	0.7	0.0	1.0	3.6	1.5	3.7
PDT	5.4	3.7	5.5	7.4	6.3	0.0	5.5	7.1	6.0	3.7
PEN/PATRIOTA	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0
PFL/DEM	18.6	14.8	14.2	3.7	8.9	7.4	5.0	0.0	4.9	7.4
PMB	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0
PMDB/MDB	22.2	18.5	19.0	25.9	21.7	18.5	18.4	25.0	18.9	11.1
PMN	0.3	0.0	0.6	0.0	0.8	3.7	0.7	0.0	0.5	0.0
PPB/PP	11.0	0.0	9.9	3.7	9.9	0.0	8.6	3.6	9.0	3.7
PPL	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.1	0.0
PPS/CIDADANIA	3.0	7.4	5.5	7.4	2.3	0.0	2.2	0.0	2.2	0.0
PR/PL	4.2	0.0	6.8	0.0	6.9	0.0	5.0	0.0	5.4	0.0
PRB/REPUBLICANOS	0.0	0.0	0.0	0.0	1.0	0.0	1.4	0.0	1.9	0.0
PRN/PTC	0.1	0.0	0.3	0.0	0.2	0.0	0.4	0.0	0.3	0.0
PRONA	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PROS	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.6	0.9	0.0
PRP	0.3	0.0	0.7	0.0	0.3	0.0	0.4	0.0	0.3	0.0
PRTB	0.1	0.0	0.2	0.0	0.2	0.0	0.3	0.0	0.2	0.0
PSB	2.6	14.8	3.1	11.1	5.6	22.2	7.9	10.7	7.4	11.1
PSC	0.5	0.0	0.4	0.0	1.0	0.0	1.5	0.0	1.6	7.4
PSD	1.9	0.0	0.0	0.0	0.0	0.0	8.9	7.1	9.7	7.4
PSDB	17.6	25.9	15.7	22.2	14.3	29.6	12.5	21.4	14.5	11.1
PSDC/DC	0.1	0.0	0.2	0.0	0.1	0.0	0.2	0.0	0.1	0.0
PSL	0.5	3.7	0.4	0.0	0.3	0.0	0.4	0.0	0.5	11.1
PSN/PHS	0.1	0.0	0.4	0.0	0.2	0.0	0.3	0.0	0.7	3.7
PSOL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PST	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PT	3.5	11.1	7.4	18.5	10.0	18.5	11.5	17.9	4.6	14.8
PT do B/AVANTE	0.1	0.0	0.4	0.0	0.1	0.0	0.4	0.0	0.2	0.0
PTB	7.5	0.0	7.6	0.0	7.4	0.0	5.3	0.0	4.6	0.0
PTN/PODEMOS	0.0	0.0	0.1	0.0	0.3	0.0	0.2	0.0	0.5	0.0
PV	0.2	0.0	1.0	0.0	1.4	0.0	1.8	0.0	1.8	0.0
REDE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0
SD/SOLIDARIEDADE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0

Source: TSE

Obs.: Parties that changed their names are considered as a unique party.

are national programs, free of charges and mandatory, during campaign time, both aired daily in radio and television, broadcasting the same content in all regions of the country. There is a fixed amount of time for electoral advertisement in these channels, 2/3 distributed according to current party presence in legislatures and 1/3 among candidates[25], and only this time is allowed to be used in these channels. Ads on newspaper are also restricted, even though being a less important media compared to TV and radio. Internet is exception, where candidates can use it, almost freely, to reach voters, since 2009, except for anonymously or paid advertisement (which includes social medias like Twitter, Facebook, Instagram and

[25]The same rules of the *Cláusula de Barreira* (Barrier Clause) is also applied here.

YouTube)[26].

So, all this set of rules concentrate resources for some parties and candidates, making the internet an important alternative in elections. As aforementioned, in the presidential election of 2018, internet was pointed out as crucial for the outcome.

Looking now to electorate, in Brazil, voting is mandatory to literate citizens aged 18 to 69. For people aged 16 to 17 and 70 and over, voting is optional. Voters absent in election must justify or pay a small fine. If they fail to justify three consecutive polls, voter registration is canceled and some rights are lost (issue or renew passports and national identification, receive wages as public servant or from any institution linked to government, participate in public competition for resources, request loans from institutions held by the government, apply to jobs as public servant, enroll in public education or engage in any public act that requests military service or income task discharge). This set raises question if this rule changes incentives to acquire information about politicians and participate in elections, specially by poorer population. Table 5 shows the total number of voters as well total population with voter registration canceled from 2005 to 2019.

Table 5: Total of voters and voter registration canceled, 2005 to 2009

Year	Voters	Turnout	Turnout %	Canceled	Canceled %
2005	121,391,631	102,526,992	84.46	1,089,662	0.89
2007	125,913,494	104,820,459	83.25	1,652,565	1.30
2009	130,604,430	110,085,172	84.29	553,406	0.42
2011	135,804,433	111,193,747	81.88	1,400,549	1.03
2013	140,646,483	115,807,514	82.34	1,358,901	0.96
2015	142,822,083	115,122,883	80.61	1,717,425	1.20
2017	146,470,911	118,757,780	81.08	1,862,665	1.27
2019	147,306,275	117,364,654	79.67	2,491,271	1.69

Source: TSE

Obs.1: Includes voters registered in Brazil and abroad.

Obs.2: Turnout in the last elections.

Despite the mandatory voting rule, turnout was 82.2% in average from 2004 to 2018, and only 1.09% of voters registration were canceled, which means that more than 90% of those who did not turnout took actions to regularize their electoral obligations. Anyhow, turnout is high when compared to USA or European countries.

According to DOWNS (1957), low probability to be pivotal in elections explain the "rational ignorance" of voters and low preference to turnout. On the other hand, mandatory vote could change these incentives, making people more inclined to vote (LIJPHART, 1997).

[26]There is a set of other rules stipulated by the Supreme Electoral Court in each election, like size of advertisement material, schedule for rallies etc.

LEON; LEITE; RIZZI (2014) finds that, for Brazilian case, mandatory voting seems not change people incentives to be more informed in voting decision. It seems that providing more information about candidates (BANERJEE et al., 2011), hence lowering the costs for collect information, is more effective than compulsory voting system.

Following FUJIWARA; MENG; VOGL (2016), we also consider possible persistent habits on voting pattern, incorporating raining information in election days in each municipality. The authors find that rainfall, both in current and past election day, reduces turnout and may possibly affect consumption value of voting. Considering the continental size of Brazil, it may be an important contribution for analysis.

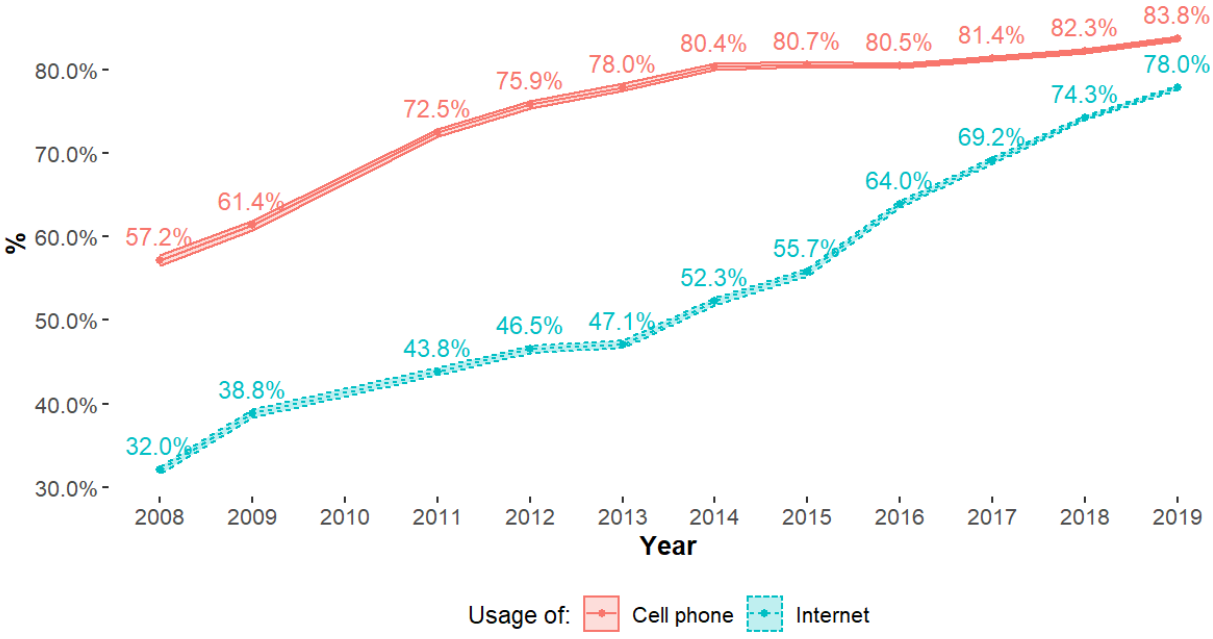
2.5 Empirical strategy and databases

In this section we describe in detail the empirical strategy, relied on the Backhaul program rules, the databases that support the analysis and the cumulative RDD set up.

2.5.1 Communication usage

As a glimpse of Brazilian communication consumption, Figure 5 presents internet and cell phone usage from 2008 to 2019.

Figure 5: Internet and cell phone usage in Brazil, % of 16+ years-old population, 2008-2009 and 2011-2019



Source: National Household Survey (PNAD e PNADc)
Obs.1: Usage of internet in the last 3 months (survey date month as reference)
Obs.2: Cell phone ownership for personal use
Obs.3: No value for 2010 (Census year)
Obs.4: Area below and above points is the confidence interval (95%).

In 2008, around 1/3 of Brazilians (16 years-old or above, i.e., population in voting age) used internet at least once in the previous three months (September as reference), while almost 58% declared cell phone ownership for personal usage. To increase these figures, the government carried out a national plan, in the begging of 2008. In 2011, these figures rose to 44% and 73%, respectively, indicating an increasing communication market in Brazil. Even in 2019, there is room remaining for internet and cell phone expansion in the country (around 22% and 16%, respectively).

Hence, this expressive change in communication consumption may have changed how Brazilians' face politics, possibly increasing opportunities for information acquisition and social interaction about this matter. Or, on the other hand, widening leisure alternatives and lowering politics information consumption.

2.5.2 Backhaul Program (National Broadband Plan)

In April 2008, the presidential Decree 6,424 changed the former National Plan of Goals for Public Switched Telephone (PST) Network Universalization, adding broadband infrastructure

as mandatory (in exchange of the PST obligation). The infrastructure mentioned in the Decree was the Backhaul, a requirement for internet implementation in the country. Backhauls are necessary to connect them to the Telephone Companies' Backbones. The plan put as target that, at least, 40% of municipalities should have the necessary infrastructure by the end of 2008, 80% by the end of 2009 and 100% by the end of 2010. Also, minimal internet velocities were set, increasing with population size (Table 6).

Table 6: Backhaul Plan – setup

Population Size	N## municipalities	%	Velocity (Mbps)
Up to 20,000	3,077	89.5	8
From 20,001 to 40,000	268	7.8	16
From 40,001 to 60,000	63	1.8	32
Above 60,001	31	0.9	64
Total	3,439	100.0	

Source: Anatel

According to the National Agency of Telecommunication (Anatel[27]) (ANATEL, 2010), the majority of municipalities to be covered by Backhaul program were up to 20,000 inhabitants, which is more than half of total municipalities of Brazil[28]. The minimal required velocity (8 Mbps[29]) guarantee improvement in navigation quality, allowing, for example, streaming (of music and videos).

The program had three types of technology to be deployed: fiber, radio, and satellite. The first is installed by cables of optical fiber, with less interference in long distances, being connected directly to the household (FTTH) or to a concentrating point (FTTC), with a higher cost of installation and maintenance. The second one is usually easier to be installed (by antennas), maintained and reaches broader areas, like rural locales, but have limitations of interference, due to physical barriers, and of internet speed, due to distance. Finally, the third needs a satellite, an antenna in the household and a base antenna to intermediate communication, a set with high costs of installation and maintenance, but capable to reach broader areas, like rural, still being susceptible to weather interference. Considering the costs, radio was the main technology chosen, for 71% of the cities, followed by fiber, for 26%, and satellite for only 3%.

[27] *Agência Nacional de Telecomunicações* in Portuguese.

[28] Today, Brazil has 5,568 municipalities and two districts. By the time when the program was created, six municipalities did not exist yet.

[29] Megabit per second.

Out of 5,570 municipalities, by 2015, only 85 remained uncovered (Table 7) and 2,125 (38%) already had broadband infrastructure before the program's start, mainly larger cities. The program focused on small cities, with average population under 15,000.

Table 7: Backhaul deployment by coverage status, 2015

Situation	## Munic	Avg Pop.
Covered	3,360	14,403
Covered before	2,125	67,151
Uncovered	85	35,372
Total	5,570	34,072

Source: Anatel and IBGE

According to program schedule, 100% of Brazilians' municipalities should has backhaul infrastructure in 2010. However, by this year, only 72% of the goal was achieved. Table 8 presents the roll out of the program by year.

Table 8: Backhaul deployment by year

Backhaul year	## Munic	Avg. Velocity	Avg Pop.
2008	1,384	13	16,911
2009	1,388	10	13,340
2010	495	9	9,026
2011	27	2	12,134
2012	7	14	25,531
2013	41	4	20,238
2014	17		38,490
2015	1		13,293
Total	3,360	11	14,403

Source: Anatel.

Obs.1: No velocity information for 2014 and 2015.

Obs.2: Information only for program participants cities.

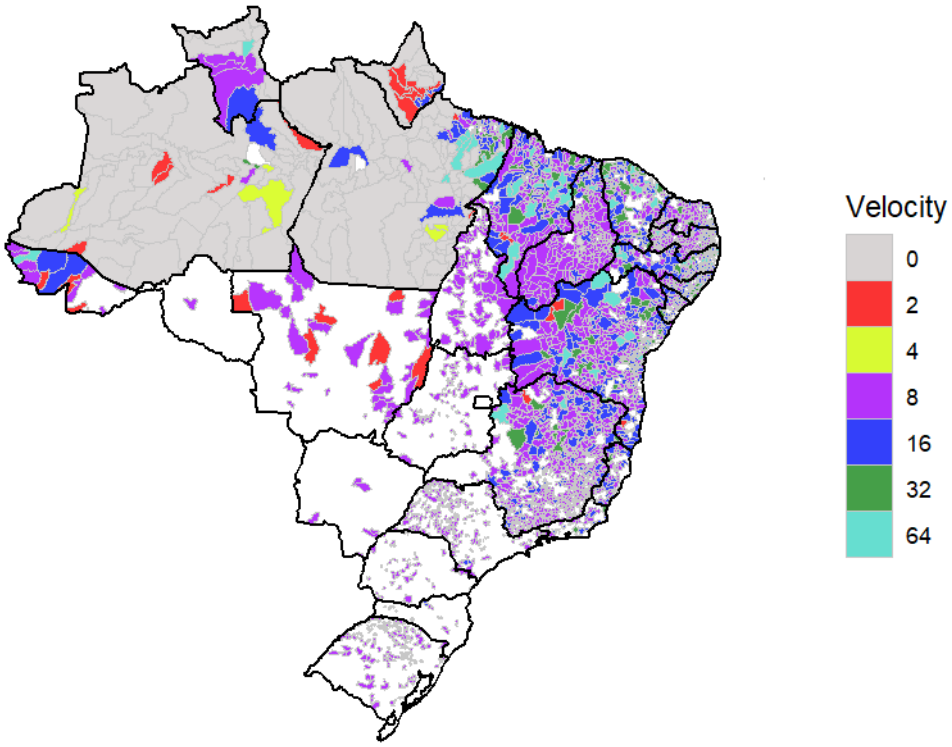
The main point of our identification strategy relies on the velocity discontinuity, which is further analyzed, geographically[30], in Figure 6. North and Northeast regions are poorer, while South and Southeast are richer[31], the very same patter when we look the program

[30]Brazil is a continental country (8,516,000 km² of area, the fifth in the world) with important regional inequalities.

[31]For example, the state of *São Paulo* was responsible for almost 1/3 of Brazilian GDP in 2017. Household income *per capita* of the richest state (Federal District) was 3.84 times greater than the poorest (Alagoas), according to 2014 National Household Survey (IBGE/PNAD). Brazilian Gini index for the same year was 0.517.

deployment in the territory.

Figure 6: Internet velocity in backhaul program by municipality



Source: Anatel and IBGE

Figure 6 shows that an important portion of cities in the south and center-west were covered before (blank areas), while the northeast had the largest number of cities in the program. Also, the north region (the Amazon area) had a lot of cities uncovered by the program (gray areas). The most common velocity was 8 Mbps, as showed before in Table 6, corresponding to cities under 20,000 inhabitants.

2.5.3 Methodology

Following CATTANEO; TITIUNIK; VAZQUEZ-BARE (2018), each municipality has a running variable X_i (the size of population) with potential outcomes $Y_i(0)$ (a lower internet velocity) and $Y_i(1)$ (higher – double – internet velocity). Municipalities face three possible cutoffs $C_i \in C$, with $C = c_1, c_2, c_3$. Ranges of population determine which type of treatment a municipality will receive: at least 8 Mpbs if $X_i \leq c_1$, at least 16 Mbps if $c_1 < X_i \leq c_2$, at least 32 Mbps if $c_2 < X_i \leq c_3$ and at least 64 Mbps if $X_i > c_3$.

Denote each treatment as d_j, so $D_i \in \{d_1, d_2, d_3\}$.

The effect for each cutoff, under standard regularity conditions, is identified by:

$$\tau_j = \mathbb{E}[Y_i(d_j) - Y_i(d_{j-1})|X_i = c_j] = \lim_{x \downarrow c_j} \mathbb{E}[Y_i|X_i = x] - \lim_{x \uparrow c_j} \mathbb{E}[Y_i|X_i = x] \quad (1)$$

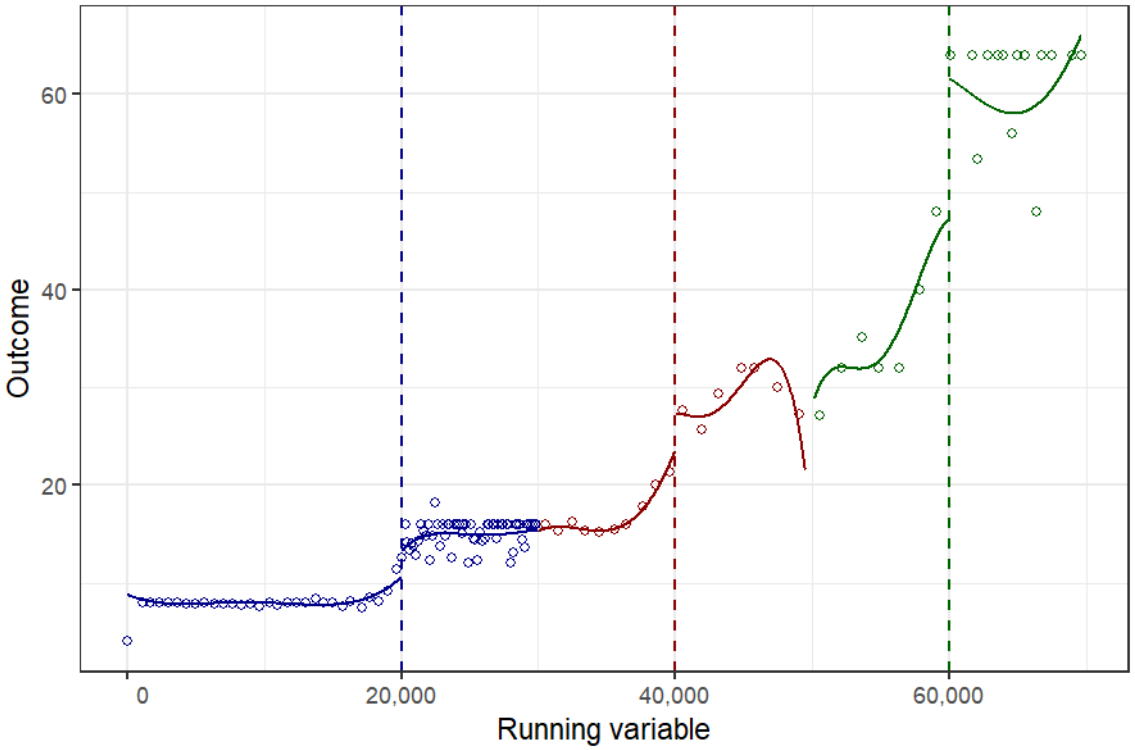
Each observation may be used to estimate two different, and contiguous, treatment effects. For example, looking to the first cutoff, 20,000, municipalities with population above that value are treated (up to the next cutoff – 40,000) when estimating τ_j, but they are controls when estimating τ_{j+1} (and, hence, below the next cutoff – 40,000). An observation can or cannot be used to estimate two effects, depending on bandwidth selection.

We consider the three cut-offs (20,000, 40,000 and 60,000), with optimal bandwidth chosen by minimizing the asymptotic mean squared error following CALONICO; CATTANEO; TITIUNIK (2014), CALONICO et al. (2017) and CALONICO; CATTANEO; FARRELL (2018).[32].

Figure 7 shows a clear jump in velocity cutoffs for the entire period. Municipalities just below the population size established in the program face lower internet velocities.

[32]Regressions are performed in R software, with rdmulti package: Matias D. Cattaneo, Rocio Titiunik and Gonzalo Vazquez-Bare (2020). rdmulti: Analysis of RD Designs with Multiple Cutoffs or Scores. R package version 0.6. https://CRAN.R-project.org/package=rdmulti

Figure 7: Discontinuity in Backhaul program velocity by population cutoffs: 20,000; 40,000; 60,000



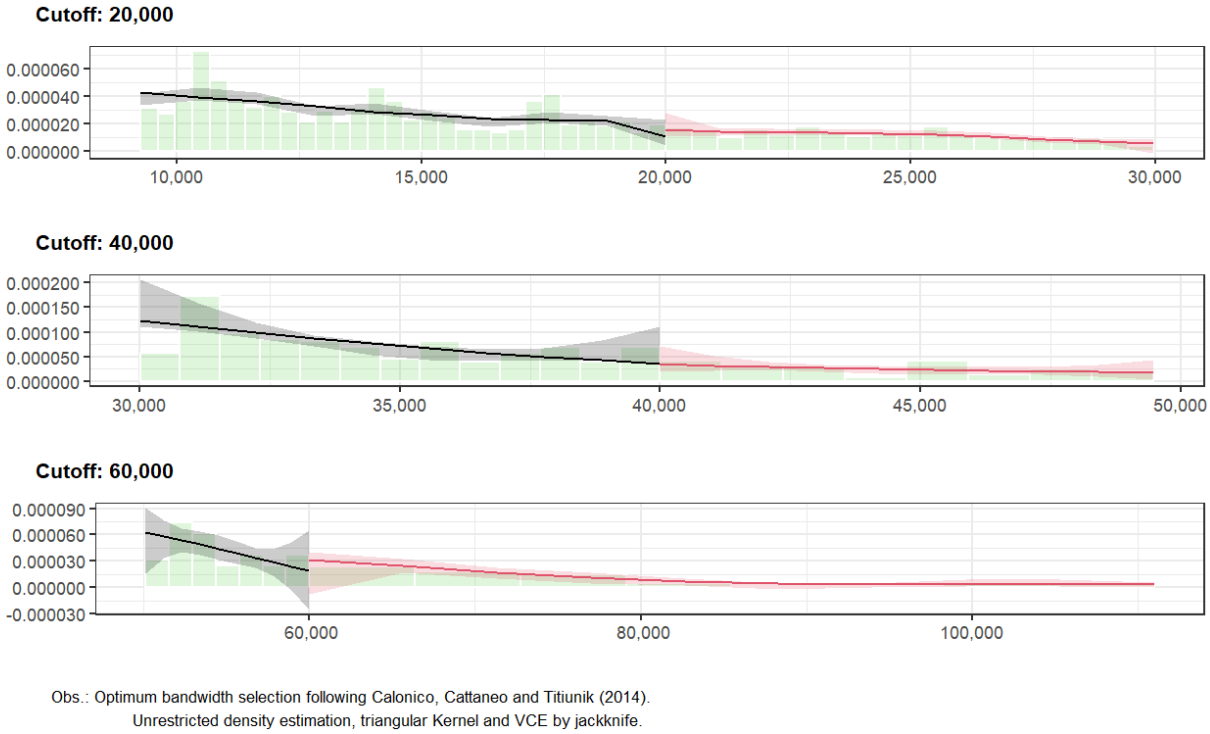
The classical McCrary manipulation test of cutoffs (MCCRARY, 2008) looks if there is a selection into treatment, analyzing the density distribution of the running variable around the cutoff. CATTANEO; JANSSON; MA (2019) and CATTANEO; JANSSON; MA (2021) developed an alternative test, where confidence bands are provided and is well suited for RDD designs. Results for this test are presented in Table 9 and Figure 8 for all three cutoffs.

Table 9: Cattaneo, Jansson and Ma manipulation test of cutoffs

Cutoff	Bw	N	Nl	Nr	T (jackknife)	P.value
20,000	3,584	3,001	239	158	0.950	0.342
40,000	11,403	226	168	58	-1.157	0.247
60,000	16,997	115	47	37	-0.151	0.880

Obs.1: Optimum bandwidth selection following Calonico, Cattaneo and Titiunik (2014).
Obs.2: Unrestricted density estimation, triangular Kernel and VCE by jackknife.
Obs.3: Bw=bandwidth; N, Nl and Nr are total n## of obs., n## on the left and n## on the right.

Figure 8: Density plot - Cattaneo, Jansson and Ma manipulation test of cutoffs



Obs.: Optimum bandwidth selection following Calonico, Cattaneo and Titiunik (2014).
Unrestricted density estimation, triangular Kernel and VCE by jackknife.

The manipulation test, together with Figure 7, suggests that our identification strategy is valid, for all three cutoffs, although the second and third have a lower number of observations. As we can see on Figure 8, the visual inspect confirms the test results. Further, the figure shows that Brazil has an odd population distribution, with unexpected jumps in some population ranges. MONASTERIO (2013) shows that these jumps occur due to a legislation regarding Federal transfers of resources to municipalities (*Fundo de Participação dos Municípios - FPM*), based on the population size[33]. Despite there are no intersection of running variable and FPM's cutoffs, we control for the later one, to avoid any possible confounder effect regarding this situation in results.

2.5.4 Databases

Outcomes are election results/information, organized by Superior Election Court (TSE)[34]. We will analyze 2008, 2010 and 2012 elections, covering two municipal and one national

[33] According to the Decree-Law 1,881/1981, there are 17 ranges of population, with increasing possibility of resources distribution for each range. The cuts are: 10,188; 13,584; 16,980; 23,772; 30,564; 37,356; 44,148; 50,940; 61,128; 71,316; 81,504; 91,692; 101,880; 115,464; 129,048; 142,632; 156,216. Available in http://www.planalto.gov.br/ccivil_03/Decreto-Lei/1965-1988/Del1881.htm

[34] *Tribunal Superior Eleitoral* in Portuguese.

suffrage. The main outcomes are: turnout, percentage of blank or null votes[35] and vote shares, for left wing parties. As mentioned before, using POWER; ZUCCO JR (2012) party index from Brazilian Legislative Survey[36]. left wing parties are those up to quantile 0.25 of the index, center parties are those between 0.25 and 0.75 and right wing are those above). We also look to a small party, young candidates (under 30 years-old) and their budget campaign.

Despite the clear discontinuities in the running variable, a set of covariates were collected, to control for any further confounders that might remain. Lack of information at municipal level is one of the weakness in Brazilian research at this territory level. Census occurs only every ten years[37], remaining just few administrative data in the between years, some of them with low quality (mainly for small cities). Even tough, considering this is the only source of the main socioeconomic variables, we use information from the last two censuses (2000 and 2010), organized by Brazilian Institute of Statistics and Geography (IBGE). Also, from IBGE, we collect total population estimates and GDP. Considering that direct cash transfers are important in Brazil, we collect data from the two major programs: *Bolsa Família* (PBF) and *Benefício de Prestação Continuada* for elders (BPC), booth organized by Ministry of Citizenship[38]. In addition, we collect the mass of wages (formal labor market) from RAIS database, organized by Ministry of Economy[39]. We also collected information from National Institute of Meteorology, to control for rain and temperature in election day, following FUJIWARA; MENG; VOGL (2016). Municipalities were joined by the nearest distance between the center of the city and the closest meteorological station. Table 10 summarizes each variable and source.

[35]In Brazilian election, people may cast a blank vote, which is not computed for any candidate and is not considered for official results, as well null votes. The difference consists in the way the registration of these votes is made: the blank vote is available as a button in the electronic ballot, while the null vote occurs when someone enters an invalid candidate number into the ballot and confirms the vote.

[36]Version 7, available in https://dataverse.harvard.edu/dataverse/bls;jsessionid=992eedb7e954a17ef718c7078cf5?widget=dataverse%40harvard&q=&types=dataverses%3Afiles%3Adatasets&sort=dateSort&order=desc&page=3

[37]When not delayed. The 1990 census was postponed to 1991, as well as 2020 census is postponed to 2021.

[38]PBF is one of the biggest conditional cash transfer program in the world. The target are families under the extreme poverty and poverty lines (in 2020, families earning up to R$ 89 by person, or U$ 17, by month are considered extremely poor, while families above that amount and up to R$ 178, or U$ 35, are considered poor), focused one children. As counterpart, school attendance and vaccination are required. PBF reached around 14 million families in Brazil in 2021. On the other hand, BPC is a program for elderly and handicapped. The poor population in this profile (people aged 65 or above and all handicapped) are eligible for a minimum wage paycheck (R$ 1.100, or U$ 216, in 2021).

[39]In Brazil, every formal company have to fill the Annual Relation of Social Information (RAIS), with the profile of all workers they had in the calendar year, including wages.

Table 10: Variables, description and source, by type

Category	Variable	Description	Source
Outcome	Turnout	Ratio between participants in elections and total electorate	TSE
	Vote share	Vote share of parties/candidates	TSE
	Blank and null votes	Percentage of blank and null votes in total	TSE
	Campaign budget	Total campaign budget declared to TSE	TSE
Running	Population	Estimated population	IBGE
Controls	Black	Percentage of blacks in population	IBGE
	College	Percentage of people with college degree	IBGE
	Married	Percentage of people married	IBGE
	Income	Median household income	IBGE
	Population over 60 years	Percentage of population over 60 years in population	IBGE
	Radio	Percentage of households with radio	IBGE
	Rural	Percentage of population in rural areas	IBGE
	Television	Percentage of households with television	IBGE
	Working age population	Percentage of population in working age	IBGE
	GDP	Gross Domestic Product	IBGE
	BPC	Ratio of BPC payments and GDP	MC and IBGE
	PBF	Ratio of PBF payments and GDP	MC and IBGE
	Formal wage	Ratio of formal wages (sum) and GDP	ME and IBGE
	Temperature	Average temperature a week before and after election day	Inmet
	Rain	Rain precipitation in election day	Inmet
	Fiber	Optical fiber infrastructure in municipality internet provision	Anatel
	FPM	Per capita FPM transfer	Treasury

Source: TSE, IBGE, Inmet, ME (Ministry of Economy), MC (Ministry of Citizenship) and Anatel.

2.5.5 Descriptive statistics

Descriptive statistics are separated by population ranges, considering the cutoffs (under 20,000, between 20,000 and 40,000, between 40,000 and 60,000 and above 60,000). Table 11 shows the figures for 2008 year.

We notice that municipalities under 20,000 inhabitants have less percentage of blacks, BPC transfers and optical fiber technology penetration, while more population in working age, in rural areas, ownership of radio, population over 60 years and married people.

In order to clarify identification validity, Table 12 presents a simple t-test for 20,000 population cutoff with a 3584 bandwidth, the same used before in the manipulation test.

Results for 2008 year in Table 12 (which, indeed, refers to 2000 census for socioeconomic variables) show that there were no significant differences for most of the characteristics between municipalities just above and just below the cutoff, except for formal wages (at 1% of significance), BPC and rural population (at 5% of significance), fiber and television (at 10% of significance). Some results, however, do not hold in 2010, year with values collected from 2010 census and, hence, closer to the year of analysis. Some covariates, like median income,

Table 11: Descriptive statistics by population size of municipality, 2008

Variable	Under 20k	Above 20k to 40k	Above 40k to 60k	Above 60k
Avg. Temperature	24.2	25.6	26.4	27.2
Black	53.4	65.5	68.1	68.6
BPC	0.3	0.8	0.8	0.9
College	0.7	0.6	0.6	0.8
Fiber	0.2	0.3	0.3	0.5
Formal Wages	11.5	11.5	12.8	13.8
FPM	742.9	326	261.3	214.3
GDP p.c.	7.8	6.9	6.3	8.2
Income	603.2	549.6	555.2	645.2
Married	30	24.1	22	22.5
PBF	2.5	3	2.6	2
Pop. over 60 years	9.9	8.5	8	7
Population	8,042	27,272	47,976	97,096
Radio	79.3	75.4	75.6	76.1
Rain (elect. day)	3.3	1.7	1.9	2.9
Rural	49.8	46.6	40.3	32.2
Television	69.2	66.4	68.1	73.6
Working Pop.	41.4	38.8	37.9	38.9
Observations	2,733	520	107	72

Source: IBGE, Inmet, ME, MC and Anatel.

Table 12: Covariates means difference t test for 20,000 cutoff, 2008, 2010 and 2012

Variable	Diff 2008	Diff 2010	Diff 2012
Median Income	5.7	-92.6***	-38.4
Pop. over 60 years	-0.11	-0.03	0.23
Rural	4.59**	6.84***	4.28**
Black	-1.51	2.44	0.29
Radio	-1.20	-1.10	0.16
Television	-3.48*	-1.91**	-1.11
College	0.03	-0.26**	-0.16
Married	0.05	0.02	0.66
Working Pop.	-0.73	-2.22***	-0.77
Rain (elect. day)	-0.2	0.1	0.0
Avg. Temperature	0.0	0.4	0.2
PBF	0.24	0.67***	0.48**
BPC	-0.22**	-0.27***	-0.14
GDP	0.8	-2.6	-1.7
Formal Wages	-1.73***	-1.18*	-0.91
Fiber	-8.00*	-3.67	-3.72
FPM transfer (pc)	51.5***	53.2***	73.4***
N. Obs	380	383	384

Source: IBGE, Inmet, ME, MC and Anatel.

Obs.1: Null hypotheses is no difference.

Obs.2: * = significant at 10%; ** = significant at 5%; *** = significant at 1%.

Obs.3: Bandwidth: 3,584

rural areas, working age population BPC and PBF seems to be different across municipalities, although in a low absolute difference for the most of them. Results for 2012 are more like those observed in 2000, for the most of variables. Overall, Table 12 results suggests that our

identification strategy should work if controlled for covariates.

2.6 Results

Considering there are two rounds for two types of offices, mayor and president, and some municipalities might not have a second round, we focus only on the first one, using, hence, the larger sample size as possible. In tables results, there are always three election years, where two are local (2008 and 2012) and the other is national (2010). A summary of first stage of fuzzy RDD are reported in Table 61 in Appendix, all of them supporting the identification strategy.

We begin our analysis of results looking to the effects of broadband internet velocity in participation. Results in Table 13 suggest no relationship between the velocity of this technology and participation in elections. For all regressions, considering the three cutoffs and the three years, only one had a slightly significant result (2010, for 60,000 inhabitants). Since participation in elections is mandatory in Brazil, and turnout is relatively high (around 80%), maybe there is not much room to improve situation. These results are in line with MENEZES (2015) for Brazil, as well as results reported by MINER (2015) for Malaysia, but differ from results reported in USA and European countries (CAMPANTE; DURANTE; SOBBRIO, 2017; FALCK; GOLD; HEBLICH, 2014; GAVAZZA; NARDOTTO; VALLETTI, 2019; JABER, 2013), where vote is not mandatory, and, hence, an important institutional difference.

Table 13: Fuzzy-RDD multi-cutoff results for turnout. Election years: 2008, 2010 and 2012

Year	Cutoff	Bw	Obs.	Coef.	SE	P.value
2008	20,000	3,584	401	0.003	0.010	0.634
	40,000	11,403	265	0.003	0.004	0.229
	60,000	16,997	126	-0.002	0.003	0.168
2010	20,000	4,924	554	-0.001	0.006	0.493
	40,000	12,191	302	0.001	0.002	0.419
	60,000	23,183	200	0.002	0.001	0.071
2012	20,000	10,735	1,335	0.002	0.003	0.507
	40,000	6,617	143	0.001	0.010	0.870
	60,000	11,631	95	0.001	0.003	0.649

Obs.: Standard Errors (SE) are clustered by regions, with heteroskedasticity-robust nearest neighbor variance estimator (three minimum neighbors). Optimal bandwidth (Bw) selection by Mean Square Error following Calonico, Cattaneo and Titiunik (2014). Triangular kernel with quadratic local-polynomial. Turnout for the first round. Results with controls listed in Table 10.

Seeing from a different perspective, the new possibility of leisure did not reduce people participation in elections. In other backgrounds, where participation in elections are not mandatory, results may be different (like in Germany and UK, showed by FALCK; GOLD; HEBLICH (2014) and GAVAZZA; NARDOTTO; VALLETTI (2019), respectively with negative effects).

The next outcome regards to the percentage of blank or null votes (Table 14). Again, there is no support in favor of the influence of broadband internet speed in this type of vote (a *proxy* for "absence of engagement with political process", since these votes can be seen as a "whatever vote"). So, results so far suggests that broadband did not encourage (nor discourage) people to turnout, neither people to place more directed votes in elections (again in accordance to MENEZES (2015) results).

Table 14: Fuzzy-RDD multi-cutoff results for blank or null votes. Election years: 2008, 2010 and 2012. Offices: president and mayors.

Year	Cutoff	Bw	Obs.	Coef.	SE	P.value
2008	20,000	4,250	471	-0.095	0.136	0.752
	40,000	11,989	282	0.000	0.003	0.565
	60,000	20,209	147	0.001	0.004	0.490
2010	20,000	4,098	461	0.004	0.004	0.182
	40,000	9,872	251	0.002	0.004	0.320
	60,000	18,868	152	0.000	0.001	0.877
2012	20,000	3,681	413	0.023	0.027	0.196
	40,000	14,408	394	-0.002	0.005	0.677
	60,000	13,801	107	0.034	0.170	0.822

Obs.: Standard Errors (SE) are clustered by regions, with heteroskedasticity-robust nearest neighbor variance estimator (three minimum neighbors). Optimal bandwidth (Bw) selection by Mean Square Error following Calonico, Cattaneo and Titiunik (2014). Triangular kernel with quadratic local-polynomial. Results with controls listed in Table 10.

In last presidential elections (2018), polarization was dramatic in Brazil. Left versus Right debate was at the center of presidential run, with the last four times winner party (the left-wing Workers Party – PT) being the main target. Before that, the 2014 elections were one of the closest seen in Brazil, when Mrs. Rouseff defeated Mr. Neves (from central right Brazilian Social Democracy Party – PSDB) with only 51.64% of the valid votes in the second round. Internet may had an important role in this scenario, since, back in 2010, Mr. Lula da Silva, the first president of Workers Party, had 80% of presidency approval, the highest value ever recorded.[40]

[40] A news about these figures are available in: http://g1.globo.com/politica/noticia/2010/12/popularidade-

Hence, a closer look at the relationship between broadband velocity and vote share of left parties back to 2008 might shed light into this turnaround in Brazil. As pointed before, vote shares were classified as left, center or right based on POWER; ZUCCO JR (2012) party index.

Results suggests, once again, no relationship between broadband internet speed and the vote share received by left wing parties in elections for president and mayors (Table 15). So, unlike results reported by previous studies (CAMPANTE; DURANTE; SOBBRIO, 2017; FALCK; GOLD; HEBLICH, 2014; GAVAZZA; NARDOTTO; VALLETTI, 2019; JABER, 2013), there is little evidence of important effects of broadband internet on vote shares, at least for fixed internet.

Table 15: Fuzzy-RDD multi-cutoff results for left-wing parties vote share. Election years: 2008, 2010 and 2012. Offices: president and mayors.

Year	Cutoff	Bw	Obs.	Coef.	SE	P.value
2008	20,000	3,723	217	-0.269	2.040	0.940
	40,000	12,974	195	-0.009	0.034	0.919
	60,000	26,281	143	-0.001	0.004	0.809
2010	20,000	4,810	542	-0.003	0.011	0.820
	40,000	7,651	175	-0.025	0.319	0.800
	60,000	21,243	183	0.003	0.003	0.243
2012	20,000	3,693	237	-0.229	1.285	0.484
	40,000	12,756	212	-0.006	0.016	0.876
	60,000	15,286	87	-0.048	0.127	0.497

Obs.: Standard Errors (SE) are clustered by regions, with heteroskedasticity-robust nearest neighbor variance estimator (three minimum neighbors). Optimal bandwidth (Bw) selection by Mean Square Error following Calonico, Cattaneo and Titiunik (2014). Triangular kernel with quadratic local-polynomial. Left wing parties: PSTU, PSOL, PC do B, PT, PSB and PCO. Results with controls listed in Table 10.

Despite results so far suggest no relationship between left wing parties and votes, some smaller parties, that face narrow campaign budgets, could use broadband internet to reach more people at lower costs. Table 16 presents the vote share of PSOL party for local legislators (*vereador*) and federal deputy (*deputado federal*), offices with more number of candidates[41]. PSOL (*Partido Socialismo e Liberdade*[42]) is a relatively recent left wing party, formed in 2004 with dissidents from PT, which makes an interesting case of study.

de-lula-bate-recorde-e-chega-87-diz-ibope.html

[41]Executive offices campaigns are more expensive and parties usually support each other to improve winning chances, forming blocks (*coligações*).

[42]Socialism and Liberty Party.

Table 16: Fuzzy-RDD multi-cutoff results for PSOL vote share. Election years: 2008, 2010 and 2012. Offices: local legislator and federal deputy.

Year	Cutoff	Bw	Obs.	Coef.	SE	P.value
2008	20,000	13,955	66	-0.0016	0.0017	0.6262
	40,000	27,243	84	0.0000	0.0006	0.3531
	60,000	38,619	71	0.0009	0.0030	0.7607
2010	20,000	6,802	772	-0.0004	0.0002	0.0119
	40,000	13,581	358	0.0002	0.0002	0.0250
	60,000	20,984	179	-0.0001	0.0001	0.0789
2012	20,000	9,708	77	0.0029	0.0187	0.8564
	40,000	12,799	62	0.0003	0.0017	0.8727
	60,000	17,647	42	0.0010	0.0038	0.5404

Obs.: Standard Errors (SE) are clustered by regions, with heteroskedasticity-robust nearest neighbor variance estimator (three minimum neighbors). Optimal bandwidth (Bw) selection by Mean Square Error following Calonico, Cattaneo and Titiunik (2014). Triangular kernel with quadratic local-polynomial. Results with controls listed in Table 10.

Results suggests a negative effect, for two out the three cutoffs, in 2010 national elections, but with a limited effect in terms of percentage of votes. MENEZES (2015), reports positive effects for small and third-placed parties, also for 2010 elections, which is, somewhat related with results found here, at least in significant results (not in magnitude nor direction).

Another possible effect is in votes for young candidates (under 30 years-old), who could take better advantage of broadband internet due to familiarity to new technologies. Table 17 presents the vote share of local legislators (*vereador*) and federal deputy, who also have more candidates running than for executive offices.

Results suggest no relationship between broadband internet speed and vote share for young candidates, in any year or cutoff, meaning this technology seems have not helped in electoral performance of younger.

We now investigate two outcomes not related to ballots directly, but with candidates' participation and budget. Table 18 present results for the first variable, only for 2008 and 2012 years, once, in national elections, candidates do not run representing cities or districts.

Table 17: Fuzzy-RDD multi-cutoff results for young candidates (under 30 years-old). Election years: 2008, 2010 and 2012. Offices: local legislator and federal deputy.

Year	Cutoff	Bw	Obs.	Coef.	SE	P.value
2008	20,000	4,621	506	-0.090	0.261	0.909
	40,000	13,638	333	0.004	0.006	0.180
	60,000	10,466	71	0.002	0.002	0.135
2010	20,000	3,147	368	0.009	0.014	0.288
	40,000	11,604	281	-0.002	0.012	0.752
	60,000	7,021	55	0.000	0.001	0.795
2012	20,000	4,319	474	-0.011	0.013	0.206
	40,000	4,806	102	-0.341	25.000	0.878
	60,000	12,816	102	0.013	0.050	1.000

Obs.: Standard Errors (SE) are clustered by regions, with heteroskedasticity-robust nearest neighbor variance estimator (three minimum neighbors). Optimal bandwidth (Bw) selection by Mean Square Error following Calonico, Cattaneo and Titiunik (2014). Triangular kernel with quadratic local-polynomial. Results with controls listed in Table 10.

Table 18: Fuzzy-RDD multi-cutoff results for number of candidates. Election years: 2008, 2010 and 2012. Offices: local legislator and federal deputy.

Year	Cutoff	Bw	Obs.	Coef.	SE	P.value
2008	20,000	4,309	476	0.645	0.969	0.846
	40,000	11,976	282	-0.103	0.190	0.519
	60,000	28,290	262	0.006	0.032	0.758
2012	20,000	12,800	1,680	0.000	0.028	0.822
	40,000	6,637	143	0.166	0.513	0.661
	60,000	10,808	89	-0.079	0.169	0.828

Obs.: Standard Errors (SE) are clustered by regions, with heteroskedasticity-robust nearest neighbor variance estimator (three minimum neighbors). Optimal bandwidth (Bw) selection by Mean Square Error following Calonico, Cattaneo and Titiunik (2014). Triangular kernel with quadratic local-polynomial. Results with controls listed in Table 10.

Like most of outcomes so far, results suggest no relationship between broadband internet speed and number of candidates running for local offices. The new possibility to reach voters seems not be sufficient to attract people to run in elections.

Regarding budget campaign, we look two outcomes: amount used by a small party (PSOL) and by young candidates (Tables 19 and 20, respectively).

Table 19: Fuzzy-RDD multi-cutoff results for PSOL campaign budget Election years: 2008, 2010 and 2012. Office: local legislator

Year	Cutoff	Bw	Obs.	Coef.	SE	P.value
2008	20,000	19,155	58	0.053	0.266	0.725
	40,000	19,261	32	5.799	217.042	0.569
	60,000	28,808	26	0.037	0.199	0.941
2012	20,000	7,267	40	-0.032	0.337	0.929
	40,000	15,803	55	-0.056	0.074	0.206
	60,000	24,205	45	0.089	0.171	0.459

Obs.: Standard Errors (SE) are clustered by regions, with heteroskedasticity-robust nearest neighbor variance estimator (three minimum neighbors). Optimal bandwidth (Bw) selection by Mean Square Error following Calonico, Cattaneo and Titiunik (2014). Triangular kernel with quadratic local-polynomial. Results with controls listed in Table 10.

Table 20: Fuzzy-RDD multi-cutoff results for young candidates (under 30 years-old) budget. Election years: 2008, 2010 and 2012. Office: local legislator

Year	Cutoff	Bw	Obs.	Coef.	SE	P.value
2008	20,000	4,739	524	-0.157	0.603	0.838
	40,000	16,976	477	0.003	0.003	0.072
	60,000	11,331	73	0.001	0.002	0.694
2012	20,000	4,892	539	-0.017	0.010	0.021
	40,000	5,467	118	0.048	0.200	0.609
	60,000	31,060	368	-0.004	0.002	0.000

Obs.: Standard Errors (SE) are clustered by regions, with heteroskedasticity-robust nearest neighbor variance estimator (three minimum neighbors). Optimal bandwidth (Bw) selection by Mean Square Error following Calonico, Cattaneo and Titiunik (2014). Triangular kernel with quadratic local-polynomial. Results with controls listed in Table 10.

For the PSOL party, there is no relationship between broadband internet speed and budget, while for young candidates' results are mix: slightly significant and positive for just one cutoff in 2008 (40,000) and indeed negative in 2012 for the first and last cutoffs. If any conclusion could be taken is that broadband internet velocity is related with lower young candidates' budgets in 2012 elections. The only parallel in literature we can make about this

outcome is regarding party donating in US elections, where JABER (2013) reports a positive impact for Democratic Party.

2.6.1 Further investigation

In previous section, due to RD design, each regression was run considering the multiple cutoff structure. Another way to estimate results is with parametric regressions, using the distance of the running variable to the cutoff and adjusting it by a polynomial. We look to parametric regressions considering two specifications: linear and quadratic. This choice follows GELMAN; IMBENS (2019), to avoid possible noisy estimates, eventual sensitivity to the degree of the polynomial and problems with the confidence intervals.

The results are presented in Table 21, for all outcomes and the three elections years and cutoffs. The lack of relationship between broadband internet and turnout, blank and null votes, left wing vote share, number of candidates, PSOL vote share and budget, and young candidates vote share and budget remains. Significant results are sparse and, sometimes, with inverted signs when linear specification is switched to quadratic.

Table 21: Parametric Fuzzy-RDD results for all outcomes. Election years: 2008, 2010 and 2012

Year	Model	Cutoff	Obs.	Coef.	SE	P.value	Outcome
2008	Linear	20,000	3,356	0.001	0.001	0.196	Turnout
	Quadratic	20,000	3,356	0.002	0.010	0.862	
	Linear	40,000	3,356	0.001	0.000	0.011	
	Quadratic	40,000	3,356	-0.001	0.001	0.222	
	Linear	60,000	3,356	0.002	0.001	0.023	
	Quadratic	60,000	3,356	-0.001	0.001	0.061	
2010	Linear	20,000	3,427	0.001	0.001	0.089	Turnout
	Quadratic	20,000	3,427	0.000	0.001	0.876	
	Linear	40,000	3,427	0.000	0.000	0.010	
	Quadratic	40,000	3,427	-0.002	0.001	0.147	
	Linear	60,000	3,427	0.002	0.000	0.000	
	Quadratic	60,000	3,427	0.000	0.001	0.977	
2012	Linear	20,000	3,428	0.001	0.001	0.329	Turnout
	Quadratic	20,000	3,428	0.000	0.001	0.579	
	Linear	40,000	3,428	0.000	0.000	0.203	
	Quadratic	40,000	3,428	-0.002	0.001	0.132	
	Linear	60,000	3,428	0.001	0.000	0.001	
	Quadratic	60,000	3,428	0.000	0.000	0.944	

Table 21: Parametric Fuzzy-RDD results for all outcomes. Election years: 2008, 2010 and 2012 *(continued)*

Year	Model	Cutoff	Obs.	Coef.	SE	P.value	Outcome
2008	Linear	20,000	3,356	0.001	0.001	0.154	Blank or Null votes
	Quadratic	20,000	3,356	0.018	0.067	0.788	
	Linear	40,000	3,356	-0.002	0.000	0.000	
	Quadratic	40,000	3,356	0.001	0.000	0.001	
	Linear	60,000	3,356	-0.002	0.002	0.267	
	Quadratic	60,000	3,356	0.001	0.002	0.590	
2010	Linear	20,000	3,427	0.000	0.000	0.690	Blank or Null votes
	Quadratic	20,000	3,427	0.002	0.001	0.068	
	Linear	40,000	3,427	0.000	0.000	0.381	
	Quadratic	40,000	3,427	0.000	0.000	0.159	
	Linear	60,000	3,427	0.000	0.000	0.784	
	Quadratic	60,000	3,427	0.001	0.000	0.009	
2012	Linear	20,000	3,428	0.002	0.001	0.052	Blank or Null votes
	Quadratic	20,000	3,428	0.007	0.004	0.047	
	Linear	40,000	3,428	0.001	0.000	0.006	
	Quadratic	40,000	3,428	0.002	0.001	0.217	
	Linear	60,000	3,428	-0.001	0.001	0.491	
	Quadratic	60,000	3,428	-0.002	0.001	0.182	
2008	Linear	20,000	1,530	0.012	0.002	0.000	Left wing vote share
	Quadratic	20,000	1,530	0.115	0.458	0.803	
	Linear	40,000	1,530	-0.003	0.001	0.000	
	Quadratic	40,000	1,530	-0.015	0.004	0.000	
	Linear	60,000	1,530	-0.005	0.003	0.085	
	Quadratic	60,000	1,530	-0.012	0.008	0.145	
2010	Linear	20,000	3,427	0.001	0.001	0.575	Left wing vote share
	Quadratic	20,000	3,427	-0.002	0.004	0.610	
	Linear	40,000	3,427	0.001	0.001	0.586	
	Quadratic	40,000	3,427	0.001	0.001	0.288	
	Linear	60,000	3,427	0.000	0.001	0.960	
	Quadratic	60,000	3,427	0.001	0.001	0.319	
2012	Linear	20,000	1,738	-0.002	0.002	0.302	Left wing vote share
	Quadratic	20,000	1,738	-0.011	0.012	0.371	
	Linear	40,000	1,738	-0.001	0.002	0.480	
	Quadratic	40,000	1,738	-0.004	0.002	0.058	
	Linear	60,000	1,738	-0.002	0.002	0.261	
	Quadratic	60,000	1,738	-0.005	0.001	0.000	

Table 21: Parametric Fuzzy-RDD results for all outcomes. Election years: 2008, 2010 and 2012 *(continued)*

Year	Model	Cutoff	Obs.	Coef.	SE	P.value	Outcome
2008	Linear	20,000	3,356	0.007	0.006	0.185	N. cand.
	Quadratic	20,000	3,356	0.022	0.152	0.883	
	Linear	40,000	3,356	-0.007	0.011	0.539	
	Quadratic	40,000	3,356	-0.017	0.006	0.004	
	Linear	60,000	3,356	-0.003	0.002	0.149	
	Quadratic	60,000	3,356	0.010	0.012	0.401	
2012	Linear	20,000	3,427	0.010	0.008	0.224	N. cand.
	Quadratic	20,000	3,427	0.033	0.048	0.489	
	Linear	40,000	3,427	-0.008	0.010	0.466	
	Quadratic	40,000	3,427	-0.006	0.014	0.658	
	Linear	60,000	3,427	-0.018	0.010	0.070	
	Quadratic	60,000	3,427	-0.014	0.009	0.101	
2008	Linear	20,000	128	0.000	0.001	0.370	PSOL vote share
	Quadratic	20,000	128	0.002	0.001	0.225	
	Linear	40,000	128	0.001	0.001	0.089	
	Quadratic	40,000	128	0.000	0.000	0.038	
	Linear	60,000	128	0.002	0.005	0.652	
	Quadratic	60,000	128	0.000	0.004	0.915	
2010	Linear	20,000	2,815	0.000	0.000	0.266	PSOL vote share
	Quadratic	20,000	2,815	0.000	0.000	0.002	
	Linear	40,000	2,815	0.000	0.000	0.719	
	Quadratic	40,000	2,815	0.000	0.000	0.003	
	Linear	60,000	2,815	0.000	0.000	0.765	
	Quadratic	60,000	2,815	0.000	0.000	0.629	
2012	Linear	20,000	211	-0.001	0.001	0.363	PSOL vote share
	Quadratic	20,000	211	-0.003	0.009	0.708	
	Linear	40,000	211	0.000	0.000	0.018	
	Quadratic	40,000	211	0.000	0.000	0.770	
	Linear	60,000	211	0.000	0.000	0.024	
	Quadratic	60,000	211	0.000	0.000	0.410	
2008	Linear	20,000	89	0.245	0.350	0.484	PSOL budget
	Quadratic	20,000	89	-0.430	1.099	0.696	
	Linear	40,000	89	-0.007	0.036	0.855	
	Quadratic	40,000	89	-0.049	0.019	0.010	
	Linear	60,000	89	0.082	0.109	0.450	
	Quadratic	60,000	89	0.162	0.109	0.137	

Table 21: Parametric Fuzzy-RDD results for all outcomes. Election years: 2008, 2010 and 2012 *(continued)*

Year	Model	Cutoff	Obs.	Coef.	SE	P.value	Outcome
2012	Linear	20,000	149	-0.006	0.021	0.776	
	Quadratic	20,000	149	-0.016	0.117	0.894	
	Linear	40,000	149	-0.046	0.007	0.000	PSOL
	Quadratic	40,000	149	-0.141	0.033	0.000	budget
	Linear	60,000	149	0.029	0.016	0.079	
	Quadratic	60,000	149	0.044	0.039	0.265	
2008	Linear	20,000	3,237	0.002	0.002	0.258	
	Quadratic	20,000	3,237	0.014	0.038	0.720	
	Linear	40,000	3,237	0.001	0.001	0.104	Young
	Quadratic	40,000	3,237	0.001	0.002	0.489	votes
	Linear	60,000	3,237	-0.001	0.001	0.525	
	Quadratic	60,000	3,237	-0.002	0.001	0.006	
2010	Linear	20,000	3,413	0.001	0.001	0.139	
	Quadratic	20,000	3,413	0.001	0.004	0.800	
	Linear	40,000	3,413	0.000	0.000	0.419	Young
	Quadratic	40,000	3,413	-0.001	0.000	0.012	votes
	Linear	60,000	3,413	-0.001	0.000	0.004	
	Quadratic	60,000	3,413	-0.002	0.000	0.000	
2012	Linear	20,000	3,369	-0.001	0.001	0.098	
	Quadratic	20,000	3,369	-0.008	0.006	0.172	
	Linear	40,000	3,369	-0.001	0.001	0.390	Young
	Quadratic	40,000	3,369	-0.001	0.002	0.699	votes
	Linear	60,000	3,369	0.000	0.000	0.090	
	Quadratic	60,000	3,369	-0.001	0.001	0.322	
2008	Linear	20,000	3,237	0.001	0.003	0.828	
	Quadratic	20,000	3,237	0.015	0.060	0.800	
	Linear	40,000	3,237	0.001	0.001	0.100	Young
	Quadratic	40,000	3,237	0.002	0.002	0.305	budget
	Linear	60,000	3,237	0.000	0.001	0.961	
	Quadratic	60,000	3,237	-0.001	0.002	0.641	
2012	Linear	20,000	3,369	-0.002	0.000	0.000	
	Quadratic	20,000	3,369	-0.011	0.009	0.230	
	Linear	40,000	3,369	0.000	0.000	0.441	Young
	Quadratic	40,000	3,369	0.000	0.001	0.866	budget
	Linear	60,000	3,369	-0.001	0.000	0.167	
	Quadratic	60,000	3,369	-0.002	0.001	0.026	

Table 21: Parametric Fuzzy-RDD results for all outcomes. Election years: 2008, 2010 and 2012 *(continued)*

Year	Model	Cutoff	Obs.	Coef.	SE	P.value	Outcome

Obs.: Standard Errors (SE) are clustered by regions, with heteroskedasticity-robust variance estimator.
Left-wing parties: PSTU, PSOL, PC do B, PT, PSB and PCO.
Results with controls listed in Table 10.

The only result we could point out as relatively consistent is the negative relationship between the internet velocity and left-wing vote share, for local elections (2008 and 2010) and for 40,000 cutoff. But being too specific, it is hard to support as a consistent result, specially taking into considerations previous section.

A possible limitation of RDD models is the bandwidth choice, which could influence results. It is possible that a narrower or wider bandwidth give different results, since fewer or more observations will be part of regressions (a tradeoff between "randomness" and "precision"). Considering this possibility, Table 22 presents only significant results (at 5%) using also half or double bandwidths of those from previous section.

Table 22: Significant Fuzzy-RDD results with half or double bandwidths for all outcomes. Election years: 2008, 2010 and 2012

Year	Cutoff	Model	Obs.	Coef.	SE	P.value	Outcome
2008	60,000	Double-Bw	423	0.000	0.001	0.010	Turnout
2010	60,000	Double-Bw	1,214	0.001	0.001	0.000	Turnout
2008	20,000	Double-Bw	549	0.013	0.010	0.001	Blank or Null votes
2012	20,000	Double-Bw	519	0.002	0.002	0.000	Blank or Null votes
	60,000	Double-Bw	195	-0.006	0.005	0.002	
	20,000	Half-Bw	191	0.351	0.804	0.029	Left wing vote share
2008	20,000	Double-Bw	1,031	-0.103	0.097	0.000	N. cand.
2010	20,000	Double-Bw	1,687	0.000	0.000	0.000	PSOL vote share
	60,000	Double-Bw	848	0.000	0.000	0.008	
2008	60,000	Double-Bw	153	0.000	0.002	0.004	Young votes
2012	20,000	Double-Bw	1,031	-0.006	0.004	0.028	Young votes
	60,000	Double-Bw	248	-0.004	0.005	0.029	
	20,000	Double-Bw	1,242	-0.008	0.003	0.000	Young budget
	60,000	Double-Bw	3,359	-0.001	0.000	0.000	

Obs.: Standard Errors (SE) are clustered by regions, with heteroskedasticity-robust variance estimator.

Left-wing parties: PSTU, PSOL, PC do B, PT, PSB and PCO.
Results with controls listed in Table 10.

Once again significant results are sparse, for some years, outcome, cutoffs and bandwidths, putting in check any solid relationship between broadband internet speed and election outcomes.

2.6.1.1 Pooled regressions

The cumulative multi-cutoff RD design, although considers all the heterogeneity that multiple cutoffs offer, reduces the sample size, especially in the 40,000 and 60,000 cutoffs. We now look for pooled regressions, to verify if combined samples (and, hence, more observations) alter results presented in the previous sections. All cutoffs are considered, and samples are divided by the mean value between the cutoffs (population up to 30,000 for the first, between 30,001 and 50,000 for the second and 50,001 and above for the third). Table 23 shows results for all outcomes.

Table 23: RDD-fuzzy pooled multi-cutoffs for all outcomes. Election years: 2008, 2010 and 2012

Year	Bw	Obs.	Coef.	SE	P.value	Outcome
2008	3,435	476	0.001	0.008	0.786	
2010	3,559	510	-0.009	0.007	0.068	Turnout
2012	4,379	605	0.013	0.010	0.121	
2008	3,727	506	0.051	0.010	0.000	
2010	4,770	676	0.008	0.005	0.063	Blank or Null votes
2012	3,513	499	0.008	0.019	0.531	
2008	3,672	275	0.102	0.033	0.001	
2010	4,784	676	-0.002	0.021	0.869	Left wing vote share
2012	2,246	188	-0.096	0.027	0.000	
2008	2,281	315	-0.636	0.117	0.000	N. cand.
2012	3,614	509	0.200	0.127	0.053	
2008	3,684	28	-0.003	0.011	0.801	
2010	7,764	1,131	-0.001	0.001	0.113	PSOL vote share
2012	3,915	65	-0.002	0.005	0.257	
2008	4,311	27	1.496	0.522	0.003	PSOL
2012	3,138	40	2.840	0.519	0.000	budget
2008	3,476	477	0.039	0.029	0.102	
2010	4,650	655	-0.009	0.009	0.318	Young votes
2012	7,498	1,116	-0.016	0.027	0.600	
2008	3,894	517	0.053	0.029	0.031	Young
2012	5,954	867	-0.022	0.012	0.085	budget

Obs.: Standard Errors (SE) are clustered by regions, with heteroskedasticity-robust nearest neighbor variance estimator (three minimum neighbors). Optimal bandwidth (Bw) selection by Mean Square Error following Calonico, Cattaneo and Titiunik (2014). Triangular kernel with quadratic local-polynomial. Results with controls listed in Table 10.

Results suggest a negative effect for turnout in 2010 and positive effects for blank or null votes in 2008 and 2010. Regarding left wing vote share, a positive effect in 2008 is reverted to negative in 2012. PSOL vote share is no longer significant, while its budget is. In 2008, we see a negative effect to number of young candidates, reverting to positive in 2012. Finally,

for younger candidates, we observe no effects on votes, but a significant effect on budget: positive in 2008 reverted to negative in 2012.

Putting all these results together (the cumulative RDD, the parametric RDD and the pooled RDD), it is hard to conclude that the Backhaul program, and, hence, broadband internet velocity, made a significant difference in 2008, 2010 and 2012 elections in terms of turnout, percentage of blank or null votes, left wing vote share, PSOL vote share, young candidates' vote share and PSOL and young candidates budget. Although only the first round and four offices were analyzed, it is not likely to expect a different result in other scenarios (second rounds or other offices). The lack of consistency across specification, years and cutoffs put in check the significant results observed in some regressions.

2.7 Discussion and conclusions

Relationship between broadband velocity and elections outcome seems to be irrelevant in Brazil, at least when fixed broadband is considered, neither for local nor national elections between 2008 and 2012. Despite our robust identification strategy, we did not find strong relationship between broadband velocity, measured by the jumps of internet speed in the Backhaul program roll out, and election outcomes. These results are in line with some finds reported by MENEZES (2015) (turnout and blank and null votes).

Our results are also different from those reported in some part of the literature, mostly concentrated in European countries and USA (CAMPANTE; DURANTE; SOBBRIO, 2017; FALCK; GOLD; HEBLICH, 2014; GAVAZZA; NARDOTTO; VALLETTI, 2019; JABER, 2013), which could indicate that the background may be important in this kind of analysis. First, vote is mandatory in Brazil, which is not necessarily true in other countries. Second, Brazil is in a presidentialism system, in a federation republic, which means that people may behavior differently than in a parliamentarian system. Third, national congress deputies and local assemblies are elected by proportional vote, while senators are elected by majority vote, situation that may differ among countries. A fourth source of variation in political background regards to the difference between unitary and federal systems, that sets different rules to be played in the "political game".

Aside the political background, the internet usage is not fully controlled in our analysis. Installing the internet infrastructure in a municipality does not mean that all population will have access to the technology. Figure 5 shows that less than a third of Brazilian population had access to the internet in 2008, rising to 46.5% in 2012. Despite the lower costs of internet in electoral campaign, less than half of population could be reached until 2012, even a less

portion in poorer cities. So, it might be the case that internet velocity had a limited room to change reality and the sparse significant results found is not sufficient to trace a pattern yet.

Also, it was not possible to address the qualification of internet usage in our analysis. First, social networks have grown in Brazil after 2010. WhatsApp, one of the most popular social media in Brazil today, was created only in 2009, the very same year internet campaign was regulated. Is it possible that, today, mobile broadband and social medias usage in smartphones are more important for communication and mobilization than older social medias (like internet forums) and connection made at home, through desktop or laptop computers and by fixed land lines. Unfortunately, the roll out of 3G and 4G technology implementation, at municipality level, is unavailable. There is only data at Direct Dialing codes (DDDs)[43] areas, which makes impossible to determine when the technology begun to operate in every city[44].

Nonetheless, our paper contributes to bring into discussion that internet and political outcomes should be viewed in a wider perspective, meaning that some relationship may be circumstantial to idiosyncrasies of the countries or the technology. Also, further investigation, like the role of the new social medias and mobile broadband, are necessary to shed light in this discussion, even because internet and social medias are still evolving.

[43]The DDD codes are numbers that divides Brazil in 67 areas.

[44]We contacted the Regulation Agency of Telecommunication – Anatel – requesting mobile internet implementation at municipality level. Unfortunately, there is no such data available.

3 The power of short-term training programs: the case of Pronatec in Brazil

3.1 Introduction

The aim of this paper is to analyze the effects of the Brazilian training program – *Pronatec* – on employment and wages, from 2012 to 2019, including the spillover effects. Pronatec flooded specific job markets with thousands of trained workers, affecting not only participants but the whole local labor market where it was carried out, which demands an analysis not only of its direct effects but also the indirect effects.

In the end of 2011, the Federal government launched the National Program for Access to Technical Education and Employment (Pronatec), a nationwide initiative to improve job opportunities and quality of education in Brazil. This type of initiative is under Active Labor Market Programs (ALMP) rationale, often used as response to fight unemployment, with a variety of policies, most commonly job search assistance and training. This latter is vastly used around the world, with its effects being deeply studied (CARD; KLUVE; WEBER, 2010, 2018; CRÉPON; VAN DEN BERG, 2016; HECKMAN; LALONDE; SMITH, 1999; KLUVE, 2010; VOOREN et al., 2019).

In the end of 2016, Pronatec reached more than 9 million people, investing almost R$ 40 billion[45] (SENADO, 2017). One arm of Pronatec program is the *Bolsa Formação*[46], which offers short standardized training courses, free of charges and with stipends for food and transportation, to workers (regardless of occupational status) and students. *Bolsa formação* sums almost 50% of participants and 30% of total budget, being an important case of study and center of government attention (SENADO, 2017).

Using Pronatec application rules, we construct a quasi-experimental database to assess the impacts of Pronatec on employment probability and wages premium (log of hourly wages), as well spillover effects in local formal labor markets. Due to reasons out of candidates control, some applicants were not able to participate in training (oversubscription or class cancellation). Additionally, we consider only a small unemployment spell (within six months), taking into consideration the "earn dip" phenomenon (ASHENFELTER, 1978). We link participants to a rich administrative database which has information of all formal workers in the country, yearly. Combining these two databases we can determine the labor path of workers in the formal labor market after training conclusion. Using *logit* (for employment) and OLS[47] (for wages), we analyze seven periods after participation in program (from six to eighty-four months), to verify short, mid and long run effects. Also, we control for the share

[45] Or US$ 7.86 billion, using the exchange rate (R$ 5.09 to dollar) of 2021-07-15. The same exchange rate will be used along the text.

[46] In English, it would be like "Scholarship-formation".

[47] Ordinary Least Squares.

of participants concluding the course in the same year, in a specific market, to verify possible spillover effects (FERRACCI; JOLIVET; BERG, 2014). As robustness tests, we estimate results for several subsets, matching data set and instrumental variables. To the best of our knowledge, this is the first study to conduct a wide analysis of Pronatec, looking to the entire *Bolsa Formação* workers arm. Booth BASTO et al. (2016) and QUINTANA; CRAVO (2019) reports positive effects of Pronatec in employment but restrict to a specific initiative which represents less than 2% of the same program arm.

We find positive and significant effects in the short, mid, and long run (between 16% and 20% in employment odds ratio for treated when compared to control), which is like effects reported in literature (CARD; KLUVE; WEBER, 2010, 2018; GREENBERG; MICHALOPOULOS; ROBINS, 2003; HECKMAN; LALONDE; SMITH, 1999; KLUVE, 2010; VOOREN et al., 2019). Pronatec shares (the number of training concluders in a specific market) negatively contributes both to employment and wages in the short run (between two and three years after training conclusion), suggesting spillover. This finding is relevant since spillover effects are not always considered in this kind of analysis (CAHUC; LE BARBANCHON, 2010; DAVIDSON; WOODBURY, 1993). Results are robust to several subsets, indicating, in general, the same pattern. We also find heterogeneous effects, with lower results for women, higher for youth, different results among training types (larger impact for industry related courses) and among course requester (no effects for on-line applicants). These last findings contribute to literature once it is a common practice consider any kind of training as a single treatment, while we show effects separately. Additionally, when costs are taken into consideration, the Program seems to be highly economical, returning R$ 1.89 for each R$ 1 spent.

The paper is organized as follows: next section presents a review of past studies about ALMP and a general framework of their mechanisms. The third section explains the Pronatec program, databases, and the empirical strategy. Fourth section presents results, considering all subsets, as well robustness checks, heterogeneous effects and cost and benefit analysis. The last section ends with conclusions and remarks.

3.2 Active Labor Market Programs review

The effects of education in economic outcomes are widely investigated (HANUSHEK; MACHIN; WOESSMANN, 2016; JOHNES; JOHNES, 2007), including its relationship with labor market. To improve workers' skill, many governments provide or encourage training initiatives, as part of Active Labor Market Programs (ALMP). This is not a novelty, considering that type of programs dates the 1960's in the United States (Manpower Demonstration and Training Act – MDTA, for example), being reshaped and substituted time

after time (BARNOW, 1993). Among the objectives that can be related to ALMP, we can cite retaining employment, creating new job opportunities, incentives for seeking jobs, human capital accumulation (in line with BECKER, 1962; BEN-PORATH, 1967; MINCER, 1958; MINCER, 1974; SCHULTZ, 1961 theories of human capital) and improvements in market matching (BROWN; KOETTL, 2015). In Brazil, labor market policies are concentrated in the passive stream, where almost all expenses are concentrated in unemployment insurance or salary bonus for low wage workers (see SILVA (2018) analyses expenses of *Fundo de Amparo ao Trabalhador*[48]). On the other hand, in OECD countries, active policies averages 40% of all expenses (between 2004 and 2018)[49]. So, despite equivalent expenses (around 1% as percentage of GDP), composition is quite different.

Today, almost sixty years after the first large scale program implemented in US, we still have these kind of initiatives and, literally, hundreds of evaluations of their effectiveness, in a variety of countries (CARD; KLUVE; WEBER, 2010, 2018; GREENBERG; MICHALOPOULOS; ROBINS, 2003; HECKMAN; LALONDE; SMITH, 1999; KLUVE, 2010; VOOREN et al., 2019) and time (covering programs implemented from 1960's to the present decade). Thanks to that, there is a much better understanding of what might work or not in ALMP. General conclusions of meta-analysis about them are:

- There is, usually, no or negative effects of training programs in employment and wages in the short run (known as "lock-in effects");
- In the mid and long run, results turn to positive, suggesting that training may have lasting effects;
- Also, there are important heterogeneous effects, with larger effects for women and negligible or even negative for youths.
- Also, the design of the program matters, especially those with job search assistance;
- Macroeconomic background (like recessions) seems to delivery more positive results;
- Course length matters in short and long run (negatively and positively, respectively).

Other important finding is that experimental versus non-experimental studies leads, in general, to the same conclusions, a valid concern pointed out about non-experimental studies (ASHENFELTER; CARD, 1984; CAHUC; LE BARBANCHON, 2010; FRAKER; MAYNARD, 1987; HECKMAN; ICHIMURA; TODD, 1997; LALONDE, 1986). ANDERSSON et al. (2013) also contribute in this direction, suggesting that non-experimental evaluation may delivery reliable results when well applied to a rich data base. In addition, no publication biases were

[48]Workers Support Fund, in English, is a special fund, bound to Ministry of Economy, to undertake labor policies.

[49]According to Public expenditure and participant stocks on LMP, from OECD.Stat.

reported in this analysis. However, despite meta-analysis cover a large number of countries[50], not all of them include emergent markets or the developing world, like African and Latin American countries. Also, general equilibrium and spillover effects are only present in a fraction of them (DAVIDSON; WOODBURY, 1993). FERRACCI; JOLIVET; BERG (2014) study the spillover effects when few participants of a market are exposed to an intervention, focusing on training programs for unemployed workers in France. It is particularly important because, due to interaction between people, peer and neighborhood effects might be present as well general equilibrium changes, considering shifts in supply and demand sides of labor[51]. This issue is widely discussed by CAHUC; LE BARBANCHON (2010), for example. Also, much of these studies focus only on an extremely specific public (e.g., in a specific region or specific demographic group, like youths), putting serious restriction in external validity. Finally, there are few literatures that relates the costs and results, i.e., cost-benefit or cost-effectiveness analysis (JESPERSEN; MUNCH; SKIPPER (2008), RICHARDSON; VAN DEN BERG (2002), ATTANASIO; KUGLER; MEGHIR (2011), ATTANASIO et al. (2017) and ROSHOLM (2008) are exceptions).

Despite these general conclusions, distinct results are found, like positive effects of training in the short run[52] (ANDERSSON et al., 2013; BLASCO; ROSHOLM, 2011) and no effects in the long run (RICHARDSON; VAN DEN BERG, 2002), which makes room for ambiguous results. The usual outcomes are probability of employment or unemployment, duration of the same variables and earnings. Duration (BLASCO; CRÉPON; KAMIONKA, 2012; BLASCO; ROSHOLM, 2011), discrete choice models (*probit* and *logit*) or OLS are the most chosen estimation methods, usually combined with matching techniques. Finally, administrative data is widely used, offering a bunch of rich information, and usually for a long period, which explains the use of non-experimental analysis with robust results (CRÉPON; FERRACCI; FOUGÈRE (2012), for example, use a large and rich data set, which gives statistical power for estimates, the same path we will pursuit in this paper).

For Brazil, to the best of our knowledge, there are only two analyses of Pronatec program (BASTO et al., 2016; QUINTANA; CRAVO, 2019), both looking to a small subset of the program, with less than two percent of participants. We pretend to contribute with a wider

[50]The following countries were covered, at least, in one of meta-analysis aforementioned: Argentina, Australia, Austria, Belgium, Canada, Chile, Colombia, Denmark, Dominican Republic, Estonia, Finland, France, Germany, Hungary, Ireland, Israel, Italy, Mexico, Netherlands, New Zealand, Norway, Panama, Peru, Poland, Portugal, Romania, Russia, Serbia, Slovak Republic, Spain, Sweden, Switzerland, UK and USA.

[51]Increasing the contingent of workers with a specific skill could alter the market equilibrium, as well changes in relative prices (wages).

[52]CARD; KLUVE; WEBER (2010) points out that duration models are more likely to report positive effects in the short run.

analysis, looking to the whole *Bolsa Formação* arm[53] and taking into account the spillover effects, an important aspect due to the range of this program and the way the program was implemented.[54]

Following MCCALL; SMITH; WUNSCH (2016) theoretical framework, we present the main rationales behind ALMP, bringing some ideas of the channels of these interventions in outcomes, to give a better understanding of results that will be later presented. Following FERRACCI; JOLIVET; BERG (2014), we contribute to spillover effects discussion, analyzing the role of the Program in the local labor markets.

3.2.1 General framework

In a simple microeconomic point of view, a worker will only participate in labor training if she evaluates that future earnings or job opportunities exceed the costs involved to participate now (discount rates considered). Costs can be the time spent in training and resources needed to participate (fees, transportation, food etc.). In this simple framework, she must know (or, at least, has some expectation) about future earns. However, it might be the case that someone assess it is worth participating in training but has no assets to pursuit that. In this scenario, the problem could be solved by market (credit) or by government (fully provided or subsidized). A hybrid mechanism, like vouchers, can also be used, despite it can emerge some undesirable informational side effects (see HIPP; WARNER (2008); BARNOW (2009); SCHWERDT et al. (2012) and DOERR; STRITTMATTER (2014) for a discussion).

Beside earnings, two additional things should also be considered: skills that some worker has now depreciate over time; and there is also a chance that these skills become obsolete in the future. Skills depreciation affects future earnings over the time, with some rate, while obsolescence can end them, with some probability. These facts must also be taken into consideration while deciding if it is worth engaging in training.

Regarding training types, if there are different options of training, it could be the case that a too specific training gives skills that only will be useful to perform tasks in a specific firm or, perhaps, in firms of the same industry. So, for some people, it might not be interesting acquiring a too specific skill if there is a high probability of that skill will become obsolete.

From the firm's point of view, in a first look, considering that workers who participate in training programs can be more productive, there might be incentives to encourage participation. On the other hand, if this training is also useful for another firms and there is a

[53] More information about Pronatec program design in be provided in next section.
[54] More details below, in section 3.3.

high probability of employees change jobs after conclusion, firms could face lack of incentive in provide training. Further, the existence of competitive and non-competitive markets (ACEMOGLU; PISCHKE, 1999) and the presence of government training offers (GÖRG; STROBL, 2006) can also play a role in firms' preference for training. For example, government training offer may substitute initiatives that would be undertaken by firms anyway.

When some worker is unemployed, general ideas of this framework are still valid, but there are other features to be considered. She must put part of her time in job search activities, which means that the cost to participate in training increases. Putting more time in job search activities could increase the chances of getting job interviews. However, if she decided to engage in training (maybe to increase chances to find a job or a better occupation), she will put less (or no effort) in job search (a "lock-in" period) while participating (LECHNER; MIQUEL; WUNSCH, 2007; ROSHOLM; SKIPPER, 2009). It may be optimal that, if she chooses to take the job training, it happens in the begging of unemployment spell, to reduce opportunity costs. Here, the length of training is important, mostly for those without income assistance when unemployed. A time to end assistance or a threat of cessation can also influence decision (see DOLTON; O'NEILL (1996) and ROSHOLM (2008) for a discussion).

In a scenario with a variety of training options, workers who intend to change job (i.e., acquiring new skills) don't know in advance if they have the right profile for the new job they are prospecting. Also, they might not know the marketing condition of this new profession, considering the lack of experience (although some expectation can be put on that). At this point, asymmetric information may have an important role, where caseworkers could help to mitigate that.

Finally, while making this decision, workers should consider other workers decisions. For example, if a large fraction of labor force engages on training for a specific job or skill, there might be an excess of offer, bringing wages down. Thus, evaluation of earnings before the training may be misleading (LECHNER; MIQUEL; WUNSCH, 2007). Also, untrained workers might see their wages raise due to the relative difference between trained and untrained workers. From the labor demandants point of view, they might have incentives to substitute untrained for trained works, due to changes in relative prices. If it happens, evaluations anchored on Stable Unit Treatment Value Assumption (SUTVA) could not hold. So, spillover effects must be considered.

3.2.2 The Pronatec program

Professional education is in Brazilian 1988's Constitution and also in a specific law about its educational system since 1996[55]. The *Programa Nacional de Acesso ao Ensino Técnico e Emprego* (Pronatec)[56] was created by Federal government in October 26th, 2011[57]. The objectives of Pronatec were expand and democratize technical education (high school level), vocational/training courses (in general, any level of schooling) and professional qualification. Also, improve school-to-work transition, public professional education, employment programs and public high school overall quality[58]. The target groups were public high school students[59], workers, beneficiaries of cash transfer programs and former public high school concluders (or also those from private school with scholarships), evidencing a clearly assistentialist vein. In broad lines, the program was designed with two main fronts: **i)** *Formação Inicial e Continuada – FIC*[60], a kind of short duration standardized courses (usually, from 160 to 400 hours, with three to six months length) cataloged by *Ministério da Educação (MEC)*[61] and without, in some cases, prior education requirements; **ii)** and *Technical Courses*, with a longer duration (from 800 to 1200 hours, with one to three years length), also regulated by MEC and with a prior high school level requirement[62].

In a narrower view, the program had five action plans (SENADO, 2017): **i) expansion of federal professional and technological network**, to enlarge and create schools offering technical education; **ii) Brasil profissionalizado (Professionalized Brazil)**, to increase integrated professional education in high schools; **iii) free agreement with S system**[63], to

[55]Law 9,394 (20 December 1996), known as *Lei de Diretrizes e Bases* (Education Guidelines and Bases Act), updated, in the technical education matter by the Law 11,741 (16 July 2008) and once again by the Law 13,415 (16 February 2017).

[56]National Program for Access to Technical Education and Employment, in English.

[57]The program's creation Law is available in http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12513.htm

[58]When the program was created, the unemployment rate ended 2011 in 4.7%, according to the *Pesquisa Mensal de Emprego – PME* (Monthly Employment Survey), from *Instituto Brasileiro de Geografia e Estatística – IBGE* (Brazilian Institute of Geography and Statistics), a survey covering metropolitan regions. It was in a falling trajectory until 2014 (reaching 4.3%). Since then, the unemployment rate rose to 8.2% in February 2016. The last month available of PME, which was replaced by the *Pesquisa Nacional por Amostra de Domicílios Contínua – PNADC* (Continuous National Household Survey), also from IBGE. The two surveys coexisted since 2012. According to PNADC, there were 7.5 million people unemployed in the first quarter of 2012, 6.4 million in the fourth quarter of 2014 and 11 million in the first quarter of 2016. It means that the program existed in two very distinct scenarios, i.e., economic boom and contraction.

[59]In Brazil, there is a modality of schooling for young and adults in school lag, called *Educação de Jovens e Adultos – EJA* – (Young and Adults Education), which was also included as target group.

[60]Initial and Continued Formation, in English.

[61]Ministry of Education.

[62]Technical education can be organized together with secondary school level, being integrated, concomitant or sequential to high school in Brazil.

[63]National Services of Apprenticeship, quasi-governmental organizations with specialized unities for

increase free courses offer; **iv) E-tec**, to expand distance education courses; and **v) Bolsa formação (scholarship-formation)**, to expand technical education and FIC courses in public and private institutions, including SNA (National Services of Apprenticeship) and Universities.

Most courses of Pronatec are FIC, being offered directly by public institutions (like universities, technological institutes, or national service of apprenticeship) or by private institutions (through agreements with MEC). Government ministries, like, and most commonly, *Ministério do Desenvolvimento Social e Combate à Fome (MDS)*[64], *Ministério do Trabalho*[65], *Ministério do Turismo*[66], *Ministério do Desenvolvimento Industrial e Comércio Exterior (MDIC)*[67], demand specific courses, as well States and Federal District Education Secretaries.

The general guideline for courses creation is to analyze job demand needs to attend them (trying to minimize the matching problem in the labor market). This task is performed by MEC, based on several criteria (working age population, needs of workers by municipality labor market, capacity to offer the courses etc.) to create a Specific Demand Map, which is consolidated in an Identified Demand Map after MEC's analysis. Thus, only vacancies approved by MEC are offered in the Sistec[68], which controls all courses offered in Brazil. However, the amount paid by government for course offer is the same (R$ 10 per hourly class – US$ 2 – since July 2012), regardless the type. This was pointed out as a flaw by *Controladoria Geral da União* (CGU)[69] and was never changed by MEC[70]. This amount must also cover assistance to participants, like transportation and food stipends. According to MEC, there was not much variation in costs by courses that could justify different values.

Due to this fact, in the beginning of the program, most courses offered were less expensive or already being offered by institutions (the so called "shelf courses"). This fact may lead to an oversupply of some kinds of training (the cheaper ones), increasing the number of workers

industry, commerce, rural, transportation and cooperativeness.

[64]Ministry of Social Development and Fight against Hunger, name at that time. Today, the same ministry is called *Ministério da Cidadania* (Ministry of Citizenship).

[65]Ministry of Labor. This ministry no longer exists, being absorbed as a secretary by the new *Ministério da Economia* (Ministry of Economy).

[66]Ministry of Tourism.

[67]Ministry of Development, Industry and Foreign Trade. This ministry is a special case, since it is a unique demand driven courses requester, making a bridge between firms needs and the program offer. O'CONNELL et al. (2017) evaluate Pronatec focusing in MDIC, finding 8.6% increasing in employment for participants indicated by this Ministry. QUINTANA; CRAVO (2019) also evaluate the same program, finding positive effects.

[68]*Sistema Nacional de Informações da Educação Profissional e Tecnológica* (National System of Professional and Technological Education Information).

[69]Audit Office of Federal Government, in English.

[70]TCU (2015a); TCU (2015b)

with correlated skills in the local markets. So, spillover effects are a possibility and must be considered in the analysis, which is exactly one of our contributions.

3.3 Databases: Pronatec and RAIS

We initiate our analysis with over 4.7 million observations, all of them regarding the *Bolsa formação* arm, specifically to workers, consisting almost 50% of the program enrollment and 30% of total budget, covering the 2011-2015 period. Table 24 shows the entire Pronatec major figures for 2011-2016 period.

Table 24: Pronatec program – application and investment by initiative, 2012-2016

Initiative	Enrollment	Investment (R$ bi)	Enrollment (%)	Investment (%)
Expansion of federal network	857,373	6.5	9	16
Brasil profissionalizado	537,032	1.3	6	3
Free agreement with S system	3,252,767	18.7	33	47
E-tec	423,106	0.7	4	2
Bolsa formação	4,657,583	11.3	48	29
Total	9,727,861	38.5	100	97

Source: SIMEC extracted from Senate report (Senado, 2017)

Due to some inconsistencies regarding the initial date of the course and its end date (or expected end date) and enrollment status, around 30,000 observations were dropped. Also, since some people could take more than one course[71], we considered only the first time of them in the program, dropping around 700,000 observations and keeping over 3,9 million distinct individuals[72]. Workers apply to Pronatec, with a "first come, first served" system, which means that not everyone ensure a sit in the classroom. Taking advantage of this rule, we construct a control group with people who did not enter in the program (so, against their will), called "involuntary" non-participants[73]. There are also people who were applicants in Pronatec, but did not confirm their participation, called here "voluntary" non-participants. We advocate that this rule creates a quasi-experimental setup, allowing us to identify the true

[71]Not, however, simultaneously, nor more than three in the same year.

[72]The normal case is one subscription per person in the program. We consider that an additional course is less important when someone already participated once in the Pronatec.

[73]"First come, first served" system may cause selection bias, since first applicants may be different from last applicants. Unfortunately, we do not have the application order to control for this possible bias selection. However, situations as cancellation of classes could minimize this issue.

effects, dealing with possible self-selection and non-observable confounders (see HECKMAN; SMITH (2004) about selection into participation issues). Also, our huge data base, with all participants of *Bolsa fomação*, give us statistic power to estimate impacts with high precision (ANDERSSON et al., 2013), allowing us to look for heterogeneous effects and test robustness in several subsets.

All courses are fully free of charges, existing a pecuniary incentive to participants, covering transportation and food expenses (around R$ 2,00 – US$ 0.39 – per hourly class). Table 25 summarizes the number of participants by statuses.

Table 25: Pronatec program – application statuses, 2011-2015

Status	N	%
Class ended	103,396	2.63
Concluded	1,845,835	47.04
Failed	254,138	6.48
Non-participants - involuntary	155,303	3.96
Non-participants - voluntary	590,991	15.06
On course	157,074	4.00
Others	16,077	0.41
Quitted	801,152	20.42
Total	3,923,966	100.00

Source: Sistec

Obs.1: Quitters are people who abandoned, dropped or did not show up in class after subscription confirmation.

Obs.2: Involuntary control are people who class or the subscription was cancelled before the beginning of classes, had lack of documentation or oversubscription.

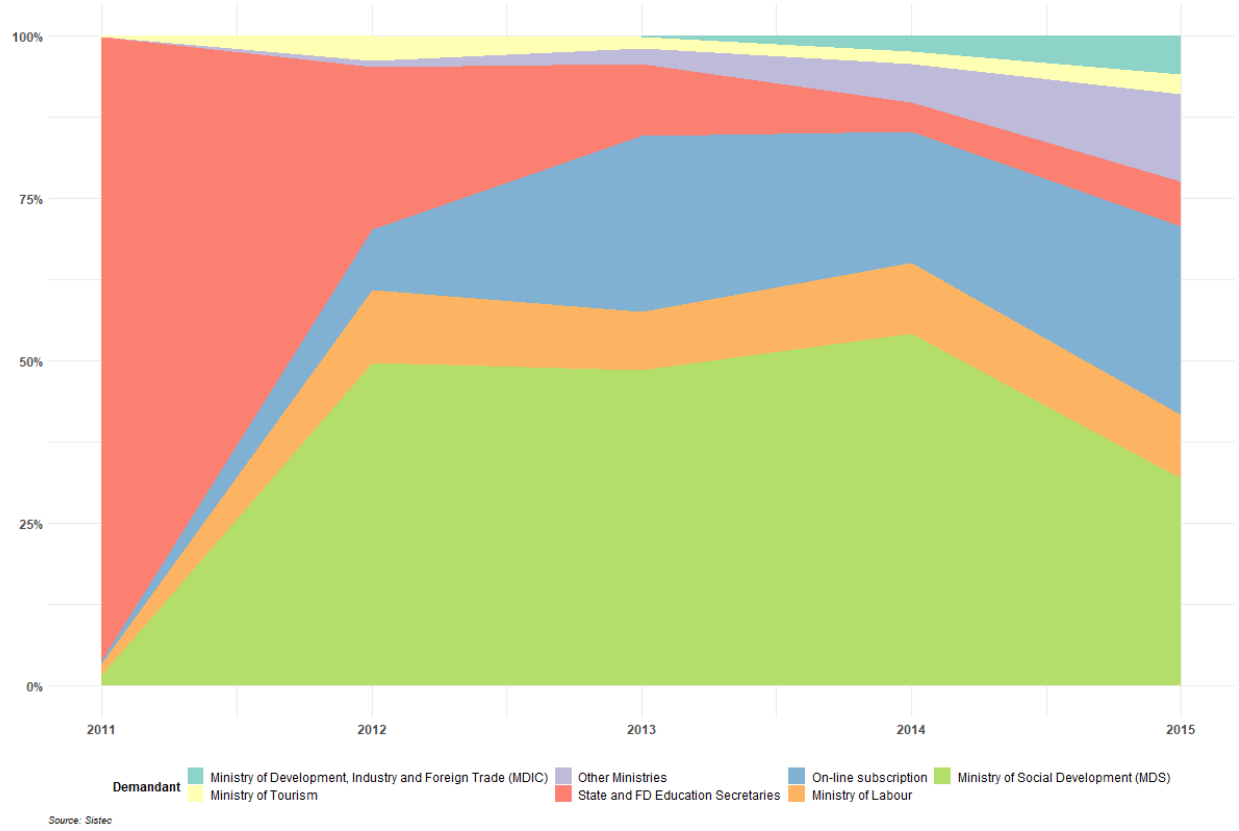
Obs.3: Voluntary control are people who did not show up to confirm subscription

We note that, for treated, i.e., people who had subscription confirmed, around 47% concluded the course (1,845,835 people). Quitters sums 20%, failed are 6.5%, still on course 4% and around 2.6% had their classes canceled after start[74]. For non-participants, i.e., people who did not have their application confirmed, around 79% was due to voluntary reasons (did not show up), while 21% was due to reasons out of their control (155,303 people).

The source of courses demand is an important feature of the program, which is presented in Figure 9 (and Table 62 in Appendix). For Pronatec it could be a relevant source of selection into the program.

[74]This group could be considered a kind of control group, since they were not able to keep the course against their will. However, since they initiate the course, they are somehow different from those who never had the chance to initiate it. Also, they must be allocated in another class when this situation happens.

Figure 9: Pronatec program – subscriptions by demandant, 2011-2015



Almost half of solicitations were made by MDS, which focuses on poor people. This is a special case, since there are possibly caseworkers directing people into the program at *Centros de Referência de Assistência Social* (CRAS)[75] and *Centros Especializados de Referência em Assistência Social* (CREAS)[76], the network of social assistance in Brazil. On-line subscriptions were almost 21%, while State and FD Education Secretaries sum 12%. MDIC, the "demand driver requester", was only 1.2% of enrollments. Therefore, we notice that almost 3/4 of this program's arm was demanded by MDS or directly by participants on-line, a relevant information regarding selection.

Another important feature is the technological axis of the courses[77], showed in Figure 20 (and Table 63 in Appendix). Most courses are in "Business and Management" axis (around 26%), followed by "Industrial Control and Processes" (around 14%), "Information and Communication" (10.5%) and "Infrastructure" (10.2%). Security axis has less than 1% of

[75]References Centers of Social Assistance.

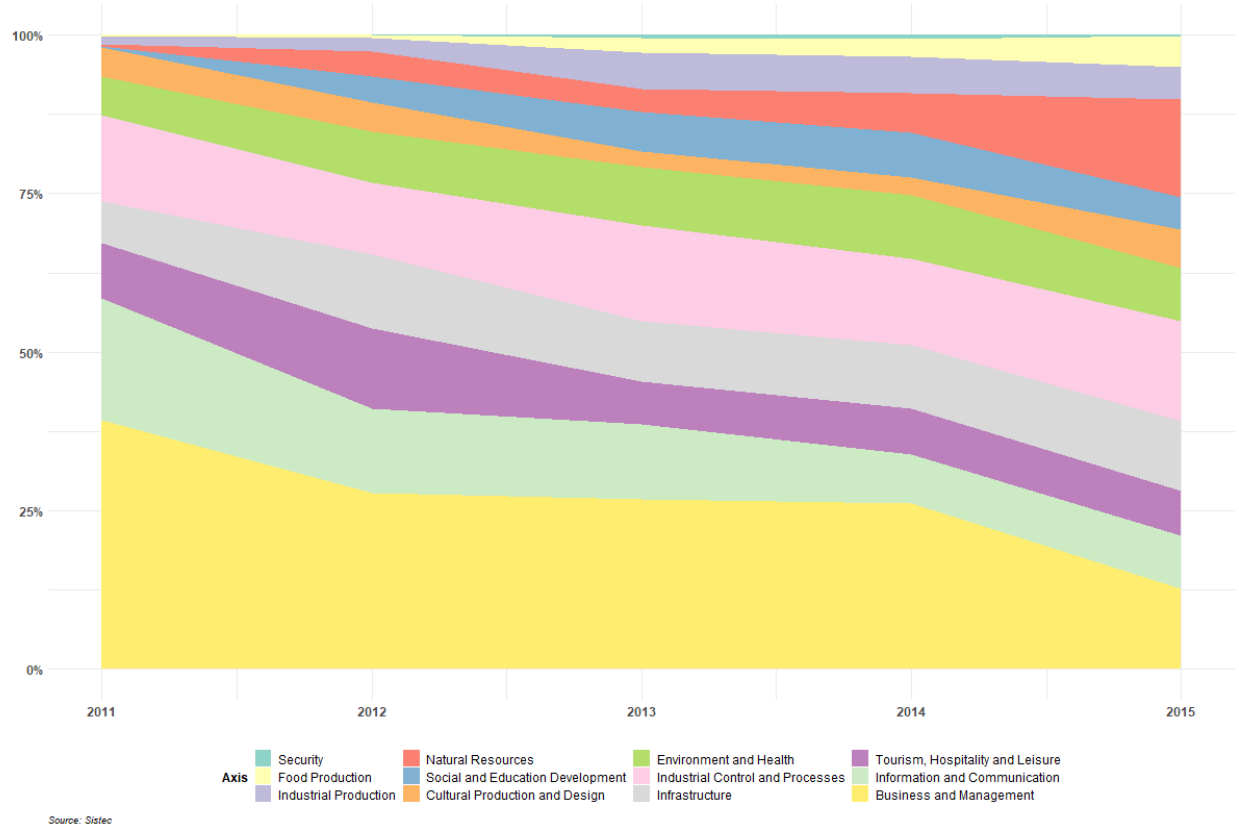
[76]Specialized Reference Centers of Social Assistance.

[77]MEC divides FIC courses in twelve axis, organized in a guide. The last FIC's guide, organized in 2016, sums 646 courses in 12 axes and is available in https://map.mec.gov.br/attachments/74900/guia_pronatec_de_cursos_fic_2016.pdf

the enrollments, with "Food Production", "Cultural Production and Design" and "Industrial Production" with less than 5%.

The evolution of course offers, by axis and demandant, however, was not constant over time. For example, "Business and Management" axis had an increase in number of applicants over the years, but in a lower rate than the other axes. The same is true for demandants, where State and FD Education Secretaries initiate the offer (due to past experiences) until the Ministries organize their offer. Overall, there was an increase of course offer from 2011 to 2014, by axis and demandant, but at different rates.

Figure 10: Pronatec program - subscriptions by axis, 2011-2015



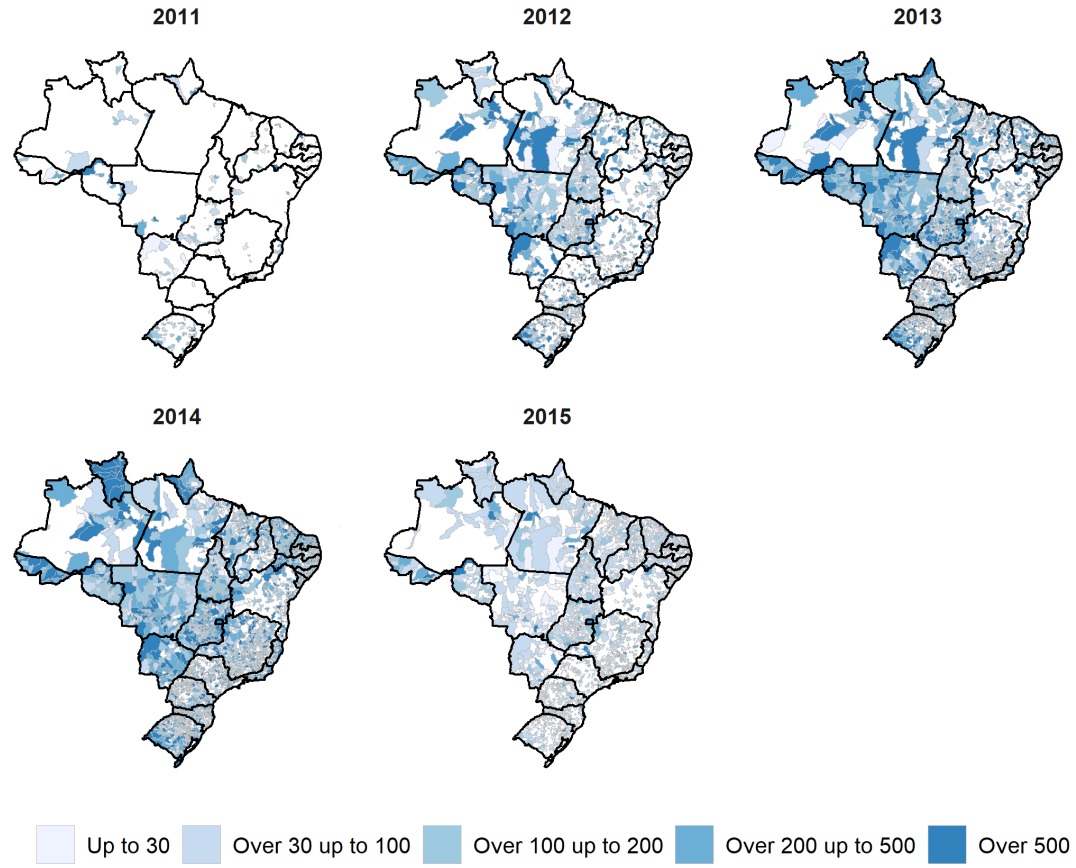
Additionally, Brazil is a continental country[78], with historical inequalities in wealth, income and opportunities distribution, where the north and northeast cities are poorer than, especially, south and southeast (NETO, 2009)[79]. So, when public programs are designed, interiorization and prioritization of such areas are usually present. Figure 11 shows that

[78]Brazil has 8,516,000 km^2 of area, the fifth in the world.

[79]Brazil is divided in five big regions to facilitate public polices design and analysis: north, northeast, center-east, southeast and south.

it was the case with Pronatec, reaching almost the majority municipalities, mainly in the poorest regions.

Figure 11: Pronatec subscriptions by municipality and year, 2011-2015



Source: Sistec and IBGE

Our data set allow us to control for different aspects of training, course length and the institutions demanding courses. Most of previous studies treats training aggregately making no difference regarding the type of course. Here, we can add controls for twelve fields areas (axes). Also, we have the course length in hours and the demandants (course requesters), giving us best control of heterogeneity.

To construct the outcomes, we use the *Relação Anual de Informações Sociais* (RAIS)[80] from 2010 to 2019, a rich administrative data base with information of Brazilian's formal labor market, that amounts over 65 million records in 2019[81]. All formal firms that has relationship with registered employees in the fiscal year (hiring or firing) are obligated to

[80] Annual Relation of Social Information.
[81] The size for each year varies between 55 and 65 million observations.

report it to the Secretary of Labor[82]. This database includes a set of variables about workers, like gender, wages, hours contracted, schooling, occupation, and about the firm, like number of employees, main activity, address, and legal nature. We link Pronatec participants to RAIS using social security numbers[83], present in both data bases.

We also use the National Guide of FIC courses to link the number of occupations existing in a municipality in a specific axis, adding the amount of Pronatec concluders in the same axis, to control for impacts in local labor market (see FERRACCI; JOLIVET; BERG (2014) about the spillovers of training in labor markets). Here, we consider every axis (twelve) in a municipality (5,570 since 2013) as a separated labor market. Thus, we have the following control variable:

$$Pronatec.share_{y,m,a} = \frac{\sum P_{y,m,a}}{\sum E_{y,m,a} + \sum_a P_{a,m,a}} \tag{2}$$

Where $Pronatec.share_{y,m,a}$ is the Pronatec share of workers in the year y, in the municipality m and in the axis a. $P_{y,m,a}$ is the number of Pronatec concluders in the year y, in municipality m, in axis a and $E_{y,m,a}$ is the number of formal employees in the year y, in municipality m and in axis a.

To further control for impacts of Pronatec in local formal labor market, we construct an indicator based on some job offer. Using the data from *Sistema Nacional de Emprego* (SINE)[84], we create the ratio of the number of Pronatec trained and the number of jobs offered in the system, by axis, municipality and year (equation 3).

$$Job.offer_{y,m,a} = \frac{\sum P_{y,m,a}}{\sum S_{y,m,a}} \tag{3}$$

Where $Job.offer_{y,m,a}$ is the job offer in the year y, in the municipality m and in the axis a. $P_{y,m,a}$ is same specified before and $S_{y,m,a}$ is the number of formal jobs offered by SINE in the year y, in municipality m and in axis a.

Based on this ratio, for each axis and year, we create an indicator variable for municipalities above de median (equation 4).

[82] Former Ministry of Labor.

[83] *Cadastro de Pessoa Física (CPF).*

[84] National System of Employment, was created in 1975, as a public service, to help matching in labor market, registering job offers and workers.

$$Job.index = \begin{cases} 1 & if \quad Job.offer_{y,m,a} > \widetilde{Job.offer}_{y,m,a} \\ 0 & otherwise \end{cases} \tag{4}$$

Where $\widetilde{Job.offer}_{y,m,a}$ is the median of $Job.offer_{y,m,a}$.

Additionally, we consider another control variable, looking to the average hourly wage in the year before of training starts (equation 5):

$$Pronatec.wage_{y-1,m,a} = \frac{\sum_i \frac{W_{y-1,m,a,i}}{H_{y-1,m,a,i}}}{\sum E_{y-1,m,a}} \tag{5}$$

Where $Pronatec.wage_{y-1,m,a}$ is the average wage per hour in year $y-1$, in municipality m in the axis a, $W_{y-1,m,a,i}$ is the wage of the worker i in the year $y-1$, municipality m and axis a, while $H_{y-1,m,a,i}$ is the number of hours in the labor contract of worker i, in year $y-1$, municipality m and axis a. $E_{y,m,a}$ is the same defined before, with a year lag.

The wage level in a market is important for several reasons. For example, high wages in a specific market may means excess of demand or lack of supply, with Pronatec offer of training possibly altering workers and firm's decision and, hence, equilibrium. On the other hand, low wages may mean the opposite, with Pronatec offer of training pushing further down wages.

Thus, we can control for three important features of labor market equilibrium: share of workers, job offer and wages per hour (in lag), in every market (municipality and axis, by year of course conclusion).

The empirical strategic relies on the following reduced model:

$$Y_{ma} = \alpha + \delta_1 T + \delta_2 Job.index + \delta_3 Pronatec.share + \delta_4 Pronatec.wage + \beta_i \boldsymbol{X_i} + \varepsilon \tag{6}$$

Where Y_{ma} is a dummy for employed in formal labor market *ma months after* Pronatec conclusion (six, twelve, eighteen, twenty-four, thirty-six, forty-eight, sixty, seventy-two and eighty-four) or the real hourly wage (in natural log)[85], T is a dummy for Pronatec concluders, $Job.index$, $Pronatec.share$ and $Pronatec.wage$ are the control variables for labor market defined before, ε is the error term and $\boldsymbol{X_i}$ are additional controls (year of training conclusion,

[85]If a person is not in RAIS in some reference, so she has no wage from formal labor market. We adjust as $log(1) = 0$.

state – 26 dummies, axis – eleven dummies, demandant – 9 dummies, schooling – 6 dummies, age (and its square), length of course – in hours, gender, presence of Federal Institutes of Technology in the city[86], GDP *per capita*, share of PBF and BPC programs in population[87]. GDP and population size were collected from IBGE while BPF and BPC were collected from Ministry of Citizenship. To consider spatial dependency, we control for the distance between the center of municipality and the center of the State capital. Table 26 summarizes all variables that will be used to estimate the effects.

[86] According to Census of Superior Education (*Censo da Educação Superior*), from the National Institute of Educational Studies and Researches (*Instituto Nacional de Estudos e Pesqusias Educacionais Anísio Teixeira – INEP*), there were 38 institutes in Brazil from 2012 to 2016.

[87] *"Programa Bolsa Família"* (PBF), organized by Ministry of Citizenship is one of the biggest conditional cash transfer programs in the world. The target are families under the extreme poverty and poverty lines (in 2020, families earning up to R$ 89 by person, or U$ 17, by month are considered extremely poor, while families above that amount and up to R$ 178, or U$ 35, are considered poor), focused one children. As counterpart, school attendance and vaccination are required. PBF reaches almost 14 million families in Brazil in 2020. On the other hand, BPC (*"Benefício de Prestação Continuada"*) is a program for elderly and handicapped. The poor population in this profile (people aged 65 or over and all handicapped) are eligible for a monthly minimum wage paycheck (R$ 1.045, or U$ 205, in 2020).

Table 26: Variables: description and source, by category

Category	Variable	Description	Source
Outcome	Employment	A dummy for employment in formal labor market m months after course conclusion	RAIS
	Wages	Real wage in formal labor market m months after course conclusion	RAIS
Main regressors	Treated	A dummy for Pronatec concluders	Sistec
	Job index	A dummy for municipalities above the median in ratio of Pronatec trained and job offer	Sistec, RAIS and SINE
	Pronatec share	Share of Pronatec concluders in a market (year, municipality and axis)	Sistec and RAIS
	Pronatec wage	Lag of average hourly wage in a market (previous year of course start, municipality and axis)	Sistec and RAIS
Controls	Year of conclusion	Dummy for year of training conclusion	Sistec
	Schooling	Dummies for schooling: College (complete and incomplete), High school (complete and incomplete), Elementary school (complete and incomplete) and No schooling.	Sistec
	Axis	Dummy for the course axis.	Sistec
	Demandant	Dummy for course demandant.	Sistec
	Course length	Course length in hours	Sistec
	Age	Age informed in application and its square	Sistec
	Gender	Dummy for gender	Sistec
	Time unemployed	Last wage for those who lost the job within 6 months before course start	RAIS
	Last wage	Unemployment spell for those who lost the job within 6 months before course start	RAIS
	State	Dummy for the 26 States and the Federal District (UF)	Sistec
	Distance	Distance from municipality to the State capital	Authors calc.
	Federal Institute	Dummy for existence of Federal Institute of Technology in the municipality	INEP
	GDP p.c.	GDP per capita	IBGE
	BPC	Beneficiaries share of BPC program in population	MC and IBGE
	PBF	Beneficiaries share of PBF program in population	MC and IBGE

Source: RAIS, Sistec, IBGE and MC (Ministry of Citizenship)

We estimate equation 6 with a *logit*[88], when the outcome is employment, and when the outcome is wage, by OLS. Wages are in real values, according to the official inflation index, using March 2020 as reference[89].

3.3.1 Subsets for estimation and descriptive statistics

The main challenge of this analysis is to create credible comparable groups of Pronatec participants and non-participants. Since these groups might not be similar in several characteristics, we will focus firstly on participants who lost their jobs within six months before course starts (see ASHENFELTER, 1978 about the "earning dip" before job training programs) (**subset A**). Then, we filter our data further to only involuntary non-participants as control (our "quasi-experimental" **subset B**). Treated are always people who concluded training.

With these subsets we are able to control for the last wage and unemployment spell, as well identify job relationships in formal labor market in all years.[90] As a robustness exercise, we will consider matching[91] data of both subsets. Additionally, we will run an instrumental variable regression for subset A, using as instrument a dummy for involuntary non-participation (for Pronatec conclusion), and its share in a market (for the Pronatec share variable).

Table 27 summarizes each subset and sample size, by treated and control (Figure 14 in Appendix presents this information schematically).

[88] TRAIN (2009) provides an excellent review of discrete choice methods and how *logit's* flexibility can be used in problems like this.

[89] *Índice Nacional de Preços ao Consumidor Amplo* (IPCA) is the official Brazilian inflation index, calculated by IBGE.

[90] Since we are using an administrative information of formal labor market, we know that if someone does not appear in a specific year in RAIS it is because she is out of formal labor market. But we do not know if she is self-employed, working informally or has any other activity, neither any other information about her work life.

[91] More details about matching procedure will be given in section 3.4.2.

Table 27: Data sets used in analysis: description and number of observations.

Data set	Description	Number of observations
Subset A	Applicants to the Pronatec Program who lost their formal jobs within six months before the training starts:"concluders" as treated and all other applicants as control.	Treated = 239,880; Control = 400,390
Subset B	Previous data set filtered to a quasi-experimental scenario: "concluders" as treated and "involuntary" non-participants as control	Treated = 239,880; Control = 39,946
Matching A	Matching (propensity score) database of subset A	Treated = 239,064; Control = 238,969
Matching B	Matching (propensity score) database of subset B	Treated = 39,781; Control = 39,729

Source: Sistec and RAIS.
Obs.: Data from 2011 to 2015.

As mentioned before, we will condition on observables available in Pronatec database to control for possible confounders: age (and its square, to capture non-linearities), course axis, course demandant, schooling[92] and gender. Also, we consider municipality controls: GDP *per capita*, PBF and BPC shares in population, presence of federal institute of technology and distance to the state capital. The year of course conclusion (or expected year of conclusion) will be considered to control for year specificities.

Table 28 presents descriptive statistics for the subset A (i.e. concluders as treated and all non-concluders or non-participants as control, considering only who lost their job within six months before training starts or should starts), with means (continuous variables) or proportions (dummy variables). Also, the same table presents some market characteristics (regarding municipalities) that will be, in general, common to both subsets.

[92]In Brazil, basic school is mandatory from 4 to 17 years-old, divided in three stages: pre-school (*ensino infantil*, children between 4 and 5 years), elementary school (*ensino fundamental*, children between 6 to 14 years) and high school (*ensino médio*, adolescents between 15 and 17 years). After that, there is college, masters and doctoral. We divide elementary school, high school, and college in "complete" or "incomplete" and no schooling, resulting in seven categories.

Table 28: Descriptive statistics for municipalities and labor market, 2011-2015, subset A

Type	Variable	Mean	SD	Min	Max
Market	Distance (to State capital)	231.24	162.16	0.00	1,485.38
	Job offer index (%)	26.04	39.72	0.00	100.00
	Federal Institute of Technology (%)	1.69	12.90	0.00	100.00
	BPC (%)	1.93	1.21	0.00	13.14
	PBF (%)	8.76	5.33	0.17	24.68
	Pronatec.share (%)	14.18	19.47	0.00	100.00
	GDP (per capita)	22	23	4	816
	Pronatec.wage	11.26	3.86	0.00	107.02
Individual	Time unemployed	2.50	1.53	0.00	6.00
	Previous wage	7.61	5.38	0.38	74.99
	Age	30.98	8.94	14.00	96.00
	Course length (hours)	193.94	45.53	160.00	400.00
	Business and Management (%)	30.59			
	Cultural Production and Design (%)	2.52			
	Environment and Health (%)	6.31			
	Food Production (%)	1.62			
	Industrial Control and Processes (%)	18.77			
	Industrial Production (%)	4.53			
	Information and Communication (%)	8.98			
	Infrastructure (%)	12.78			
	Natural Resources (%)	1.3			
	Security (%)	0.48			
	Social and Education Development (%)	4.3			
	Tourism, Hospitality and Leisure (%)	7.82			
	Min. of Dev., Ind. and F. Trade (MDIC) (%)	0.68			
	Ministry of Labor (%)	45.21			
	Ministry of Social Development (MDS) (%)	33.81			
	Ministry of Tourism (%)	1.16			
	On-line subscription (%)	15.83			
	Other Ministries (%)	0.91			
	State and FD Education Secretaries (%)	2.4			
	College (complete) (%)	2.1			
	College (incomplete) (%)	3.91			
	Elementary school (complete) (%)	10.39			
	Elementary school (incomplete) (%)	10.84			
	High school (complete) (%)	54.28			
	High school (incomplete) (%)	18.21			
	No schooling (%)	0.26			
	Female (%)	45.24			

Source: Sistec, IBGE, ME and MC.

Obs.1: Means for continuous variables and proportions for dummies.

Obs.2: Distance in km. SD = Standard Deviation.

We notice that PBF covers more people than BPC in average (8.7%), and there are municipalities where program reaches almost a quarter of population. Pronatec share indicates that program affects local formal labor market, since it represents, in average, around 14% of their markets (i.e., in a specific axis and municipality). We see that Pronatec could also "create" a specific market (the only potential workers for that market are Pronatec concluders), while Job index are above median for one quarter of municipalities. GDP *per capita* has a wide range as well past Pronatec wages (hourly). Finally, we see that federal institutes of technology are present in a small fraction of municipalities and there are a wide range regarding distance to State capital[93].

Regarding individual characteristics, we notice that the majority of people are male, in the beginning of adult life, with high school, applied to business and management courses, with a length around 190 hours, requested by Ministry of Labor. For those who had previous information of labor market, the average unemployment spell before Pronatec application was 2.5 months, with an average wage of R$ 7.61 hourly.

Table 29 shows the means and proportions difference test between treated and control, for each variable, by subset. We also present difference test for a random sample taken from RAIS, to assess if it would be a potential alternative control group.

[93]Brazil has 8,515,767 km^2, the fifth largest country in the world.

Table 29: Mean/proportion difference between Treated and Control, by data set, 2011-2015

Variable	Subset A	Subset B	Random RAIS
Age	0.350***	-0.960***	-1.550***
Business and Management (%)	4.410***	1.180***	
Cultural Production and Design (%)	-0.520***	0.030	
Environment and Health (%)	0.560***	1.040***	
Food Production (%)	-0.200***	-0.410***	
Industrial Control and Processes (%)	-1.250***	-0.800***	
Industrial Production (%)	0.080	0.420***	
Information and Communication (%)	-1.210***	0.000	
Infrastructure (%)	1.240***	3.240***	
Natural Resources (%)	0.210***	0.900***	
Security (%)	0.040**	0.100**	
Social and Education Development (%)	-1.000***	-1.560***	
Tourism, Hospitality and Leisure (%)	-2.330***	-4.120***	
Course length (hours)	-6.420***	-3.920***	
Min. of Dev., Ind. and F. Trade (MDIC) (%)	0.000	0.420***	
Ministry of Labor (%)	-9.220***	-33.190***	
Ministry of Social Development (MDS) (%)	1.640***	11.780***	
Ministry of Tourism (%)	-0.640***	-0.420***	
On-line subscription (%)	8.250***	20.870***	
Other Ministries (%)	-0.020	0.370***	
State and FD Education Secretaries (%)	-0.020	0.170**	
College (complete) (%)	0.080**	-1.210***	-1.060***
College (incomplete) (%)	-0.680***	-2.280***	-4.790***
Elementary school (complete) (%)	-1.870***	-1.070***	-3.890***
Elementary school (incomplete) (%)	-5.580***	-7.680***	-9.340***
High school (complete) (%)	9.670***	11.590***	11.810***
High school (incomplete) (%)	-1.470***	0.770***	7.680***
No schooling (%)	-0.150***	-0.120***	-0.410***
Previous wage	0.110***	-0.040	-3.000***
Female (%)	5.750***	7.460***	12.480***
Time unemployed	-0.020***	0.180***	-0.570***

Source: SIMEC, IBGE, ME and MC.

Obs.1: Means for continuous variables and proportions for dummies.

Obs.2: Null hypothesis is no difference.

Obs.3: * = significant at 10%; ** = significant at 5%; *** = significant at 1%.

Overall, mean and proportion difference between treated and control are, in general, extremely low in absolute values, although there are some important differences. Two demandants, "Ministry of Labor" and "On-line subscription", are underrepresented and

overrepresented, respectively, among treated[94]. There are also differences regarding education, specially "elementary school (incomplete)" and "high school (complete)", with lower and higher proportion for treated, respectively. Proportion and mean difference tests show significant statistical differences in virtually all variables. Considering our sample size, this result is expected, even when means/proportions are closer in their point estimates, which is the case in most variables. The sample taken from RAIS seems not to be better than our subsets, so we will not consider it in results section[95]. It is important to point out that previous wages are statistically the same for both subsets. Nonetheless, covariates are necessary not only due to efficiency purposes, but to avoid eventual biases.

3.4 Results

In this section, we will present results for subsets A and B, looking to employment probability and wages, as well heterogeneous effects and robustness tests (matching databases). We also present an instrumental variable regression. We end with a cost and benefit exercise for matching of subset B.

3.4.1 Subset A

Results of equation 6 are in Table 30. We will present results for nine periods after course conclusion (six, twelve, eighteen, twenty-four, thirty-six, forty-eight, sixty, seventy-two and eighty-four months after). Since the range of course conclusion is 2012 to 2016, each year after the first year of conclusion has fewer observations after three years (RAIS' range is 2010 to 2019). We report the number of observations for each estimation and, for employment probability, odds ratio are reported[96].

[94]Negative differences indicates the control mean/proportion is higher.

[95]Results, not reported, suggests even stronger effects where this sample is considered.

[96]The interpretation of odds ratio x is: values above unity means $1 - x$ higher odds than base group; values below unity means $x - 1$ lower odds while unity indicates equal odds.

Table 30: Subset A – Pronatec impact on employment after 6 to 84 months of course conclusion

Variable	6 months	12 months	18 months	24 months	36 months	48 months	60 months	72 months	84 months
Treated (O.R.)	1.010	1.117	1.156	1.159	1.181	1.178	1.176	1.172	1.182
	(0.005)*	(0.005)***	(0.005)***	(0.005)***	(0.005)***	(0.005)***	(0.006)***	(0.008)***	(0.030)***
Job index	0.013	0.015	0.018	0.007	0.008	0.011	0.027	0.031	0.062
	(0.006)**	(0.006)***	(0.006)***	(0.006)	(0.006)	(0.006)*	(0.006)***	(0.009)***	(0.032)*
Pronatec.share	-0.282	-0.307	-0.267	-0.203	-0.099	-0.041	-0.024	0.018	0.405
	(0.038)***	(0.037)***	(0.037)***	(0.037)***	(0.037)***	(0.038)	(0.040)	(0.067)	(0.452)
Pronatec.wages	-0.001	-0.003	-0.004	-0.002	-0.004	-0.003	-0.003	0.001	-0.008
	(0.001)	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)	(0.004)*
N. obs.	640,270	640,270	640,270	640,270	640,270	622,225	567,059	272,755	26,294

Obs.1: O.R.= Odds ratio. Standard Deviation in parenthesis.

Obs.2: * = significant at 10%; ** = significant at 5%; *** = significant at 1%

Obs.3: Standard errors clustered by State.

We notice a very same chance to be employed just six months after course conclusion between treated and control (just 1%, significant only at 10%). Although not negative, this result could be interpreted as a kind of lock-in effect, where training did not improve chances of holding a formal job just after conclusion. A year after, return goes up to 11.7%, increasing until 36 months after (18.1%). Then, effects vary around this level, ending with the highest magnitude (18.2%). It is important to notice, however, that only a small fraction of concluders had time to reach seven years of training conclusion.

Job index always presents a positive effect, but barely significant in some periods (like 48 and 84 months after). Pronatec share effect is negative until the third year after conclusion, fading away during time. In the fourth year onward, estimates are not different from zero. Pronatec wages, on the other hand, is almost always negative (except for the first and sixth year) and relatively constant over time. All these results together suggest that stable unit treatment value assumption (SUTVA) may not be valid once the share of concluders matters for employment probability. This result is in accordance to FERRACCI; JOLIVET; BERG (2014), so, the greater the share and wages of Pronatec concluders are, the lower is the chance to be employed in the formal labor market, with more lasting effects for wages control variable.

Looking to the impact on wages (Table 31), we also see a positive effect, but only a year after, with the lowest significant return estimated (6.1%)[97]. For the rest of periods, the average return is 9.3%, with little variation. The peak is reached in the last period, 10%, but

[97]Due to log transformation, actual effect is $e^{\beta} - 1$.

estimated with the smallest sample. The job index now is not significant in seven out the nine periods, and only significant at 5% in one of them, which suggests that its relevance is more attached to employment status than wages. Pronatec share shows a similar behavior as aforementioned, while Pronatec wages are barely significant with a mixture of signs, always with a lower magnitude.

Table 31: Subset A – Pronatec impact on wages after 6 to 84 months of course conclusion

term	6 months	12 months	18 months	24 months	36 months	48 months	60 months	72 months	84 months
Treated	0.003	0.059	0.080	0.082	0.093	0.092	0.089	0.090	0.095
	(0.003)	(0.003)***	(0.003)***	(0.003)***	(0.003)***	(0.003)***	(0.003)***	(0.005)***	(0.017)***
Job index	0.000	0.000	0.004	-0.004	-0.002	0.001	0.009	0.011	0.029
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)***	(0.005)**	(0.018)
Pronatec.share	-0.132	-0.147	-0.129	-0.100	-0.049	-0.009	-0.004	0.013	0.274
	(0.019)***	(0.019)***	(0.020)***	(0.020)***	(0.020)**	(0.021)	(0.022)	(0.037)	(0.247)
Pronatec.wages	0.001	0.000	-0.001	0.000	-0.001	-0.001	-0.001	0.001	-0.004
	(0.000)***	(0.000)	(0.000)	(0.000)	(0.000)**	(0.000)	(0.000)**	(0.001)*	(0.002)
N. obs.	640,270	640,270	640,270	640,270	640,270	622,225	567,059	272,755	26,294

Obs.1: Standard Deviation in parenthesis.

Obs.2: * = significant at 10%; ** = significant at 5%; *** = significant at 1%

Obs.3: Standard errors clustered by State.

First results are in accordance with previous studies regarding mid and long run (CARD; KLUVE; WEBER, 2010, 2018; VOOREN et al., 2019), but differs in the shortest run, with a light "lock-in effect". Considering that subset A may carry some self-section into the program (HECKMAN; SMITH, 2004), next section looks to the subset B: our "quasi-experimental" data set, with specific program statuses for control.

3.4.1.1 Subset B: quasi-experimental data base (specific statuses) We now consider as control who did not start the course due to factors out of candidate's control (the "involuntary" control group), keeping as treated concluders. We now have 279826 individuals in total, from which 85.72% were treated. So, our control group in just a fraction of the previous subset (around 10%).

When we look to employment results (Table 32), the effect is stronger and significant since six months after conclusion (9%) as well in the rest of the periods. A year and a half after, effect reaches 19.2% and, thenceforth, remains steady until the last period, the peak, 22.1%. Job index stills positive, but sparsely significant. Pronatec share and Pronatec wages show a similar pattern of the previous subset, but effect of the former vanishes after two

years now. In general, conclusions remain: a higher number of Pronatec concluders in the market as well higher wages make find a job a harder task, particularly in the first two years.

Table 32: Subset B - Pronatec impact on employment after 6 to 84 months of course conclusion

term	6 months	12 months	18 months	24 months	36 months	48 months	60 months	72 months	84 months
Treated	1.090	1.157	1.192	1.194	1.187	1.189	1.192	1.197	1.221
(O.R.)	(0.012)***	(0.012)***	(0.012)***	(0.012)***	(0.012)***	(0.012)***	(0.012)***	(0.020)***	(0.076)***
Job index	0.007	0.013	0.025	0.012	0.017	0.007	0.020	0.058	0.001
	(0.009)	(0.009)	(0.009)***	(0.009)	(0.009)*	(0.009)	(0.009)**	(0.014)***	(0.057)
Pronatec.share	-0.322	-0.341	-0.325	-0.261	-0.075	-0.055	-0.027	0.093	0.811
	(0.053)***	(0.052)***	(0.052)***	(0.052)***	(0.052)	(0.052)	(0.055)	(0.090)	(0.673)
Pronatec.wages	-0.002	-0.005	-0.005	-0.002	-0.004	-0.002	-0.003	0.003	-0.006
	(0.001)*	(0.001)***	(0.001)***	(0.001)*	(0.001)***	(0.001)**	(0.001)**	(0.002)*	(0.007)
N. obs.	279,826	279,826	279,826	279,826	279,826	277,664	254,882	117,308	8,847

Obs.1: O.R.= Odds ratio. Standard Deviation in parenthesis.

Obs.2: * = significant at 10%; ** = significant at 5%; *** = significant at 1%

Obs.3: Standard errors clustered by State.

Looking to wages (Table 33), results are similar, with stronger estimates for this subset and significant since the beginning. The peak is reached in the last but one period (10.7%), however effect is less precisely estimated. After the first year, the average effect until the last period is 10.2%, with little variation over the years. Job index seems not to be important for wages, with closer to zero estimates in virtually all periods. Pronatec share is negative until the second year, becoming inconclusive since then, reinforcing previous conclusion: more trained people by Pronatec in a market is associated with lower wages in the short run, with a fading effect over the years. For Pronatec wages control, we observe virtually no contribution, with estimates always closer to zero, suggesting a limited influence of past market wages levels in individual wages.

Table 33: Subset B - Pronatec impact on wages after 6 to 84 months of course conclusion

term	6 months	12 months	18 months	24 months	36 months	48 months	60 months	72 months	84 months
Treated	0.045	0.081	0.102	0.099	0.093	0.094	0.094	0.107	0.104
	(0.007)***	(0.007)***	(0.007)***	(0.007)***	(0.007)***	(0.007)***	(0.007)***	(0.011)***	(0.043)**
Job index	-0.003	-0.001	0.007	-0.001	0.002	-0.001	0.006	0.027	0.010
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.008)***	(0.033)
Pronatec.share	-0.147	-0.162	-0.162	-0.131	-0.037	-0.025	-0.014	0.047	0.395
	(0.027)***	(0.027)***	(0.027)***	(0.028)***	(0.028)	(0.028)	(0.030)	(0.050)	(0.362)
Pronatec.wages	0.001	-0.001	-0.001	0.000	0.000	0.000	0.000	0.003	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)**	(0.004)
N. obs.	279,826	279,826	279,826	279,826	279,826	277,664	254,882	117,308	8,847

Obs.1: Standard Deviation in parenthesis.

Obs.2: * = significant at 10%; ** = significant at 5%; *** = significant at 1%

Obs.3: Standard errors clustered by State.

In general, with this subset B, we notice that patterns are somewhat the same reported for subset A, however with stronger magnitudes, both for employment probability and wages premium.

3.4.2 Robustness tests: matching data bases and IV regression

In this section, we will consider two robustness tests: a matching of Pronatec subsets A and B; and an IV regression of subset A.

3.4.2.1 Matching of subset A Despite the richness of our data set, it is possible that, even controlling for observables available, controls and treated are not comparable enough. In order minimize eventual differences about groups composition, we create a new subset, matching participants.

It is important to point out that matching procedure is time consuming in large data sets, which is our case here. As strategy to surpass this obstacle, and considering methodological coherence, we apply matching separately by state (27 in Brazil[98]). Conducting matching by state, we consider more homogeneous groups, getting better matches with smaller subgroups. We use the nearest neighbor method, with a logit distance (propensity score)[99], using all the covariates mentioned before[100]. Figure 15 (Appendix) presents densities distribution of

[98] Actually, Brazil has 26 states and one Federal District, that has similar state's attributions.

[99] Matching are conduct with "MatchIt" R package (HO et al., 2018).

[100] For more details about matching procedures, see HECKMAN; ICHIMURA; TODD (1998) and ABADIE; IMBENS (2006).

propensity score to be "treated" before matching, while Figure 16 (Appendix) after matching, both by state (UF)[101]. Table 34 presents estimation results.

Table 34: Matching of subset A – Pronatec impact on employment after 6 to 84 months of course conclusion

term	6 months	12 months	18 months	24 months	36 months	48 months	60 months	72 months	84 months
Treated	0.998	1.105	1.143	1.147	1.169	1.165	1.164	1.163	1.168
(O.R.)	(0.006)	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.009)***	(0.034)***
Job index	0.009	0.014	0.025	0.008	0.014	0.015	0.029	0.041	0.008
	(0.007)	(0.007)**	(0.007)***	(0.007)	(0.007)**	(0.007)**	(0.007)***	(0.011)***	(0.041)
Pronatec.share	-0.277	-0.319	-0.290	-0.238	-0.131	-0.072	-0.052	0.054	0.233
	(0.041)***	(0.040)***	(0.040)***	(0.040)***	(0.040)***	(0.040)*	(0.043)	(0.071)	(0.536)
Pronatec.wages	0.000	-0.003	-0.004	-0.002	-0.003	-0.002	-0.003	0.003	-0.005
	(0.001)	(0.001)***	(0.001)***	(0.001)**	(0.001)***	(0.001)***	(0.001)***	(0.001)*	(0.005)
N. obs.	478,033	478,033	478,033	478,033	478,033	474,926	434,472	207,071	15,115

Obs.1: O.R.= Odds ratio. Standard Deviation in parenthesis.

Obs.2: * = significant at 10%; ** = significant at 5%; *** = significant at 1%

Obs.3: Standard errors clustered by State.

Estimates are very similar to those presented before (Table 30), but with a slightly lower magnitude. First period now is not significant, being positive only a year after conclusion (10.5%). In the next two periods, effects increase to 14.3% and 14.7%, reaching 16.5% in the fourth year, remaining similar to this level thereafter. Job index has a positive effect during all period (but significant only in some years), while Pronatec share is now negative until the fourth year (only at 10% in this year). Pronatec wages has no influence just after conclusion but has a negative relationship from the first until the fifth year. Table 35 presents results for wages.

[101]Federation Unity (*Unidade da Federação*).

Table 35: Matching of subset A – Pronatec impact on wages after 6 to 84 months of course conclusion

term	6 months	12 months	18 months	24 months	36 months	48 months	60 months	72 months	84 months
Treated	-0.004	0.051	0.073	0.076	0.087	0.085	0.083	0.085	0.088
	(0.003)	(0.003)***	(0.003)***	(0.003)***	(0.003)***	(0.003)***	(0.004)***	(0.005)***	(0.020)***
Job index	-0.002	-0.001	0.007	-0.003	0.001	0.003	0.010	0.017	0.009
	(0.004)	(0.004)	(0.004)**	(0.004)	(0.004)	(0.004)	(0.004)**	(0.006)***	(0.024)
Pronatec.share	-0.125	-0.151	-0.138	-0.120	-0.068	-0.027	-0.019	0.034	0.087
	(0.021)***	(0.021)***	(0.021)***	(0.021)***	(0.022)***	(0.022)	(0.023)	(0.039)	(0.286)
Pronatec.wages	0.002	0.000	0.000	0.000	-0.001	0.000	-0.001	0.002	-0.002
	(0.000)***	(0.000)	(0.000)	(0.000)	(0.001)*	(0.001)	(0.001)	(0.001)**	(0.003)
N. obs.	478,033	478,033	478,033	478,033	478,033	474,926	434,472	207,071	15,115

Obs.1: Standard Deviation in parenthesis.

Obs.2: * = significant at 10%; ** = significant at 5%; *** = significant at 1%

Obs.3: Standard errors clustered by State.

Remarkably similar results are observed for wages too, with no positive effect just after course conclusion. It is positive after a year, reaching the maximum value seven years latter (9.2%). Once again, out of the market control variables, only Pronatec share seems to be important for wages, affecting negatively in the first three years.

Overall, the picture remains. Results are similar with or without matching, suggesting consistency among estimates, regardless the technique or the subset used.

3.4.2.2 Matching of subset B For subset B, since we have fewer control units than treated, we reverse the matching: for each control, we match one treated (in Appendix, Figure 17 presents densities distribution of propensity score to be "control" before matching, while Figure 18 after matching). The rest of the matching procedure is the same as mentioned for subset A. Table 36 present results for employment and Table 37 for wages.

Table 36: Matching of subset B – Pronatec impact on employment after 6 to 84 months of course conclusion

term	6 months	12 months	18 months	24 months	36 months	48 months	60 months	72 months	84 months
Treated	1.106	1.165	1.206	1.208	1.182	1.189	1.194	1.179	1.162
(O.R.)	(0.014)***	(0.014)***	(0.014)***	(0.014)***	(0.014)***	(0.015)***	(0.015)***	(0.026)***	(0.096)
Job index	0.005	0.013	0.036	0.004	-0.009	-0.001	0.005	0.044	0.122
	(0.017)	(0.017)	(0.017)**	(0.017)	(0.017)	(0.018)	(0.019)	(0.034)	(0.122)
Pronatec.share	-0.220	-0.191	-0.263	-0.183	0.060	-0.089	0.067	0.052	3.335
	(0.130)*	(0.127)	(0.128)**	(0.128)	(0.127)	(0.128)	(0.135)	(0.236)	(2.049)
Pronatec.wages	-0.002	-0.003	-0.005	0.000	-0.003	-0.005	-0.002	0.001	0.016
	(0.002)	(0.002)	(0.002)**	(0.002)	(0.002)	(0.002)**	(0.002)	(0.005)	(0.021)
N. obs.	79,510	79,510	79,510	79,510	79,510	78,157	72,550	25,870	2,523

Obs.1: O.R.= Odds ratio. Standard Deviation in parenthesis.

Obs.2: * = significant at 10%; ** = significant at 5%; *** = significant at 1%

Obs.3: Standard errors clustered by State.

Table 37: Matching of subset B – Pronatec impact on wages after 6 to 84 months of course conclusion

term	6 months	12 months	18 months	24 months	36 months	48 months	60 months	72 months	84 months
Treated	0.050	0.082	0.106	0.104	0.089	0.092	0.094	0.099	0.060
	(0.008)***	(0.008)***	(0.008)***	(0.008)***	(0.008)***	(0.008)***	(0.009)***	(0.015)***	(0.053)
Job index	-0.005	-0.001	0.016	-0.005	-0.009	-0.003	0.002	0.026	0.053
	(0.010)	(0.010)	(0.010)*	(0.010)	(0.010)	(0.010)	(0.011)	(0.019)	(0.067)
Pronatec.share	-0.054	-0.049	-0.108	-0.075	0.047	-0.047	0.013	-0.007	1.983
	(0.069)	(0.069)	(0.069)	(0.070)	(0.070)	(0.070)	(0.073)	(0.129)	(1.084)*
Pronatec.wages	0.001	0.000	-0.001	0.001	0.000	-0.002	-0.001	0.002	0.008
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.012)
N. obs.	79,510	79,510	79,510	79,510	79,510	78,157	72,550	25,870	2,523

Obs.1: Standard Deviation in parenthesis.

Obs.2: * = significant at 10%; ** = significant at 5%; *** = significant at 1%

Obs.3: Standard errors clustered by State.

In sum, we see that conclusions remain the same: no lock-in effects, treated has, in average, 17.9% higher chance to be employed, with 9.4% higher wages (excluding the last period, which is no longer significant, likely due to sample size). Job index seems not to be important both for employment nor wages, Pronatec share still negative but less significant, and Pronatec wages affects negatively only employment (in two of the periods). Thus, estimates are also consistent with this subset B, but we lost precision for market controls (possibly due to sample size).

3.4.2.3 IV regression In this subsection, we run an instrumental variable regression, to deal with some possibility of endogeneity not addressed by previous approaches. Equations 7 and 8 show the strategy for the first stages, while Equation 9 for the second stage.

$$Pronatec = \alpha_1 + \gamma_1 Z_1 + \gamma_2 Z_2 + \kappa_i \boldsymbol{X_i} + \mu \tag{7}$$

$$Pronatec.share = \alpha_2 + \psi_1 Z_1 + \psi_2 Z_2 + \phi_i \boldsymbol{X_i} + \nu \tag{8}$$

$$Y_{ma} = \alpha_3 + \delta_1 \widehat{Pronatec} + \delta_2 Job.index + \delta_3 \widehat{Pronatec.share} + \delta_4 Pronatec.wage + \beta_i \boldsymbol{X_i} + \varepsilon \tag{9}$$

In the above equations, Z_1 is a dummy for involuntary non-participants and Z_2 is the share of Z_1 in the market. Weak instrument (F-test) and Wu-Hausman test for endogeneity are reported in Tables 64 and 64 (in Appendix), suggesting that both instruments should be valid in all periods (but the last one). However, for periods longer than two years, Wu-Hausman test suggests that OLS should be more suitable. Nonetheless, Table 38 and 39 present results for employment and wages, respectively.

Table 38: IV regression of subset A – Pronatec impact on employment after 6 to 84 months of course conclusion

term	6 months	12 months	18 months	24 months	36 months	48 months	60 months	72 months	84 months
Treated (O.R)	1.058	1.051	1.059	1.058	1.042	1.046	1.044	1.045	1.114
	(0.007)***	(0.007)***	(0.007)***	(0.007)***	(0.007)***	(0.007)***	(0.008)***	(0.012)***	(0.071)
Job index	0.000	0.002	0.003	0.000	0.002	0.002	0.006	0.005	0.010
	(0.001)	(0.001)	(0.001)**	(0.001)	(0.001)	(0.001)	(0.002)***	(0.003)	(0.016)
Pronatec.share	-0.234	-0.226	-0.315	-0.165	0.054	0.032	0.099	0.271	-3.633
	(0.154)	(0.154)	(0.155)**	(0.154)	(0.154)	(0.150)	(0.152)	(0.261)	(12.230)
Pronatec.wages	0.000	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	0.001	-0.003
	(0.000)*	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)	(0.006)
N. obs.	640,270	640,270	640,270	640,270	640,270	622,225	567,059	272,755	26,294

Obs.1: O.R.= Odds ratio. Standard Deviation in parenthesis.

Obs.2: * = significant at 10%; ** = significant at 5%; *** = significant at 1%

Obs.3: Standard errors clustered by State.

For employment, results are analogous in directions, but lower in magnitude. Treated has, since the first period, 5.8% higher chances to be employed, a figure that lowers to 4.2%

three years after conclusion, remaining near this level until the sixth year. The negative effect of Pronatec share disappears, remaining only the negative effect of Pronatec wages.

Table 39: IV regression of subset A – Pronatec impact on wages after 6 to 84 months of course conclusion

term	6 months	12 months	18 months	24 months	36 months	48 months	60 months	72 months	84 months
Treated	0.125	0.118	0.140	0.127	0.088	0.094	0.092	0.109	0.294
	(0.017)***	(0.017)***	(0.017)***	(0.017)***	(0.018)***	(0.018)***	(0.018)***	(0.029)***	(0.169)*
Job index	-0.005	-0.002	0.002	-0.005	-0.002	0.002	0.010	0.003	0.005
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)***	(0.007)	(0.036)
Pronatec.share	-0.402	-0.224	-0.392	-0.080	0.342	0.309	0.535	0.959	4.106
	(0.352)	(0.354)	(0.357)	(0.358)	(0.360)	(0.352)	(0.354)	(0.613)	(28.571)
Pronatec.wages	0.001	0.000	-0.001	0.000	-0.001	0.000	0.000	0.002	-0.001
	(0.001)*	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)**	(0.014)
N. obs.	640,270	640,270	640,270	640,270	640,270	622,225	567,059	272,755	26,294

Obs.1: Standard Deviation in parenthesis.

Obs.2: * = significant at 10%; ** = significant at 5%; *** = significant at 1%

Obs.3: Standard errors clustered by State.

For wages, results are similar in directions too, but higher in magnitude in the first two years: treated has, in average, 13.6% higher wages with IV regression (against 5.8% in OLS regression). From the third year onward (except the last one), the average return for IV regression is 10% (against 9.5% in OLS). None of the market variables are relevant for wages.

Overall, the picture is quite the same, regardless of the technique used: standard logit/OLS, matching data combined with logit/OLS or instrumental variables. As Wu-Hausman test suggests, OLS results should provide us more efficiency in estimation, which is relevant for our market control variables.

3.4.3 Heterogeneous effects

As mentioned in review section, the literature reports heterogeneous effects for job training programs. Here, we explore some of them: gender; age, looking if the effect is different for youth; course type, looking if different competencies result in distinct job placement; and course requester, looking if there are difference between approaches of demandants. Considering that we have twelve axes, we concentrate analysis in the most important ones in terms of applications (Business and Management, Industrial Control and Processes, Infrastructure, Information and Communication, Tourism and Environment). We use same strategy for demandant, focusing on Ministry of Labor – the most important in

our subsets, Ministry of Social Development (MDS) – which has poor people as target – and On-line subscriptions, overrepresented on treated. Results, for gender and youth, are in Tables 40. Table 41 and 42 show axes results and Table 43 has results by demandant (always applied to subset B and regarding wages).

Table 40: Heterogeneous effect: women, men and young (18 to 29 years-old) – Pronatec impact on wages after 6 to 84 months of course conclusion, subset B

Axis	term	6 months	12 months	18 months	24 months	36 months	48 months	60 months	72 months	84 months
Women	Treated	0.026	0.062	0.078	0.085	0.072	0.067	0.067	0.059	0.008
		(0.009)***	(0.009)***	(0.010)***	(0.010)***	(0.010)***	(0.010)***	(0.010)***	(0.017)***	(0.076)
	Job index	-0.014	-0.015	-0.005	-0.009	-0.001	-0.003	0.008	0.021	0.020
		(0.007)**	(0.007)**	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.011)*	(0.051)
	P. share	-0.187	-0.234	-0.193	-0.178	-0.091	-0.090	0.032	0.185	0.912
		(0.038)***	(0.039)***	(0.040)***	(0.040)***	(0.041)**	(0.041)**	(0.044)	(0.076)**	(0.509)*
	P. wages	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.003	0.003	-0.001
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)***	(0.002)*	(0.007)
	N. obs.	133,664	133,664	133,664	133,664	133,664	132,940	122,490	54,083	3,346
Men	Treated	0.057	0.093	0.116	0.105	0.104	0.107	0.107	0.129	0.128
		(0.009)***	(0.009)***	(0.009)***	(0.009)***	(0.009)***	(0.009)***	(0.010)***	(0.015)***	(0.053)**
	Job index	0.007	0.011	0.017	0.006	0.007	0.003	0.006	0.034	0.020
		(0.007)	(0.007)	(0.007)**	(0.007)	(0.007)	(0.007)	(0.008)	(0.011)***	(0.043)
	P. share	-0.132	-0.123	-0.151	-0.104	0.003	0.014	-0.051	-0.050	0.216
		(0.037)***	(0.038)***	(0.038)***	(0.038)***	(0.038)	(0.039)	(0.041)	(0.066)	(0.526)
	P. wages	0.002	-0.001	-0.002	0.001	-0.001	0.000	0.000	0.001	-0.005
		(0.001)**	(0.001)	(0.001)*	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.006)
	N. obs.	146,162	146,162	146,162	146,162	146,162	144,724	132,392	63,225	5,501
Young	Treated	0.035	0.081	0.118	0.099	0.106	0.103	0.108	0.111	0.134
		(0.010)***	(0.010)***	(0.010)***	(0.010)***	(0.010)***	(0.010)***	(0.010)***	(0.016)***	(0.063)**
	Job index	0.004	0.002	0.005	-0.001	-0.003	-0.002	0.004	0.031	-0.029
		(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.011)***	(0.046)
	P. share	-0.096	-0.155	-0.148	-0.108	-0.008	-0.013	-0.018	0.027	-0.092
		(0.034)***	(0.035)***	(0.036)***	(0.036)***	(0.037)	(0.037)	(0.039)	(0.065)	(0.493)
	P. wages	0.001	-0.001	0.000	0.001	0.000	0.001	-0.001	0.002	0.001
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.006)
	N. obs.	133,572	133,572	133,572	133,572	133,572	132,575	120,931	55,996	4,153

Obs.1: Standard errors clustered by State.

Obs.2: * = significant at 10%; ** = significant at 5%; *** = significant at 1%.

Obs.3: Standard errors clustered by State.

We see that treated women have a positive and significant wage premium, in average, 6.7% higher than control, while, for men, effect is 10.8%, 4.1 p.p higher for them. Also, it seems that Pronatec share has more harm for women than men. For young (between 18 and 29 year-old), the average return is 10.5%, which is slightly above than for general population (9.5%). Also, Pronatec share is less significant for them (lower in magnitudes).

Regarding course type (axis), the average returns (excluding the last period due to lack of precision) for the three main axes, Business and Management, Industrial Control and Processes and Infrastructure, are 8.5%, 11.6% and 11% respectively, suggesting a slightly higher return for industrial related courses (Table 41).

Table 41: Heterogeneous effect by axis: Business and Management, Industrial Control and Processes, Infrastructure – Pronatec impact on employment after 6 to 84 months of course conclusion, subset B

Axis	term	6 months	12 months	18 months	24 months	36 months	48 months	60 months	72 months	84 months
B&M	Treated	0.039	0.066	0.085	0.107	0.078	0.091	0.090	0.094	0.161
		(0.011)***	(0.011)***	(0.011)***	(0.012)***	(0.012)***	(0.012)***	(0.012)***	(0.021)***	(0.089)*
	Job index	-0.010	-0.017	0.004	0.002	0.003	0.002	0.016	0.024	0.064
		(0.009)	(0.009)**	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)*	(0.015)	(0.064)
	P. share	-0.246	-0.282	-0.258	-0.190	-0.068	-0.135	-0.080	-0.073	0.558
		(0.066)***	(0.067)***	(0.068)***	(0.069)***	(0.070)	(0.070)*	(0.072)	(0.139)	(0.975)
	P. wages	0.000	-0.002	0.000	0.000	-0.002	-0.002	-0.003	0.001	0.004
		(0.001)	(0.001)*	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)*	(0.003)	(0.011)
	N. obs.	92,841	92,841	92,841	92,841	92,841	92,575	86,837	38,253	2,535
IC&P	Treated	0.056	0.089	0.133	0.091	0.120	0.116	0.127	0.147	0.074
		(0.016)***	(0.016)***	(0.016)***	(0.016)***	(0.016)***	(0.017)***	(0.017)***	(0.024)***	(0.093)
	Job index	0.010	0.013	0.021	-0.006	-0.008	0.001	-0.010	-0.003	0.027
		(0.012)	(0.012)	(0.013)*	(0.013)	(0.013)	(0.013)	(0.014)	(0.019)	(0.078)
	P. share	-0.199	-0.221	-0.287	-0.191	-0.117	-0.119	-0.082	-0.059	0.131
		(0.059)***	(0.061)***	(0.061)***	(0.062)***	(0.062)*	(0.063)*	(0.067)	(0.104)	(0.800)
	P. wages	0.003	0.000	-0.002	0.001	-0.001	-0.002	0.002	0.005	0.003
		(0.001)**	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)**	(0.009)
	N. obs.	50,656	50,656	50,656	50,656	50,656	49,953	45,889	24,710	2,354
Infra.	Treated	0.086	0.094	0.114	0.122	0.129	0.117	0.092	0.083	0.065
		(0.020)***	(0.020)***	(0.020)***	(0.020)***	(0.020)***	(0.020)***	(0.021)***	(0.030)***	(0.092)
	Job index	-0.005	0.006	0.012	0.009	0.024	-0.003	-0.009	0.040	-0.063
		(0.014)	(0.014)	(0.014)	(0.014)	(0.014)*	(0.014)	(0.015)	(0.021)*	(0.077)
	P. share	-0.207	-0.277	-0.261	-0.213	-0.078	-0.203	-0.138	-0.055	-0.018
		(0.119)*	(0.116)**	(0.115)**	(0.115)*	(0.117)	(0.116)*	(0.124)	(0.179)	(0.990)
	P. wages	0.011	0.004	0.003	0.002	0.000	0.004	0.001	0.000	-0.023
		(0.003)***	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.020)
	N. obs.	36,641	36,641	36,641	36,641	36,641	36,292	32,669	17,613	1,904

Obs.1: * = significant at 10%; ** = significant at 5%; *** = significant at 1%.

Obs.2: Standard errors clustered by State.

Obs.3: B&M = Business and Management; IC&P = Industrial Control and Processes; Infra. = Infrastructure

For the next three axes with more applicants, Information and Communication, Tourism and Environment, the average returns (also excluding the last period, and additionally the first) are 11%, 9.4% and 6.5% respectively, which suggest a much lower effect to the last axis

(Table 42).

Table 42: Heterogeneous effect by axis: Information and Communication, Tourism and Environment – Pronatec impact on wages after 6 to 84 months of course conclusion, subset B

Axis	term	6 months	12 months	18 months	24 months	36 months	48 months	60 months	72 months	84 months
I&C	Treated	0.031	0.087	0.122	0.115	0.085	0.078	0.111	0.130	0.346
		(0.023)	(0.023)***	(0.023)***	(0.023)***	(0.024)***	(0.024)***	(0.025)***	(0.038)***	(0.150)**
	Job index	-0.002	0.021	-0.008	-0.017	-0.008	-0.021	0.007	0.016	-0.050
		(0.019)	(0.019)	(0.019)	(0.019)	(0.020)	(0.020)	(0.021)	(0.030)	(0.160)
	P. share	0.009	0.065	-0.034	0.017	0.231	0.178	0.205	0.251	3.005
		(0.087)	(0.090)	(0.090)	(0.092)	(0.093)**	(0.094)*	(0.102)**	(0.148)*	(0.796)***
	P. wages	0.019	0.012	0.020	0.010	-0.001	0.002	0.005	0.007	-0.014
		(0.006)***	(0.007)*	(0.006)***	(0.007)	(0.007)	(0.007)	(0.007)	(0.009)	(0.051)
	N. obs.	23,000	23,000	23,000	23,000	23,000	22,809	20,713	11,460	727
Tour.	Treated	0.058	0.121	0.081	0.090	0.086	0.098	0.069	0.081	-0.021
		(0.021)***	(0.021)***	(0.022)***	(0.022)***	(0.022)***	(0.022)***	(0.023)***	(0.042)*	(0.270)
	Job index	-0.003	-0.020	0.010	0.023	-0.003	-0.012	-0.018	0.003	-0.356
		(0.020)	(0.020)	(0.021)	(0.021)	(0.021)	(0.021)	(0.022)	(0.035)	(0.328)
	P. share	-0.226	-0.189	-0.160	-0.123	-0.045	0.032	0.063	0.168	1.250
		(0.087)***	(0.088)**	(0.089)*	(0.091)	(0.094)	(0.093)	(0.095)	(0.158)	(1.203)
	P. wages	0.008	0.005	0.011	0.013	0.003	-0.006	-0.004	-0.012	0.098
		(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.014)	(0.075)
	N. obs.	19,437	19,437	19,437	19,437	19,437	19,289	17,790	6,632	391
Env.	Treated	0.034	0.077	0.067	0.053	0.077	0.047	0.059	0.062	-1.099
		(0.026)	(0.026)***	(0.026)**	(0.026)**	(0.027)***	(0.027)*	(0.028)**	(0.047)	(0.443)**
	Job index	0.014	0.009	0.002	0.034	0.016	0.027	0.020	0.085	1.193
		(0.018)	(0.018)	(0.018)	(0.018)*	(0.019)	(0.019)	(0.020)	(0.032)***	(1.461)
	P. share	-0.208	-0.340	-0.330	-0.233	-0.109	-0.060	-0.050	0.482	-39.694
		(0.127)	(0.135)**	(0.136)**	(0.137)*	(0.141)	(0.135)	(0.142)	(0.291)*	(22.643)*
	P. wages	0.003	0.001	0.001	0.007	0.006	-0.001	-0.003	0.002	-0.280
		(0.003)	(0.003)	(0.003)	(0.003)**	(0.004)	(0.003)	(0.003)	(0.005)	(0.141)**
	N. obs.	18,225	18,225	18,225	18,225	18,225	18,129	16,509	6,071	132

Obs.1: * = significant at 10%; ** = significant at 5%; *** = significant at 1%.

Obs.2: Standard errors clustered by State.

Obs.3: I&C = Information and Communication; Tour. = Tourism; Env. = Environment

Finally, looking to the training demandant (Table 43), we see that the average return (excluding last period) for Ministry of Labor is 8.5%, while for MDS is 10.8%, hence a higher return for the latter public, which focus on poor population. When we look to on-line applicants, the self-demand, we see no significant effect (last period estimate is problematic due to sample size). So, it seems that actions taken by the government are more important than self-demand by workers.

Table 43: Heterogeneous effect by demandant: Ministry of Labor, Ministry of Social Development and On-line applicants – Pronatec impact on employment after 6 to 84 months of course conclusion, subset B

Axis	term	6 months	12 months	18 months	24 months	36 months	48 months	60 months	72 months	84 months
ML	Treated	0.055	0.081	0.093	0.088	0.084	0.081	0.077	0.092	0.139
		(0.008)***	(0.008)***	(0.008)***	(0.008)***	(0.009)***	(0.009)***	(0.009)***	(0.018)***	(0.068)**
	Job index	-0.029	-0.029	-0.014	-0.016	-0.007	-0.016	-0.008	-0.010	0.128
		(0.009)***	(0.009)***	(0.009)	(0.009)*	(0.009)	(0.009)*	(0.009)	(0.017)	(0.070)*
	P. share	-0.113	-0.016	-0.124	-0.113	0.037	-0.073	-0.119	0.175	0.088
		(0.085)	(0.085)	(0.084)	(0.086)	(0.087)	(0.086)	(0.089)	(0.217)	(3.132)
	P. wages	0.000	-0.002	-0.002	0.000	-0.001	0.000	0.000	0.004	-0.007
		(0.001)	(0.001)*	(0.001)**	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)*	(0.007)
	N. obs.	123,632	123,632	123,632	123,632	123,632	122,416	113,944	45,066	3,711
MDS	Treated	0.022	0.086	0.125	0.125	0.114	0.117	0.123	0.108	0.021
		(0.012)*	(0.012)***	(0.012)***	(0.012)***	(0.012)***	(0.013)***	(0.013)***	(0.016)***	(0.067)
	Job index	0.027	0.022	0.027	0.020	0.010	0.015	0.019	0.048	-0.040
		(0.008)***	(0.008)***	(0.008)***	(0.008)**	(0.008)	(0.008)*	(0.008)**	(0.012)***	(0.047)
	P. share	-0.144	-0.121	-0.161	-0.112	-0.070	-0.075	-0.032	0.109	1.249
		(0.036)***	(0.037)***	(0.037)***	(0.038)***	(0.038)*	(0.038)**	(0.040)	(0.065)*	(0.434)***
	P. wages	0.002	0.000	-0.002	0.000	0.000	-0.001	0.000	0.002	0.000
		(0.001)	(0.001)	(0.001)*	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.007)
	N. obs.	92,756	92,756	92,756	92,756	92,756	92,308	84,451	43,858	3,936
On-line	Treated	-0.030	0.198	0.211	0.038	0.046	-0.070	-0.014	0.191	-1.384
		(0.155)	(0.152)	(0.156)	(0.166)	(0.172)	(0.183)	(0.183)	(0.260)	(0.239)***
	Job index	0.014	0.013	0.011	-0.004	0.006	0.003	0.011	0.028	-0.185
		(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.012)	(0.012)	(0.018)	(0.111)*
	P. share	-0.124	-0.282	-0.195	-0.170	-0.039	0.079	0.066	-0.007	-2.757
		(0.055)**	(0.057)***	(0.057)***	(0.058)***	(0.059)	(0.060)	(0.065)	(0.105)	(0.943)***
	P. wages	-0.001	0.000	0.002	0.002	0.001	0.001	-0.001	0.005	0.023
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)**	(0.019)
	N. obs.	50,411	50,411	50,411	50,411	50,411	50,123	45,882	22,633	677

Obs.1: * = significant at 10%; ** = significant at 5%; *** = significant at 1%.

Obs.2: Standard errors clustered by State.

Obs.3: ML = Ministry of Labor; MDS = Ministry of Social Development.

In sum, we notice differences in returns according to gender, age, course type and demandant, reinforcing the view that training programs have heterogeneous effects that must be taken into consideration (which is in accordance to CARD; KLUVE; WEBER (2010) and CARD; KLUVE; WEBER (2018) findings).

3.4.4 Benefit versus costs

When public policies are analyzed, not only impacts matter. The costs necessary to create that impact must be assessed to inform society if it was a good investment. Since all

the benefits are not known, it is not a trivial task (despite its importance). Looking to our matching of subset B, considering the value paid by the government by hourly course, R$ 76.64 million was invested only to treated who concluded training, which gives R$ 1926.63 by person. Considering the six years spell we are analyzing here[102], it would require R$ 26.76 higher wage every month to cover all the direct cost.

Taking the effects of matching subset B, Table 44 shows estimated returns in hourly wages for each period[103]. Assuming the return estimated in each period would be valid for the past period (for example, estimates for the 6th month is valid for the past five months and estimates for the 36th month is valid for the past eleven months) and for all participants, total return is R$ 199.04 million, which gives R$ 5003.51 by person, or R$ 69.49 by month and person. So, program is highly cost effective in this first analysis. When present value is considered[104], the net real return is R$ 66.12 million (or R$ 1.86 for each R$ 1.00). If we extend our time of analysis, assuming that the wage difference observed after the sixth year will last until retirement (thirty one more years, given the average age of treated), the program is highly economic (a net benefit of R$ 249.5 million, or R$ 4.26 for each R$ 1.00).

Table 44: Returns to Pronatec considering the matching of subset B, 6 to 72 months after course conclusion

N# treated	Time	Effect (w/h)	Effect (w/m)	Total effect	Total effect PV
39,781	6 months	0.04	7.72	1,843,601	1,788,590
	12 months	0.25	48.25	11,516,364	10,605,332
	18 months	0.47	89.60	21,385,476	18,693,588
	24 months	0.46	88.09	21,025,070	17,445,205
	36 months	0.38	71.93	34,335,144	26,355,562
	48 months	0.37	70.90	33,846,167	23,408,452
	60 months	0.41	78.29	37,375,639	23,290,649
	72 months	0.41	79.01	37,717,248	21,176,917

Obs.1: w/h = wage per hour.

Obs.2: w/m = wage per month, considering 4.35 weeks monthly and 44 hours per week.

Obs.3: PV = Present Value, considering an interest rate of 10.99% per year.

If heterogeneous effects were taken into consideration, higher or lower returns would emerge. This result is particularly important to reassess the program design, since there are good opportunities for higher returns to society, adjusting it when necessary.

[102]The last period was dropped because it was less precisely estimated and was not significant.

[103]We rerun regressions in level, not in log.

[104]The average of basic interest rate (Brazilian Selic) considering the first and the last date of training conclusion was 10.99% a year.

It is also important to say that this analysis is not considering other possible effects of the program, like improvements of earning in informal labor market, self-employment, or entrepreneurship. Also, Pronatec has an assistentialist purpose, so we are not able to assess benefits regarding the opportunity of education that program offers to population. On the other hand, we neither account for costs involving dropouts and administration of the program (systems maintained to organize the program, offices and staff employed to analyze the implementation, advertising etc.). As said before, although its necessity, it is not a simple task to assess the economic return of the program.

3.5 Conclusions and remarks

Active Labor Market Programs (ALMP) are exceptionally popular in the world, with relevant initiatives since the 1960's. The meta-analysis of these programs, specifically training programs, suggest an initial lock-in effect in the short run with positive and lasting effects in the mid and long runs, with a variety of magnitudes.

In Brazil, a large-scale program in this fashion is Pronatec, launched in the end of 2011, with its peak in 2014. Today, the program is still running, but in a quite lower scale, much as consequence of economy recession and lower availability in government budget. Despite the amount invested in the program, there are few studies available about its effectiveness (as pointed before, BASTO et al. (2016) and QUINTANA; CRAVO (2019) only looked to MDIC, while SENADO (2017) focuses on the program implementation aspects). In this study we looked to the entire *Bolsa Formação* arm, consisting almost 50% of the program participants and 30% of total budget (R$ 11 billion in 2012-2016 period), covering almost 4 million people.

Results are in accordance with literature: when considering a short unemployment spell (within six months), there seems to be a sort of lock-in effect right after training conclusion (until six after), with positive and increasing results thereafter, usually with a peak in the long run. However, when we use a more similar subset (with involuntary non-participants), the lock-in effects disappear, and results are even higher. Results are robust to several subsets considered here, including matching and instrumental variables. Heterogeneous effects indicate that, even though women are most participants (in all Pronatec), impact is lower for them, but still positive. Young seems to benefit more than general population too. Also, we see that course type matters, with stronger results for those related with industry. This is an important result because previous studies, usually, consider all different kind of training programs as a unique treatment. Demandant of course is also relevant for Pronatec, with a stronger impact for MDS applicants, while self-demand (on-line application) seems not improve formal labor opportunities.

The benefit versus cost exercise suggests that, at least when the formal labor market is looked for outcomes of the *Bolsa Formação* worker arm and putting all limitations of this kind of analysis aside, the program is highly economic. Returns, in terms of hourly wages, cover all direct costs. If the length of effects is extended until retirement, the return is even higher. Additionally, heterogeneous effects reveal that parts of the program might be even more economic.

Also, we must keep in mind that, at least for Pronatec, there is an assistencialist objective, which is exactly the target group of MDS, the majority of participants. Eventual positive externalities are not being considered here, neither the simple fact that including people with low schooling level in education programs may change other opportunities for them, not analyzed here. Finally, we see that Pronatec shares are important for outcomes in the short run, indicating that an increase in trained labor offer in a specific market results in lower chances to find a job. In this scenario, SUTVA seems not hold, suggesting that changes caused by the program in the formal local labor market must be considered, echoing finds reported by FERRACCI; JOLIVET; BERG (2014) for France.

Once Pronatec is a relatively young program, long run effects did not happen yet for several participants. So, future research can extend the period of analysis, increasing the number of observations in the mid and in a longer run. Also, there are other arms of Pronatec that must be evaluated, specially the one related to technical education (see BISHOP; MANE (2004) and VASCONCELLOS; LIMA; MENEZES-FILHO (2010), for a discussion about this subject). Ultimately, we only considered here outcomes in the formal labor market, while there is room to verify effects in the informal labor market (which is particularly important in Brazil), self-employment or entrepreneurship of concluders.

4 Returns to technical education: a longitudinal and cross-sectional study of Brazil, 2007 to 2018[105]

[105] In co-authorship with Bruno de Oliveira Cruz (IPEA) and Luiz Rubens Câmara de Araújo (CODEPLAN).

4.1 Introduction

School-to-work is one of the main policy makers' challenges in many countries, including Brazil[106]. The National Household Survey (PNAD)[107] shows, for example, that unemployment rate was 10.95% in Brazil in 2014, for young people aged 18 to 24 (while the general rate was 6.25), so these juveniles were an important part of all unemployed individuals in Brazil. Moreover, estimates shows that 23.5% of this public were neither in school nor holding down a job (among them, unemployed people are included[108]). The absence of job or any kind of occupation can potentially be a source of various social problems, like poverty and incidence of crime. At the same time, there is a growing effort by local government to increase vocational training and technical education opportunity to this population.

On the other hand, 83% of above 25 years-old had, at most, high school degree (about eleven years of schooling). People with primary education (about eight years of schooling) were 10%, while people with primary incomplete were 32% and no educated were 12%[109]. Thus, there still a huge part of the working age population with low qualification and, therefore, potential low job placement. In this scenario, technical education is often an alternative to enhance labor market opportunities for youth, decreasing unemployment and providing a source of income (FRIGOTTO, 2005). Hence, identify the return to technical education on earnings is important to answer the following question: are there any gains to young people, regarding wages or professional careers, if they have a technical education certificate?

Using a panel with all Brazilian formal workers, registered in Annual Relation of Social Information (RAIS)[110], from 2007 to 2018, we estimate the effect of occupations related to technical education on wages in formal labor market. Our database covers all registered workers, using occupational information based on Brazilian Classification of Occupations (CBO), a codification organized by IBGE, to identify technical workers. Previous studies using panel data are scarce, especially about technical education. Regarding Brazilian literature, we find papers using cross-sectional databases, mainly focused on the 2007 National

[106]Ryan (2001) compares countries from Europe and United States in school-to-work transition and several public policies undertaken to improve youth labor Market insertion. In 2019, the Provisional Measure 905 aimed to improve job insertion of young aged 18 to 29, but is no longer active. Available in: http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2019/Mpv/mpv905.htm

[107]Survey carried out by *Insituto Brasileiro de Geografia e Estatística – IBGE*.

[108]People who is looking for a job

[109]This means that 54% of people would need, necessarily, increase their schooling level in order to be able to get a technical certificate (plus 4% with incomplete high school).

[110]RAIS is a mandatory administrative database. All formal companies in Brazil must inform Ministry of Economy all active employees and/or who was an employee, but left the company during the year.

Household Survey, when a special supplement about professional qualification was applied[111]. Nevertheless, longitudinal studies, in general, can control efficiently non-observable effects (constant in time), which may affect econometric studies, a feature absent cross section databases (STEVENS; KURLAENDER; GROSZ, 2019). However, due to structure of our data, with an invariant characteristic in time, we are not able to estimate the effects of technical education by fixed-effects. To overcome this issue, we rely on random-effects, with several alternative estimation by complex random-effect-within-between (REBW) of matching data. Hence, this paper contributes to improve the debate about the relationship between education and wages, specifically technical qualification. Recent Brazilian high school reform, carried out in 2017, praises the importance of this debate. We also use PNAD 2007 and 2014 waves about technical education to assess the consistency of our estimates with an alternative database, taking into consideration the selection to labor market.

Our panel estimates shows that workers who were in occupations related to technical education in 2007 had a positive and significant wage difference, between 21.3% and 24.9%, when compared to other workers, controlling for several characteristics (e.g. schooling, tenure, sex and age). When we focus on a more specific group, young with 18 years-old (and hence with no past wage trends), who had only high school in the beginning, we find a positive and significant wage difference, but in a lower magnitude – between 5.8% and 7.8%. Results are robust to matching data and random effects within-between estimates. The cross-section results point to same direction, although the effect seems to fade away, with no effect for the young generation of 2014, suggesting the importance in reassessing the effects over time and generations. Heterogeneous effects suggests that industry workers benefit more from technical education than the general workers (only in cross-sectional data). A cost-benefit analysis estimates the maximum value could be spent, given some interest rates, to cover the costs with technical education. Results suggest, given an ideal cost (estimated by ARAÚJO et al. (2016)), the benefits would cover between 13.6 times and 5.2 times of high school costs, or between 10.5 times and 4 times of the technical education costs (depending on rate of return considered).

This paper is organized as follows. Section two presents a brief review of technical education and previous papers' estimation of its returns. The next one brings methodological aspects, including data sources and its manipulation. Fourth section presents the main results, while the last concludes with remarks.

[111]Panel starts in 2007 in order to compare our results with literature

4.2 Importance of education, technical education and its impacts on labor markets

In a wider perspective, correlation between education and economic development exists for a long time (HANUSHEK et al., 2008), gathered under human capital theory, which has been an important research area backing to the middle of XX century (BECKER, 1962). Since then, many studies flourished trying to estimate returns to education investments (PSACHAROPOULOS, 1994), where income equation, based on wages, is a common tool (MINCER, 1974). Among schooling types, or phases, technical education is the one conceived to improve school-to-work transition.

An important characteristic of technical education in Brazil is its narrow connection with high school, being, at same time, integrated or consecutive, with a recognized certificate[112]. In the first kind of technical education, students are enrolled in two courses, high school and technical education, and take parallel classes. In the second type, students are enrolled only in one course, taking classes simultaneously. Finally, after concluding high school, they can enroll in an isolated technical course, which requires high school certification. One of the most important features of technical education is its association with basic education, giving a wider qualification to future workers, since it demands a range of previous abilities from students. In opposition, professional qualification is usually simpler and shorter, needing, in general, no specific background knowledge (ALVES; SANTOS VIEIRA, 2009). FRIGOTTO (2005) points to the necessity of educated citizens with critical thinking, able not just to execute technical tasks, but capable to exercise citizenship, only possible through high quality basic education. Also, general education may help people adapt or acquire new skills due to changes in industry, a common scenario with technological changes.

Association between basic education, i.e. conclusion of all three levels of education[113], and technical education was not always prevalent in Brazil. The Decree-Law 2,208/1997[114] separated technical education from regular high school, stopping its federal expansion, dissociating these two type of teaching. This rule changed in July 2004, with Decree-Law

[112]Also known as Training and Vocational Education and Training (TVET) or Career and Technical Education (CTE), these terms are interchangeable here, meaning the same type of education. Short term vocational training, or professional qualification, on the other hand, does not have, in general, a formal time schedule neither demands a basic educational background. In addition, government authorization is not required to offer this type of qualification. Thus, professional qualification and technical education are conceptually different, with distinct objectives in Brazil and will be distinguished when necessary.

[113]Today, there are three levels of basic education in Brazil. For children under 5 years, called *Educação Infantil*; for children aged 6 to 14, called *Ensino fundamental*; and for 15 to 17, called *Ensino Médio*. These three levels are mandatory to all children/adolescents. They are equivalent, respectively, to preschool and kindergarten, elementary and middle school and high school in countries like United States.

[114]Available in: http://www.planalto.gov.br/ccivil_03/decreto/d2208.htm

5,154/2004[115]. Today, Brazil lives a new phase in this area, when Provisional Measure 746/2016[116], implemented High School Reorganization, later converted in law[117], which makes high school more flexible, establishing also common topics to be covered in all Brazilian schools. All these changes aim to help school-to-work transition, increase high-school graduation and improve quality of education. It is important to point out that, in 2014, 98.7% of children aged 7 to 14 were in school in Brazil, lowering to 84.3% for adolescents aged 15 to 17 and only 48.5% of people aged 18[118]. So, while the *ensino fundamental* (7 to 14 years) could be considered universalized, the *ensino médio* (15 to 17) was still behind, with around 15% of young out of school.

In line with school-to-work discussion, CORSEUIL; FOGUEL; GONZAGA (2019) evaluate the Brazilian Apprenticeship program[119], which subsidies firms to hire and train young, finding positive effects in the chances to hold a formal job five years after the program, also finding a positive effect in acquiring higher levels of formal education. Endogeneity was dealt with the age rules to participate in the program and the year of its implementation. To answer the same question, FERSTERER; PISCHKE; WINTER-EBMER (2008), using failed firms, found positive effect of training on employment, with results being similar between OLS and instrumental variables estimation. Despite technical education is not necessarily like an apprenticeship program, with on the job learning regularly, its concept is connected by the school-to-work transition purpose, making these results relevant to our discussion.

An important aspect of formal education curriculum is the academic versus technical track in the end of basic education. Some countries tend to put some emphasis to academic track, aiming to provide more general skills and, also, to open a door to higher education. In this line, family decision plays a central role, where expected return to the academic track, when compared to the technical education, is relevant in decision making (see BIAVASCHI et al. (2012) and KAHYARARA; TEAL (2008)). Sometimes people simply do not know the real return to technical education and prefer to enroll their children to academic path, the "safer choice", making this an informational issue. Indeed, PSACHAROPOULOS (1994) reports that rates of returns (ROR) of general education are higher than technical education, especially due to the later be more expensive (in unitary costs) than the former. However, BENNELL (1996) points out that costs of general education might be different in developing countries, and, hence, the higher ROR of general education must be taken with caution due

[115] Available in: http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2004/decreto/d5154.htm
[116] Available in: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/mpv/mpv746.htm
[117] Law 13,415/2017, available in: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2017/lei/l13415.htm
[118] According to PNAD/IBGE.
[119] *Programa Jovem Aprendiz.*

to countries heterogeneity. MEER (2007) brings this discussion showing that, for United State, the path chosen by students of technical education would not be higher if they have chosen differently, suggesting that the alternative path is not necessarily better. On the other hand, KRAFFT (2018), using longitudinal men sibling's information for Egypt (15-34 years), shows that returns to formal vocational secondary education are not different for those without formal education, and, in fact, there are substantial returns to vocational skills acquired outside school. So, the academic versus technical education debate (and even outside school skills acquisition) seems to be still open. There are relevant country idiosyncrasies, which makes our study important to bring new evidence to discussion. Despite of this, the technical education framework in Brazil has a close relationship with formal education, which approximate paths of academic and technical education.

The path difference between academic and technical education can be related not only to outcomes in the labor market, but also to level of education. In this sense, DOUGHERTY (2018), taking advantage of oversubscription, shows that CTE programs increase the probability of high school graduation on-time, with higher effects for low-income students. This is an important aspect for countries where high school dropout is relatively high and the quality of education is low, which is the case of Brazil as mentioned before (see LEON; MENEZES-FILHO (2002) and NERI (2015) for a discussion about school dropout in Brazil). BISHOP; MANE (2004), using USA data, also suggest positive impacts when technical courses track are offered in high school on: attendance, graduation, even higher levels of education, and labor market outcomes. All of them with positive rates of return considering the costs. Still according to BISHOP; MANE (2004), students who spent about 1/6 of their time dedicated to occupational courses in high school had, at least, 12% higher wages after school, and about 8% seven year later, controlling for previous ability and family background. GRUBB (1996) shows similar evidences, suggesting that two years of qualification programs can improve students' economic status, while short-term programs of qualification have lower results, recommending, therefore, unification of professional training and educational programs. So, not only labor market outcomes may be improved, but also the education level, suggesting a double benefit of technical education.

A common discussion regarding job market and education is signaling (introduced by SPENCE, 1978). Due to job market imperfect information, more educated people (with higher levels of education or some training), who have a certificate, give a signal to the market that they have higher abilities, and, hence, are more productive, regardless of the quality of education received. In this discussion, CARRUTHERS; SANFORD (2018) show that people with a certificate from Tennessee Colleges of Applied Technology had higher quarterly earnings

than non-completers, who, in their turn, earn higher earnings than matched non-students. So, it seems that technical education indeed provides skills that are useful in labor activities. Also, there are important heterogeneous effects, with stronger returns for health associated courses and improvement in mobility between industries. Similarly, STEVENS; KURLAENDER; GROSZ (2019), using a fixed-effect and individual-specific trends, show positive effects from California community colleges career and technical education concluders, ranging from 14 to 45 percent, also with higher returns to healthcare sector programs.

Another relevant aspect is the heterogeneous effects of technical education, not only due to course type. SAKELLARIOU (2003) shows that, while formal education offer higher returns to males, the scenario is the opposite for females. Also, the social return reported to vocational secondary education exceeds the academic by 10%. Once technical education has a variety of courses, and some of them are traditionally gender related, this kind of heterogeneous effects analysis is important.

Regarding short training programs, while distinct of technical education, they also have relevance in this discussion due to their similar aims. ATTANASIO; KUGLER; MEGHIR (2011), using a randomized setup, show that this type of initiative has high rates of return in terms of earnings (19.6%) and employment in the short-run (over the year after course conclusion), for women. In the long run, the follow-up of the same study (ATTANASIO et al., 2017), and with a larger sample, shows that the results are persistent and positive for men. The analysis took into consideration the formal job market, finding about 12% higher earnings overall which gives an economical status to the program. In a different analysis, BRUNELLO; COMI; SONEDDA (2012) use regional variation in Italian training subsidies to show a positive effect in monthly earnings, with different results by firm size (higher for smaller firms). GREENBERG; MICHALOPOULOS; ROBINS (2003), HECKMAN; SMITH (2004), CARD; KLUVE; WEBER (2010), CARD; KLUVE; WEBER (2018) and VOOREN et al. (2019) survey several studies relating training policies to labor outcomes, like wages or finding a job. In sum, they conclude that exists a significant middle/long term effects on earnings for workers with some professional/technical qualification.

In Brazilian literature, the main focus is on a special supplement of 2007 Brazilian Household Survey (PNAD)[120], dedicated to this topic. Results, in general, show a positive and significant effect of technical/professional qualification on wages, between 12% and 14% (VASCONCELLOS; LIMA; MENEZES-FILHO, 2010). BARROS et al. (2011) compares

[120] *Pesquisa Nacional por Amostra de Domicílios* from Brazilian Institute of Statistics and Geography (IBGE).

people between 25 and 65 year-old in *Espírito Santo*[121] to other people with same age in Southeast Region, finding 11% higher wages for workers who had technical certificate. AGUAS (2014), also using 2007 PNAD, estimates returns to technical qualification by three approaches: OLS, treatment effect and propensity score. She finds a positive and significant wage premium, between 21% and 24%, for technical education certificate holders. Regarding longitudinal data, OLIVEIRA; RIOS-NETO (2007) analyze the impact of National Plan for Professional Qualificatio[122] held in Belo Holizonte between 1996 and 2000, using a CEDEPLAR/UFMG's[123] database, finding a reduction of unemployment spell. REIS (2015), using Monthly Employment Survey (PME)[124], from January, 2006 to December, 2012, finds a positive impact of technical education in wages per hour (8%).

Thus, literature about relationship between technical education and outcomes in labor market is not exhaustive in Brazil. We find only two papers using longitudinal databases to evaluate this effect, none of them using the RAIS database, a rich administrative register with detailed information of formal workers. We also update PNAD results with 2014 wave of the same topic, hoping to provide a wider evidence regarding technical education in Brazil. So, with a longitudinal and two cross-sectional databases, we hope to update estimates and give a wide perspective regarding technical education in Brazil.

4.3 Database description and methodological aspects

In this section we will present the two databases (RAIS and PNAD) and the estimation strategy for each one.

4.4 Annual Relation of Social Information (RAIS)

RAIS is an administrative database organized by Ministry of Economy, which contains personal information of all Brazilian workers in formal labor market. Every firm that had a relationship with a worker during the year are obligated to inform the Ministry this relationship.

We cover the 2007 to 2018 period, considering only workers who was still working in the company at the end of the reference year (last day)[125]. Also, we filter workers who

[121] One of 26 Brazilian States.
[122] PLANFOR – *Programa Nacional de Qualificação Técnica.*
[123] Federal University of Minas Gerais.
[124] *Pesquisa Mensal de Emprego*, a survey from IBGE.
[125] Information related to defense activity (CNAE 8422-1) was excluded, since all army servants are registered in Federal District (e.g. all army soldiers of each state are computed in DF), not reflecting their real job location.

had contracts with, at least, 10 hours of working time and a hourly wage above the federal minimum[126]. Also, we restrict data to people with at least 18 years, the minimum age for a full-time job, excluding those under apprenticeship contract.

To construct our panel, we created a unique key to identify workers every year. We joined first and last names with all personal identification numbers (PIS, CTPS and CPF), using 2007 as base year. This procedure was necessary due to changes in CPF or PIS in the subsequent years, due to error or real changes[127]. Thus, with this unique key, we could ensure that we find the exact same workers in the following years. When a worker had two or more jobs, we kept only the one with higher wage. Therefore, based on this unique key constructed for all workers in Brazil in 2007, we followed them during the next years until 2018. We use national database to consider possible migration of workers between municipalities.

We define technical occupation, the treatment, as follows: we match all CBOs[128] with National Catalog of Technical Courses (CNCT[129]), finding every worker in this condition in 2007 RAIS. Considering that high school in mandatory for those who have technical education, we identify workers who had, at least, this schooling degree or incomplete higher education. Thus, all workers with compatible schooling and in technical related occupation were considered as potentially workers with a technical certification.

Currently, Brazilian technical education has 227 courses with 800, 1,000 or 1,200 hours class load, divided in 13 groups/axes (Environment and Wealth; Industry processes and control; Education and Social Development; Business and management; Information and communication; Infrastructure; Military; Food production; Cultural production and design; Industrial production; Natural resources; Safety; and Tourism and leisure)[130]. National Catalog brings expected profile for all professionals formed by the courses, minimum required infrastructure, job field, related norms and laws, certification possibilities, further qualifications path and formation integration, and, associated CBOs.

We focus on workers in occupations related to these courses holding high school or incomplete higher school degree, since college concluders have a superior degree of education, with different competences. All other works were grouped as "outside technical occupation" (controls). We follow these two groups, starting in 2007, until 2018, and, using this panel,

[126]Companies are in charge to provide information, so incorrect report, although rare, is a possibility.

[127]Also, these are marginal occurrences.

[128]Brazilian Classification of Occupation.

[129]This national catalog is currently in third edition, available in: http://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=41271-cnct-3-edicao-pdf&category_slug=maio-2016-pdf&Itemid=30192.

[130]The codes for each axis in regressions is 1 to 13 in this same order.

we hope capture differences in the two groups' trajectory, in a more efficient way than using cross-section data. Once we are interest in a time-invariant characteristic (workers who started the period in technical education), it is not possible to estimate parameters with a fixed-effect model. So, we use a random effect model with some alternative specification to deal with possible endogeneity. However, even in a scenario where fixed-effect approach were possible, if the wage tendency of technical and non-technical job holders were different, the estimator would not be able to give the real effect of treatment. To deal with it, in an alternative estimation, we keep only people with 18 years-old in 2007, with no experience in the formal labor market. By doing this, we assure that all of them have the same initial wage tendency (starting from zero). As a robustness exercise, we also match treated and control to create a more homogeneous group.

So, equation 10 refers to the standard model to be estimated.

$$y_{it} = \alpha + \beta^T \boldsymbol{X_{it}} + \upsilon_{it} \tag{10}$$

where y_{it} is the log of hourly wages and $\boldsymbol{X_{it}}$ are regressors.

As said before, it is usual to treat $\upsilon_{it} = \mu_i + \epsilon_{it}$ in panel analysis (BALTAGI, 2008), a fixed-effect context, where the time-invariant parameter of unobservable characteristics (μ_i) vanishes with a demeaned data (within estimator) or with first-difference (FD) estimator. However, pursuing this path, all the other time-invariant characteristics, like gender and our variable of interest, vanishes as well. To overcome this, with the random-effect strategy, a *"quasi-demeaned"* model is defined in equation 11 (CROISSANT; MILLO, 2008).

$$y_{it} - \theta\bar{y} = (\boldsymbol{X_{it}} - \theta\bar{\boldsymbol{X}}_{\boldsymbol{i}})\beta + (\upsilon_{it} - \theta\bar{\upsilon_{it}}) \tag{11}$$

where $\theta = 1 - [\frac{\sigma_\upsilon^2}{\sigma_\upsilon^2 + T\sigma_e^2}]^{1/2}$, $\bar{y}$ and $\bar{\boldsymbol{X}}_{\boldsymbol{i}}$ are time means of y and $\boldsymbol{X}$. When $\theta = 1$ we have the fixed-effect estimator, while when $\theta = 0$ we simply have a pooled OLS.

In the robustness section, we estimate a complex random effect within-between model (REWB) following BELL; FAIRBROTHER; JONES (2019), that incorporate within and between effects, allowing us to estimate time-invariant parameters (equation 12).

$$y_{it} = \alpha + \beta_{nW}(\boldsymbol{X_{it}} - \bar{\boldsymbol{X}}_{\boldsymbol{i}}) + \beta_{nB}\bar{\boldsymbol{X}}_{\boldsymbol{i}} + \beta nz_i + \mu_i(\boldsymbol{X_{it}} - \bar{\boldsymbol{X}}_{\boldsymbol{i}}) + \epsilon_{it} \tag{12}$$

where $(\boldsymbol{X_{it}} - \bar{\boldsymbol{X}}_{\boldsymbol{i}})$ is the demeaned within predictor of time-variant variables, β_{nW} is

the coefficient within, β_{nB} is the coefficient between and βn and z_i are the coefficients and time-invariant variables, respectively.

We also introduce a market control to consider eventual spillover effects of technical education (FERRACCI; JOLIVET; BERG, 2014). Every year, municipality and axis is considered a market, with the share of technical workers over the total workers. Equation 13 shows the controls, where $Axis.share_{y,m,a}$ is the share in year y, municipality m and axis a, $T_{y,m,a}$ are the technical job holders and $W_{y,m,a}$ are all the workers.

$$Axis.share_{y,m,a} = \frac{\sum T_{y,m,a}}{\sum W_{y,m,a}} \tag{13}$$

Table 45 shows the variables we will use and their description[131]. All characteristics listed are controls usually found in wage equations.

[131]PBF is one of the biggest conditional cash transfer program in the world. The target are families under the extreme poverty and poverty lines (in 2020, families earning up to R$ 89 by person, or U$ 17, by month are considered extremely poor, while families above that amount and up to R$ 178, or U$ 35, are considered poor), focused one children. As counterpart, school attendance and vaccination are required. PBF reaches around 14 million families in Brazil in 2021. On the other hand, BPC is a program for elderly and handicapped. The poor population in this profile (people aged 65 or over and all handicapped) are eligible for a minimum wage paycheck (R$ 1.100, or U$ 216, in 2021).

Table 45: RAIS Variables: description by type

Type	Variable	Description
Panel IDs	Worker information	name, Social Security Number (PIS/NIS) and Personal Identification number (CPF)
	Year	The year of information, from 2007 to 2018
Outcome	Hourly wage	Average of hourly real wage in the year
Treatment	Technical CBO	Brazilian Occupational Classification (CBO) codes used to identify technical education related jobs
Controls	Age	On December 31st, and its quadratic to capture non-linearities
	Gender	Male or female
	Schooling	Two dummies: high school or higher education
	Tenure	Length, in year, working in the current job, and its quadratic to capture non-linearities
	Firm size	A dummy for big companies (above 500 employees)
	Occupation	Dummies for the ten groups of Brazilian Occupational Classification (CBO) codes
	Industry	A dummy for the Industry sector, based on National Classification of Economic Activities (CNAE)
	Commerce	A dummy for the Commerce activity, based on National Classification of Economic Activities (CNAE)
	Public Sector	A dummy for the Public sector, based on Legal Nature of the firm
	GDP PP	Gross Domestic Product per person in the municipality
	PBF PP	Annual transfers per person from Bolsa-família Program (PBF)
	BPC PP	Annual transfers per person from Benefício de prestação Continuada (BPC)
	Metropolitan Area	A dummy for municipalities inside a Metropolitan Areas
	Distance	Distance, in km, to the state capital
	FIT	A dummy for municipalities with Federal Insitute of Technology in the year (taken from Higher Education Census).
	Axis	Share of workers in a given market. The market is a year, a municipality and one of the thirteen technological axes.

Source: RAIS, INEP and IBGE

The entire database has 294.7 million observations, which brings a computational limitation to the analyzes[132]. To overcome this situation, we take about 1.4% random

[132]In RAIS, one person may have as many entries as jobs she has. Hence, the information is about *job*, not person. If a person has more than one job in a year, we keep the one with higher wage.

sample of 2007 workers and follow them in the subsequent years, resulting in 3.97 million observations[133]. Table 46 shows these figures.

Table 46: RAIS – total observations, sample and share of Technical CBO, 2007-2018

Year	Observations	Sample	% sample	% Tech CBO
2007	35,996,861	502,061	1.39	5.2
2008	28,511,965	272,393	0.96	5.4
2009	26,855,478	374,185	1.39	5.5
2010	26,208,774	357,283	1.36	5.6
2011	25,370,347	353,519	1.39	5.6
2012	24,538,620	341,963	1.39	5.6
2013	23,732,256	329,943	1.39	5.7
2014	22,936,243	318,885	1.39	5.7
2015	21,711,226	301,382	1.39	5.8
2016	20,444,715	283,311	1.39	5.8
2017	19,553,765	270,904	1.39	5.8
2018	18,848,543	261,709	1.39	5.8

Source: RAIS

We see that, except for the first two years (2007 to 2008), we lose, on average, less than one million observations over the years. It is likely that, due to the 2008 global crisis, a huge portion of workers lost their formal jobs and never returned. In our sample, we see a reversion of the crisis from 2008 to 2009; the share of workers with technical CBO starts with 5.2% of total, rising to 5.8% in the last period. So, overall, the share of technical occupations is low in Brazil.

4.4.1 Brazilian National Household Survey (PNAD)

The PNAD is a yearly investigation of labor force characteristics. It is undertaken by IBGE since 1967 and, in some years, the survey has special supplements to investigate other matters (TRAVASSOS; VIACAVA; LAGUARDIA, 2008). Since 2004, PNAD is representative to the whole country[134] and, regarding technical education, PNAD had two special questionnaires: in 2007 and 2014.

As a household survey, we have information about the workers' family, which gives us the opportunity to take into consideration a wider range of characteristics, including some of the determinants for participation in labor market. On the other hand, we do not have a

[133]Considering computer memory constraints, this figure was the maximum value that made this analysis feasible.

[134]Before 2004, in the North, only urban areas were in the sample. The other areas of the country were fully representative.

panel structure to handle the time component, which makes both analyses complementary. Table 47 shows the variables we use to estimate the returns, considering also the selection into labor market.

Table 47: PNAD Variables: description by type

Type	Variable	Description
Outcome	Occupational status*	Dummy indicating if the person worked in the week of reference
	Hourly wage	Hourly real wage in the previous month
Treatment	Technical school	Dummy for people with technical education
Controls	Age*	In the survey's reference date
	Gender*	Male or female
	Black*	Dummy for blacks
	Schooling*	Two dummies: high school or higher education
	Tenure	Length, in year, working in the current job
	Occupation	Brazilian Occupational Classification (CBO) groups
	Urban*	Dummy for urban areas
	MR*	Dummy for if the municipality is in a Metropolitan Areas
	UF*	Dummy for 26 states and the Federal District (minus the reference)
	RP*	Dummy for the reference person in the family
	School attendance*	Dummy for people still studying (any level of education)
	Children u14*	Number of people under or with 14 years-old in the family
	Children o14*	Number of people over 14 years-old in the family
	Married Couple*	Dummy for families with married couples
	Household income*	Household income without the personal income (if she have income)
	Public sector	A dummy for public sector jobs
	Industry	A dummy for Industry sector
	Commerce	A dummy for Commerce activity
	Formal job	A dummy for formal jobs

Source: PNAD.
Obs.: * Variables present in selection equation.

In 2007, the sample size was 399,964, while in 2014 the sample was 362,627, both representative to the entire country. The results here take into consideration the complex survey design of PNAD (see SILVA; PESSOA; LILA, 2002), and Table 48 shows the big

figures for both years.

Table 48: PNAD 2007 and 2014 – population over 18 years by technical education status and total population

	2007		2014	
Variable	N	%	N	%
People with Technical Education	7,451,167	5.7	8,606,057	5.8
People without Technical Education	123,797,241	94.3	140,089,887	94.2
Total of people 18 years-old and above	131,248,408	69.1	148,695,945	73.2
Total of population	189,955,482	100.0	203,190,817	100.0

Source: IBGE/PNAD

Brazil had almost 190 million people in 2007, increasing this figure in 13 million by 2014. The share of people with or above 18 years-old is increasing, while the share of those with technical education remained steady, around 5.8%. These figures are in accordance with RAIS figures, suggesting that both analyses will have similar target groups.

To contextualize the overall situation of labor market in Brazil, Table 49 shows the unemployment numbers of 2007 and 2014.

Table 49: PNAD 2007 and 2014 – unemployment for population over 18 years by technical education status

	2007		2014	
Variable	N	%	N	%
Unemployed (18 and above)	6,928,989	7.45	6,410,849	6.25
Unemployed (18 years)	464,625	21.06	471,476	22.96
Unemployed with Tech (18 years and above)	458,885	7.07	474,176	6.57
Unemployed with Tech (18 years)	14,820	28.64	26,328	26.24

Source: IBGE/PNAD

Overall, we see that the unemployment rate of young people is about three times of general's population in 2007 and a little higher in 2014. While for all population the unemployment rate is similar between people with or without technical education, for young with 18 years-old, the rate is higher for the technical group (which can be related to the "lock-in" effects, where people engaged in qualification may suffer a lower job opportunity in

the short run; see CARD; KLUVE; WEBER (2010) and CARD; KLUVE; WEBER (2018) for a discussion).

4.5 Results

In this section, we present first results of longitudinal study (from RAIS database), both for the unrestricted and filtered only to young with 18 years-old in 2007. Then, we present results for the two waves of PNAD (2007 and 2014), also for the unfiltered and filtered data. Next, we present some heterogeneous effects, followed by alternative estimates as a robustness exercise. Finally, we present a benefit versus costs analysis to bring the social return to discussion.

4.5.1 Longitudinal analysis: technical education returns from 2007 to 2018

Figure 12 shows the changes of the two groups' characteristics, technical workers and others.

Figure 12: Descriptive for all workers, 2007 to 2018

We notice that none of technical worker, by definition, had complete higher education in

2007. During the years, we find a qualification raise with the gap between two groups shortening. By the end of the period (2016 onward), technical workers are the majority with higher education. This result suggest that technical education is not the end of qualification for those workers, but the beginning of a higher education (which is in accordance with BISHOP; MANE (2004) and CORSEUIL; FOGUEL; GONZAGA (2019) findings). This raise in education can be related to lifelong learning discussion, where (**purcell_hard_2007?**) discuss about mature graduates and their difficulties in labor market, while (**jenkins_determinants_2003?**) conduction a longitudinal study of UK, with positive effects on wages[135]. Technical workers are a little younger, more male, work more in big companies, in the public sector and in industry (after 2009), but less in Commerce. While tenure is almost the same for both groups, location of jobs is quite different: technical workers are closer to the state capital, in Metropolitan Areas (MA) and in municipalities with Federal Institutes of Technology (FIT). Figures 19, 20 and 21 (in Appendix) show descriptive by CBO, Axis and CNAE, respectively.

Response variable is natural logarithm of real hourly wage. To calculate this variable we transform weekly hired hours into monthly hours, since wages are informed in month base[136]. Ratio between wage and hours worked was adjusted by consumer price index, taking as reference December of 2020[137].

We estimate our results with random effects[138], with SWAMY; ARORA (1972) parameter transformation.

Table 50 shows results for wage gains, considering all workers.

Table 50: Random effect estimation results for all workers, 2007 to 2018

Term	Estimate	Std. Error	T	p-value
Intercept	0.699	0.005	133.930	0.000
Age	0.063	0.000	370.402	0.000
Age²	-0.001	0.000	-290.848	0.000
Tenure	0.018	0.000	215.653	0.000
Tenure²	0.000	0.000	-51.170	0.000
Axis 1	0.145	0.018	8.278	0.000
Axis 2	0.700	0.028	25.283	0.000
Axis 3	0.390	0.123	3.162	0.002
Axis 4	1.502	0.023	66.378	0.000

[135] A wider discussion regarding lifelong learning can be found in UNESCO Institute for Lifelong Learning (https://uil.unesco.org/). In 2017, this subject was present a special report in The Economist: https://www.economist.com/special-report/2017-01-14.

[136] We consider 4.35 weeks per month.

[137] IPCA, calculated by IBGE

[138] More information about random effects are available in GREENE (2003), WOOLDRIDGE (2010), BALTAGI (2008) and CROISSANT; MILLO (2008)

Axis 5	0.024	0.012	2.091	0.037
Axis 6	2.840	0.058	49.336	0.000
Axis 7	0.110	0.018	6.129	0.000
Axis 8	-0.372	0.129	-2.879	0.004
Axis 9	7.995	0.137	58.457	0.000
Axis 10	-0.081	0.041	-1.954	0.051
Axis 11	-0.177	0.084	-2.106	0.035
Axis 12	18.449	0.185	99.524	0.000
Axis 13	0.978	0.108	9.055	0.000
Firm's size	0.076	0.001	139.584	0.000
Public Sector	0.135	0.001	129.543	0.000
CBO1	-0.026	0.004	-7.205	0.000
CBO2	-0.030	0.004	-8.141	0.000
CBO3	-0.161	0.004	-44.547	0.000
CBO4	-0.246	0.004	-68.869	0.000
CBO5	-0.325	0.004	-90.219	0.000
CBO6	-0.368	0.004	-93.469	0.000
CBO7	-0.274	0.004	-75.575	0.000
CBO8	-0.247	0.004	-65.509	0.000
CBO9	-0.266	0.004	-69.868	0.000
Industry	0.064	0.001	81.404	0.000
Commerce	-0.035	0.001	-49.248	0.000
GDP PP	0.002	0.000	135.609	0.000
PBF	0.000	0.000	38.457	0.000
BPC	0.000	0.000	129.789	0.000
MA	0.036	0.001	36.209	0.000
Distance	0.000	0.000	-49.691	0.000
FIT	0.046	0.001	73.483	0.000
High School	0.022	0.001	43.702	0.000
Higher education	0.224	0.001	283.299	0.000
Male	0.178	0.001	124.209	0.000
Technical CBO	0.193	0.003	61.687	0.000

Source: RAIS/ME

Obs.: R^2-Adj: 0.36; F = 1,977,380; DF1: 41; DF2: 3,967,497.

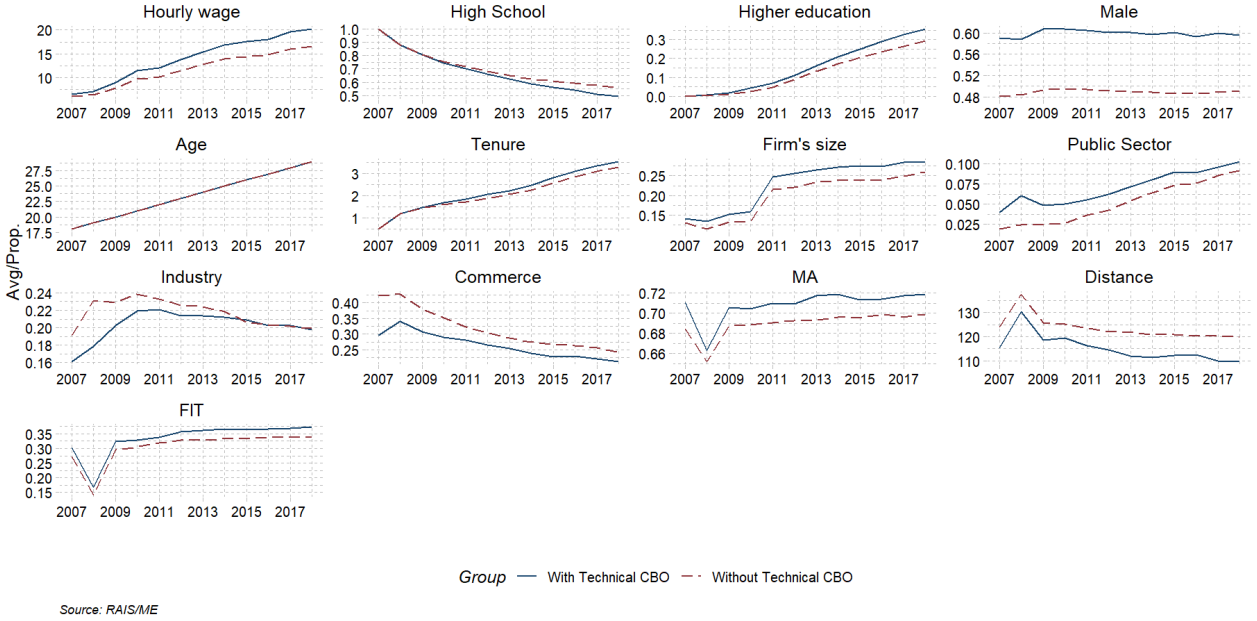
The return estimated is 21.3%[139], which is very similar to the effects reported in Brazilian literature (AGUAS, 2014; BARROS et al., 2011; OLIVEIRA; RIOS-NETO, 2007; VASCONCELLOS; LIMA; MENEZES-FILHO, 2010). Among the other characteristics contributing to higher wages, higher education and male are the most relevant. It is also interesting to notice that market variables (share of Axis workers in the total) are important in all the cases. Inside FERRACCI; JOLIVET; BERG (2014) spill-over effects of treatment discussion, these results suggest this effect is, in general, positive for wages. So, the expect

[139]Due to log transformation, the actual effect is $e^{\beta} - 1$.

result of higher shares of technical education in a market (and, hence, the competition) meaning lower wages is not prevalent. It might be the case that there is still room to increase technical jobs in Brazilian formal labor market (in other words, participation of technical jobs is low in market).

Despite the panel structure gives a temporal analysis, it might be the case that past tendency of wages is not the same for both groups. To address this limitation, we now look only to young with 18 years-old and high school in 2007. With this filter, we are dealing with people with no experience in formal (full time) job and, hence, no past wage trends[140] (Figure 13).

Figure 13: Descriptive of young 18 years-old, 2007 to 2018



Trends show remarkably similar characteristics in the beginning of the period: almost the same hourly wage, and, by definition, schooling, age and tenure. We still have some difference in formal labor market entrance between groups, where young technical workers are more in public sector and less in Industry and Commerce. It is interesting to notice that the gap in industrial sector is shortened, and, by the end of the period, figures are virtually the same. Like unfiltered database, technical workers are closer to state capital, live more in Metropolitan Areas (MA) and in municipalities with FIT. Table 51 shows the estimation for youth.

[140] By definition, they have the very same past wage history, i.e. no formal wage in a full time job before 2007.

Table 51: Random effect estimation results for young with 18 years-old in 2007, 2007 to 2018

Term	Estimate	Std. Error	t	p-value
Intercept	-1.840	0.014	-135.342	0.000
Age	0.303	0.001	273.121	0.000
Age²	-0.005	0.000	-221.638	0.000
Tenure	0.037	0.000	111.082	0.000
Tenure²	-0.002	0.000	-43.109	0.000
Axis 1	-1.076	0.036	-29.888	0.000
Axis 2	0.705	0.047	15.163	0.000
Axis 3	1.665	0.218	7.622	0.000
Axis 4	1.170	0.039	30.046	0.000
Axis 5	0.256	0.027	9.644	0.000
Axis 6	1.881	0.100	18.896	0.000
Axis 7	0.019	0.037	0.517	0.605
Axis 8	1.611	0.201	8.001	0.000
Axis 9	3.608	0.209	17.305	0.000
Axis 10	0.975	0.076	12.808	0.000
Axis 11	-1.038	0.156	-6.658	0.000
Axis 12	6.468	0.341	18.965	0.000
Axis 13	0.449	0.168	2.671	0.008
Firm's size	0.095	0.001	108.587	0.000
Public Sector	0.171	0.002	93.424	0.000
CBO1	-0.143	0.005	-26.338	0.000
CBO2	-0.127	0.005	-23.642	0.000
CBO3	-0.261	0.005	-49.193	0.000
CBO4	-0.365	0.005	-69.225	0.000
CBO5	-0.353	0.005	-66.719	0.000
CBO6	-0.349	0.006	-55.373	0.000
CBO7	-0.343	0.005	-64.231	0.000
CBO8	-0.310	0.006	-55.778	0.000
CBO9	-0.286	0.006	-50.651	0.000
Industry	0.066	0.001	64.272	0.000
Commerce	-0.007	0.001	-8.422	0.000
GDP PP	0.001	0.000	45.564	0.000
PBF	-0.001	0.000	-44.374	0.000
BPC	0.000	0.000	-3.134	0.002
MA	0.027	0.001	20.546	0.000
Distance	0.000	0.000	-28.840	0.000
FIT	0.039	0.001	37.370	0.000
Higher education	0.227	0.001	204.875	0.000
Male	0.100	0.001	75.359	0.000
Technical CBO	0.056	0.003	20.850	0.000

Source: RAIS/ME

Obs.: R^2-Adj: 0.57; F = 1,605,544; DF1: 40; DF2: 1,308,728.

When we focus on the restrict data base, i.e. young (aged 18), with high school degree, and who were in technical occupation in the beginning of the period (2007), wage premium still positive, but in a much lower magnitude, 5.8%. Results concerning other characteristics suggests that higher education is the most important, followed by holding a job in public sector. These workers are in a privileged position, once the initial wage premium is large for young people in public sector *vis-à-vis* private sector, with, in general, well defined career progression. The job market variables still have, in general, positive effects, suggesting the same spillover effect observed for the unfiltered data.

So far, when we analyze all workers, we find effects of technical jobs compatible with those estimated before in Brazil. However, when we target the analysis on young workers, with no experience in formal labor market, an important part of the effect vanishes, suggesting that estimations without taking into consideration the trend of earnings may be misleading. Also, it seems to exist an important generation difference in technical education returns, with disadvantage to the most recent cohorts.

4.5.2 Cross-section analysis: technical education returns in 2007 and 2014

We start the analysis showing the descriptive statistics for the two waves of survey, along with a mean difference t-test between groups (Table 52).

Table 52: Descriptive statistics and mean difference test, 2007 and 2014

	2007			2014		
Variable	Tech	No Tech	P.value	Tech	No Tech	P.value
Formal Job (%)	40.7	23.2	0.000	41.0	26.4	0.000
Married Couple (%)	74.2	73.4	0.066	72.8	71.6	0.013
Commerce (%)	13.7	11.6	0.000	12.5	11.6	0.004
High School (%)	65.2	21.2	0.000	60.6	25.7	0.000
Higher Education (%)	20.2	7.3	0.000	24.7	10.9	0.000
Children over 14	1.4	1.5	0.000	1.1	1.1	0.001
Children under 14	1.3	1.6	0.000	0.9	1.0	0.000
Student (%)	13.9	9.9	0.000	11.6	7.8	0.000
Public Sector (%)	13.1	4.0	0.000	10.5	4.2	0.000
Age	36.7	41.3	0.000	39.6	43.4	0.000
Industry (%)	14.0	9.6	0.000	12.7	8.0	0.000
Male (%)	47.8	47.7	0.936	50.7	47.3	0.000
Black (%)	37.1	48.8	0.000	42.2	53.2	0.000
Ref. Person (%)	42.5	43.0	0.306	47.2	45.0	0.000
Household Income (WPI)	2,024.9	1,372.3	0.000	3,223.7	2,356.8	0.000
MR (%)	41.2	31.8	0.000	40.4	31.3	0.000
Hourly wage	41.0	22.2	0.000	78.9	48.0	0.000

Tenure (%)	690.1	811.3	0.000	774.4	843.9	0.000
N# household members	3.6	3.9	0.000	3.3	3.5	0.000
Urban (%)	96.2	83.7	0.000	96.0	85.3	0.000

Source: PNAD/IBGE

In 2007, except for male, reference person and families with married couple, the two groups are different. By 2014, differences are present in all variables. So, profiles are statistically different, especially in hourly wage differences.

Once we have a potential selection problem, we run a first stage to create an Inverse Mills Ratio (IMR), following the seminal HECKMAN (1977) paper. Table 53 presents this first stage results.

Table 53: First stage regression – probit on occupation status, 2007 and 2014

	2007		2014	
Variable	Estimate	SD. Error	Estimate	SD. Error
Intercept	-0.996***	0.043	-1.276***	0.046
Age	0.089***	0.002	0.100***	0.002
Age²	-0.001***	0.000	-0.001***	0.000
High School	0.269***	0.009	0.249***	0.008
Higher Education	0.788***	0.018	0.719***	0.014
Male	0.661***	0.008	0.693***	0.008
Black	-0.008	0.007	-0.010	0.007
Married Couple	0.094***	0.008	0.065***	0.008
Urban	-0.450***	0.019	-0.306***	0.017
MR	-0.040***	0.010	0.022**	0.010
Ref. Person	0.382***	0.009	0.259***	0.008
Children under 14	-0.001	0.002	-0.003	0.003
Children over 14	0.009***	0.003	0.004	0.003
Student	-0.009	0.012	-0.060***	0.014
Household Income (WPI)	0.000***	0.000	0.000***	0.000
N# household members	-0.007**	0.002	-0.015***	0.003

Source: PNAD/IBGE

Obs.: significance level (* = at 10%; ** = at 5%; *** = at 1%)

As expected, age contributes positively (at decreasing rates) in job status, as well education (mainly higher education), gender (for males), families with married couples, the position in household (reference person), have older children and higher household income not considering the personal one (but not in a big magnitude). All positive estimates are consistent in the two years, except for children (no longer significant). Regarding the negative influence, we have urban and Metropolitan Areas and the number of household members in

2007. In 2014, Metropolitan Area changes its sign while be a student now lowers the chance to also have a job. It is interesting to notice that neither blacks nor young children have significant effects on occupational status. Overall, we suspect that it is likely that the IMR will be relevant in the wage equation, presented in Table 54.

Table 54: Wage equation for all workers, 2007 and 2014

	2007		2014	
Variable	Estimate	SD. Error	Estimate	SD. Error
Intercept	1.727***	0.057	2.417***	0.058
Technical Education	0.133***	0.008	0.112***	0.008
Age	0.044***	0.002	0.047***	0.002
Age²	0.000***	0.000	-0.001***	0.000
High School	0.208***	0.006	0.180***	0.006
Higher Education	0.824***	0.014	0.753***	0.014
Male	0.347***	0.011	0.393***	0.011
Black	-0.141***	0.005	-0.116***	0.004
Tenure	0.020***	0.001	0.017***	0.001
Tenure²	0.000***	0.000	0.000***	0.000
Urban	0.104***	0.014	0.085***	0.011
MR	0.140***	0.009	0.145***	0.008
CBO1	0.234***	0.025	0.058**	0.023
CBO2	0.020	0.024	-0.117***	0.021
CBO3	-0.084***	0.023	-0.234***	0.021
CBO4	-0.369***	0.023	-0.509***	0.021
CBO5	-0.556***	0.023	-0.622***	0.020
CBO6	-0.808***	0.028	-0.911***	0.024
CBO7	-0.504***	0.023	-0.525***	0.020
CBO8	-0.514***	0.025	-0.572***	0.024
CBO9	-0.432***	0.025	-0.473***	0.023
Public Sector	0.248***	0.009	0.209***	0.009
Industry	-0.014*	0.008	-0.092***	0.007
Commerce	-0.008	0.006	-0.072***	0.006
Formal Job	0.121***	0.006	0.065***	0.005
IMR	0.238***	0.028	0.404***	0.030

Source: PNAD/IBGE

Obs.: significance level (* = at 10%; ** = at 5%; *** = at 1%)

Indeed, in both years, we see that IMR enters with positive and significant sign. Also, we see that the effect of technical education is 14.3% in 2007 and 11.8% in 2014, both lower than the return estimate in panel analysis (which can be related to selection correction). So, even when selection is taken into consideration, return remain positive and compatible with Brazilian literature OLIVEIRA; RIOS-NETO (2007).

Although we have results compatible between panel and cross-section data for the unfiltered database, it could not be true when we look just for young with 18 years-old in the start point. Since we do not have a panel structure, but we have the age, we run an additional regression for people aged 25, in order to compare the situation of generation, that were 18 in 2007, seven years later (Table 55).

Table 55: Wage equation for people with 18 years-old (2007 and 2014) and for people aged 25 in 2014

	2007		2014		2014/25 years	
Variable	Estimate	SD. Error	Estimate	SD. Error	Estimate	SD. Error
Intercept	1.391***	0.228	2.688***	0.327	3.118***	0.136
Technical Education	0.249***	0.064	-0.068	0.060	0.097**	0.038
Male	0.290**	0.106	0.170	0.133	0.320***	0.052
Black	-0.064*	0.034	-0.052	0.032	-0.084***	0.023
Tenure	0.034	0.023	0.029	0.026	0.044***	0.009
Tenure^2	-0.004	0.003	-0.013**	0.004	-0.005***	0.001
Urban	-0.070	0.094	-0.073	0.105	0.105**	0.043
MR	0.110**	0.040	0.012	0.041	0.073**	0.026
CBO1	0.830***	0.220	0.037	0.233	0.101	0.098
CBO2	0.568***	0.169	0.106	0.196	0.024	0.088
CBO3	0.716***	0.144	0.299	0.208	-0.061	0.085
CBO4	0.445***	0.132	-0.065	0.185	-0.327***	0.082
CBO5	0.270**	0.137	-0.151	0.188	-0.438***	0.082
CBO6	0.326**	0.163	-0.241	0.219	-0.633***	0.096
CBO7	0.410**	0.135	-0.075	0.192	-0.349***	0.082
CBO8	0.461**	0.145	-0.200	0.209	-0.410***	0.109
CBO9	0.188	0.158	-0.135	0.208	-0.343***	0.094
Public Sector	0.050	0.108	-0.075	0.113	0.191***	0.054
Industry	-0.055	0.040	-0.037	0.054	-0.050	0.033
Commerce	0.004	0.041	-0.083**	0.036	-0.069**	0.026
Formal Job	0.163***	0.034	0.100**	0.037	0.072**	0.025
IMR	0.609**	0.295	0.367	0.366	0.364**	0.139

Source: PNAD/IBGE

Obs.: significance level (* = at 10%; ** = at 5%; *** = at 1%)

While we see a huge effect for youth in 2007, 28.2%, roughly the double of the full sample, the effect in 2014 is not different from zero (with a negative coefficient). When we look to results for people aged 25 in 2014, we see that positive effect is much lower, 10.2%, a little more than one third of the 2007's effect. If we admit that, in general, the sample of two years are comparable (same generation looked in two times), effect is decreasing in time. Comparing with longitudinal results, with a lower positive effect for young generation of 2007 until 2018, it suggests that effects may have decreased over the years. It is important to keep

in mind that 2014 was the beginning of a severe recession that afflicted Brazil in 2015 and 2016[141], so the adverse labor market may explain, in parts, the absence of effect observed. In this context, the so called "scarring effects" might be relevant, since the relationship between youth unemployment and permanent wage/job losses are reported in literature (**gregg_wage_2005?**; **schmillen_scars_2017?**). In a context of recession, which is the case of Brazil in the late of 2014, scarring effect may be boosted (**eliason_lasting_2006?**; **ouyang_scarring_2009?**). According to (**gangl_scar_2006?**), labor network protection plays a crucial role to alleviate this situation, even more because repeated unemployment exposure increases losses.

Hence, it is important to keep in mind that returns to technical education is neither constant over time nor is the same for different generations. It is necessary to update estimates when new data is available in order to assess if the conclusion remains or has changed.

4.5.3 Heterogeneous effects

It is common that treatment effects differ for subgroups of population from average effects. We have some features regarding technical education that should give different effects, mainly related to the type of activity. First, the public sector has an important role in the Brazilian labor market, with important wage premium that attracts workers (see BELLUZO; ANUATTI-NETO; PAZELLO (2005); HOLANDA BARBOSA; HOLANDA BARBOSA FILHO (2012); HOLANDA (2009); SOUZA; MEDEIROS (2013) for a discussion). Second, technical education has several activities usually related more to industrial jobs than the other sectors of economy. In order to investigate these possibilities, we rerun the results with three filters: only to private sector, only to industrial sector and only to commerce activity (results in Table 56 both for longitudinal and cross-section data).

Table 56: Heterogeneous effects of Technical CBO/Education on Private Sector, Industry and Commerce. All workers and young (18 years-old)

	Longitudinal		Cross-section			
Group	Full	Young	Full 2007	Young 2007	Full 2014	Young 2014
Private	0.146***	0.048***	0.152***	0.206***	0.121***	-0.024
	(0.002)	(0.002)	(0.009)	(0.060)	(0.009)	(0.058)
Industry	0.152***	0.068***	0.228***	0.362***	0.212***	0.092
	(0.003)	(0.003)	(0.019)	(0.082)	(0.020)	(0.119)

[141]In 2014, the GDP growth was only 0.5%, and, in the next two years, it felt -3.3% and -3.5%. The unemployment rate rose from 6.5% in the last quarter of 2014 to 12% in the same period of 2016, according to PNADC/IBGE.

Commerce	0.118***	0.029***	0.120***	0.221*	0.091***	0.107
	(0.003)	(0.003)	(0.019)	(0.124)	(0.019)	(0.068)

Source: RAIS and PNAD/IBGE

Obs.1: significance level (* = at 10%; ** = at 5%; *** = at 1%)

Obs.2: SE in parenthesis

For all workers and in longitudinal data, estimates are, in general, similar among sectors. When we consider only private sector, difference from general return is 5.7 p.p, when we consider industry, difference is 4.9 p.p. and for commerce, difference is 8.8 p.p. suggesting the highest difference for last sector. When we look to young with 18 years-old in 2007, we see that differences for private, industry and commerce are 0.8 p.p, -1.2 p.p and 2.9 p.p respectively, a lower difference than for unfiltered data.

Looking now to cross-section data, we observe that wage premium is higher for industry, both for all workers and for young, with 11.4 p.p. difference (25.6% return) and 15.3 p.p. difference (43.6% return) in 2007, respectively. In 2014, return to industry is also higher than to general estimate, but only for all workers (lower than 2007), with 11.8 p.p. difference (23.6% return). For young with 18 years-old, effects remain no significant, suggesting a different picture among generations.

Thus, technical education seems to enjoy higher returns from industry in a given year (when cross sectional data is considered), but not in the longitudinal context. Regarding the higher returns for industry, it is natural result if we think these activities as more technological than others, like commerce. These heterogeneous effects are in line with results reported by CARRUTHERS; SANFORD (2018) and STEVENS; KURLAENDER; GROSZ (2019).

4.5.4 Robustness tests

Despite all richness and sample size of our data, there might remain some selection into technical education not fully controlled. So, to deal with this issue, we match people with technical occupation to the rest of the database in the ratio of two to one. We use the nearest neighbor match, with a *logit* distance (propensity score)[142], considering all variables used in wage regression. Also, we increase the number of observations, both to unfiltered and filtered data (using 2007 data as reference, 5% for the former and 20% for the latter) and, alternatively, we restrict our data only to people who appears in all the years (i.e., we have 12 observations, from 2007 to 2018, to all sample), constructing a balanced panel. Descriptive are in Figures 22 and 23 (longitudinal data) and Tables 68 and 69 (cross-section data) in Appendix, while results are in Table 57.

[142]Matching are conduct with "MatchIt" R package (HO et al., 2018).

Table 57: Matching results for longitudinal and cross-section database. All workers and young (18 years-old)

Database	Estimate	SE	t	P-value
RAIS - all workers	0.223	0.002	114.6	0.000
RAIS - all workers (balanced)	0.224	0.002	123.8	0.000
RAIS - Young (18 years-old)	0.075	0.002	36.6	0.000
PNAD 2007 - all workers	0.128	0.009	14.7	0.000
PNAD 2007 - Young (18 years-old)	0.230	0.082	2.8	0.006
PNAD 2014 - all workers	0.122	0.009	13.0	0.000
PNAD 2014 - Young (18 years-old)	-0.081	0.066	-1.2	0.225

Source: RAIS/ME and PNAD/IBGE
Obs.: SE = Standard Error

We observe that conclusions remain quite the same. Results seems to be robust both for longitudinal and cross-sectional data, offering, virtually, no changes in conclusion made in previous sections.

To investigate furthermore, Table 58 presents estimates for the random effects within-between model for all workers. Due to computational limitation, associated with the necessity of a more parsimonious specification (BELL; FAIRBROTHER; JONES, 2019), we restrict estimation to the matched data and with fewer covariates (omitting the quadratic of age and tenure, industry and commerce dummies, CBO, axes, BPC, MA and distance). We looked to keep all variables with a closer relationship to workers characteristics, focusing in dropping environmental variables.

Table 58: Random-effects-between-within for all workers, 2007 to 2018

Variable	Estimate	SE	t	P value
Intercept	1.468	0.008	194.2	0.000
Year	0.041	0.001	54.2	0.000
Age Bw	0.006	0.000	38.0	0.000
Age Wi	-0.008	0.001	-9.5	0.000
Tenure Bw	0.038	0.000	154.3	0.000
Tenure Wi	0.005	0.000	23.3	0.000
FIT Bw	-0.017	0.003	-5.5	0.000
FIT Wi	0.041	0.002	26.7	0.000
Public sector Bw	-0.034	0.005	-7.3	0.000
Public sector Wi	0.117	0.004	32.4	0.000
Firms' size Bw	0.246	0.004	60.1	0.000
Firms' size Wi	0.058	0.002	37.0	0.000
GDP pp Bw	0.006	0.000	68.7	0.000
GDP pp Wi	0.003	0.000	55.0	0.000

PBF Wi	-0.002	0.000	-62.4	0.000
PBF Bw	0.001	0.000	38.5	0.000
Higher education Bw	0.853	0.008	112.9	0.000
Higher education Wi	0.094	0.002	40.1	0.000
High school Bw	-0.011	0.004	-2.6	0.008
High school Wi	-0.016	0.002	-9.7	0.000
Male	0.259	0.003	90.8	0.000
Technical CBO	0.211	0.003	74.4	0.000

Source: RAIS/ME

Obs.1: SE = Standard Error; Bw = Between; Wi = Within.

Obs.2: N: 1,068,146; n: 129,754; Marg. R^2: 0.43; Cond. R^2: 0.95.

Once again, estimates are very closer to our standard model. Here, return to technical occupation is 23.5%, which is 2.2 p.p. higher than standard estimates, and -1.4 p.p. lower than matched database estimates. Since this model allows us to distinguish within (individual related) from between coefficient (time related), we observe that the former class is more important for FIT and public sector, while the latter class is more relevant to age, tenure, firm's size and higher education. Table 59 presents the same results for young workers (18 years-old in 2007).

Table 59: Random-effects-between-within for young workers (18 years-old in 2007), 2007 to 2018

Variable	Estimate	SE	t	P value
Intercept	1.351	0.018	75.4	0.000
Year	0.067	0.001	127.4	0.000
Age Bw	0.009	0.001	10.1	0.000
Tenure Bw	0.048	0.001	37.1	0.000
Tenure Wi	0.009	0.001	14.7	0.000
FIT Bw	0.021	0.004	5.7	0.000
FIT Wi	0.044	0.003	12.9	0.000
Public sector Bw	0.042	0.009	4.7	0.000
Public sector Wi	0.184	0.008	22.2	0.000
Firms' size Bw	0.172	0.006	30.9	0.000
Firms' size Wi	0.079	0.003	24.0	0.000
GDP pp Bw	0.003	0.000	37.7	0.000
GDP pp Wi	0.002	0.000	26.8	0.000
PBF Wi	-0.002	0.000	-31.6	0.000
PBF Bw	0.000	0.000	7.6	0.000
Higher education Bw	0.472	0.008	56.4	0.000
Higher education Wi	0.132	0.005	28.4	0.000
Male	0.097	0.003	32.5	0.000
Technical CBO	0.072	0.003	23.3	0.000

Source: RAIS/ME

Obs.1: SE = Standard Error; Bw = Between; Wi = Within. Age whithin was dropped due to lack of variance.
Obs.2: N: 230,638; n: 34,806; Marg. R^2: 0.42; Cond. R^2: 0.89

Like the unfiltered data, estimates are very closer to the standard model. The return to technical occupation estimated by REWB is 7.5%, which is 1.7 p.p. higher than the standard estimates, and -0.3 p.p. lower than match estimates. Conclusions regarding the between and within parameters are quite the same.

In sum, we observe that our estimates are very robust to alternative models (standard random effects, with or without matched data, and the within-between version of random effects), all of them suggesting the same pattern for technical education returns. For all workers, returns are higher (between 21.3% and 24.9%) than for young workers (between 5.8% and 7.8%), suggesting differences by generation.

4.5.5 Cost and benefit analysis

As said before, technical education, usually, has a lower Rate of Return (ROR) due to its higher unit costs (PSACHAROPOULOS, 1994). Thus, a natural question is if returns estimated here are enough to face costs of technical education. Unfortunately, we are not able to estimate all costs related to this type of education[143], but we try to give some figures, helping the debate about ROR.

First, we will make some assumptions: (i) the average effect estimate with matching database in the 12 years of our panel will be the same for, at least, 35 years of the labor life; (ii) we take three interest rates to bring the flow of benefits to present values (6%, 12% and 18%); (iii) we assume, during this period, that people will work the same average weekly number of hours. Taking the current Brazilian inflation target (3.75% in 2021 down to 3.25% in 2023), even with its 1.5% tolerance, the rates chosen are compatible with real positive rates of return. Finally, assuming that technical education usually demands 18 months, we estimate the maximum monthly amount someone could expend to reach the rates of return suggested (Table 60).

Table 60: Cost and benefit analysis of technical education returns according to longitudinal matching database

Estimate	Effect per hour	Return in PV	Cost per month	Rate
		154,712	8,595	6%

[143]Brazil has public and private schools offering technical education. In public system, we may have schools maintained by municipal, states or federal governments, with several transfers between them. Also, some schools offer academic and technical courses, making hard to distinguish the costs.

All workers	5.01	87,241	4,847	12%
		59,103	3,283	18%
		38,832	2,157	6%
Young 18 years-old	1.22	21,897	1,217	12%
		14,834	824	18%

Source: RAIS/ME

Obs.1: Number of weekly hours worked are 40.8 and 42.2 for all workers and young, respectively.

Obs.2: We consider 4.35 weeks in a month and 35 years of work.

Obs.3: Costs considering 18 months of a standard technical course.

Obs.4: Estimates from standard random effects with matched data.

The National Education's Plan[144] established minimum parameters for educational quality and ARAÚJO et al. (2016) estimated the costs by educational phase. The yearly cost per student estimated in 2015, for high school and technical education, was R$ 6,111.16 and R$ 7,944.50 (which gives in Dec/2020 figures R$ 7,562.93 and R$ 9,831.80, respectively). So, the return estimated here covers between 13.6 times and 5.2 times of high school costs, or between 10.5 times and 4 times of technical education costs (depending on rate of return considered). Nonetheless, high school is mandatory to all young Brazilian, so if we consider only the difference between the ideal cost of high school and technical education, R$ 2,268.87, return would be even higher. It is important to remember that ideal cost is not the actual amount spent by the government, so returns are enough to cover the current costs.

However, if we consider the effect just for young workers, returns covers between 3.4 times and 1.3 times the high school costs, and between 2.6 times and 1.01 times technical education costs. It is important to remember that we are not considering the possibility of economies of scale, like when student takes, at the same time, high school and technical education. Also, we are not able to assess any externality, positive or negative, that may alter results. Even though, our results are important to bring to debate the limits until the technical education might be economic and offering update estimates to young interested in pursuit this path.

4.6 Conclusions and remarks

We update estimate of technical education on wages using longitudinal and cross-section data. We were able to consider time dimension with RAIS data, while deal with the selection into labor market with PNAD data. As proxy for technical education (in panel data), we consider people who, in 2007, was holding a job related to technical course, according to

[144] *Plano Nacional de Educação*, available in: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l13005.htm

National Catalog of Technical Courses, and had high school diploma or incomplete college. Our results suggest a positive and significant wage difference in favor of technical workers. Wages premium, between 21.3% and 24.9% for all workers, and between 5.8% and 7.8% for young workers, indicate that technical education offers, in short or in long run, a good job opportunity. Cross-section analysis confirms such numbers, although not for the young generation in 2014, suggesting that generations may be affected differently over the years. The magnitude of estimates is consistent with previous works about this issue OLIVEIRA; RIOS-NETO (2007).

For students who are finishing academic life and looking for a quick entry into labor market, technical education and, therefore, a technical education, could be an interesting alternative to improve school-to-work transition. Also, the positive effect found here could encourage students to pursuit the technical education path. For Brazil, with only less than 6% of its adults with technical education, public policies towards strengthening qualification of work force may enhance average schooling of labor force, increasing, simultaneously, wage opportunities.

5 Conclusions

Public evaluation is a key task to achieve efficiency, one of the principles of the public administration in Brazil. We contribute studying two programs, Backhaul and Pronatec, and its impacts on political and labor outcomes, respectively. In addition, we also studied the situation of technical education in Brazil in the past decade. As was shown in introduction, roughly one third of GDP passes through government, and, even with this size, little is known about its efficiency or program effects.

In the first essay, we showed that outcomes present in other countries, regarding the relationship between broadband internet and political outcomes, are not present in Brazil. However, our study focused only on the beginning of fixed broadband connections expansion. At that time (2008 to 2012), personal computers were required to connect, which was not a wide reality in the country. Also, social media was still expanding, boosted by mobile broadband connection (3G) and the launch of smartphones. Despite we found no evidence of relationship between high-speed internet and political outcomes – turnout, percentage of blank or null votes, changes in vote share of some parties nor changes in budget campaign of young candidates – the question still relevant. The last Brazilian election was heavily based on internet and social media, which now reach the majority of electorate, enabling a candidate with just eight seconds in the traditional television campaign conquer the presidency. Everything points toward a highly competitively race in 2022, where the internet may still be the decisive battleground, even more in a scenario where misleading information has flourished in recent years, so keeping the research question open.

In the second essay, we showed that Pronatec program is highly economical (between six to eighty-four months after training conclusion). Taking advantage of the quasi-experimental database emerged from program implementation rules, we showed higher odds to hold a formal job in favor of participants as well wage premiums, in short, middle, and long run. This is an important evidence of program effectiveness, considering the darkness scenario for the labor market in the recent years in Brazil. Also, we saw that, by the size of the program, Pronatec had important spillover effects, affecting, in general, negatively both wages and employment in the short run. The beginning of Pronatec suggests that job market was flooded with low-cost training ("shelf courses"), which seems not to be the best way in applying public resources. Considering heterogeneous effects, the program can be improved to maximize returns. However, it is important to remind that Pronatec has an assistentialist purpose, so possible external benefits, like educational inclusion, was not included in analysis.

In the third essay, we saw that technical workers are a small fraction of Brazilian work

force (less than 6%), despite the considerable high wage premium (around 20%). Analyzing the formal labor market with a panel from 2007 to 2018, we saw a favorable scenario for workers that initiate the period in a technical occupation. Cross-section analysis, considering the entire labor market, confirms results, although with a lower magnitude. However, when a young generation is considered, results are less favorable. In the panel analysis, wage premium is reduced to around 7% for those aged 18 in the beginning of the period, while, in the cross-sectional data for 2014, the wage-premium was not different from zero for this cohort. Another important result is related to education, where descriptive statistics showed that technical education seems to be a door to higher education. Returns estimated with panel data suggest that, given an ideal cost for technical education, returns covers by a huge margin their costs. Putting all results together, considering that today unemployment rate for young is twice the general rate, reaching three times for people aged 18, technical education may still be an alternative, because it has short duration when compared to higher education and, also, because it seems not to be the end of educational life. Also, looking to Pronatec results, a well target training program could be presented as a short/mid run alternative to face the unemployment harm afflicting Brazilian workers in the recent years.

Finally, it is important to remember that our analysis relied on quasi-experimental (the first two essays) and non-experimental designs to bring evidence. These are not the best ways in providing solid evidence about public policies, but what are currently available in Brazil. We hope that, by contributing with these results, public programs consider incorporating evaluation strategies in the design phase. With well-defined target indicator the program or police wants to change, the collection of data and identification strategy embodied in its conception, evaluation would be more robust and easier to be carried out. So, the principle of efficiency in public sector could be truly fulfilled, helping to deliver the many services Brazilian constitution promises to citizens.

References

ABADIE, A.; IMBENS, G. W. Large sample properties of matching estimators for average treatment effects. **econometrica**, v. 74, n. 1, p. 235–267, 2006.

ACEMOGLU, D.; PISCHKE, J.-S. Beyond Becker: Training in imperfect labour markets. **The economic journal**, v. 109, n. 453, p. 112–142, 1999.

AGUAS, M. F. F. **Ensaios sobre a educação profissional e os rendimentos do trabalho: uma análise para o Brasil**. phdthesis—[s.l: s.n.].

ALDRICH, J. H. Rational choice and turnout. **American journal of political science**, p. 246–278, 1993.

ALI, S. N.; LIN, C. Why people vote: Ethical motives and social incentives. **American economic Journal: microeconomics**, v. 5, n. 2, p. 73–98, 2013.

ALLCOTT, H.; GENTZKOW, M. Social media and fake news in the 2016 election. **Journal of economic perspectives**, v. 31, n. 2, p. 211–36, 2017.

ALVES, E. L. G.; SANTOS VIEIRA, C. A. DOS. Qualificação profissional: uma proposta de política pública. **Planejamento e políticas públicas**, n. 12, 2009.

ANATEL. **Relatório de acompanhamento das metas de implementação da infraestrutura de rede de suporte do STFC para conexão em banda larga (Backhaul)**Superintendência de Universalização, 2010. Disponível em: <https://www.anatel.gov.br/Portal/verificaDocumentos/documento.asp?numeroPublicacao=240258>

ANDERSSON, F. et al. **Does federally-funded job training work? Nonexperimental estimates of WIA training impacts using longitudinal data on workers and firms**. [s.l.] National Bureau of Economic Research, 2013.

ANGRIST, J. D. American education research changes tack. **Oxford review of economic policy**, v. 20, n. 2, p. 198–212, 2004.

ARAÚJO, H. E. et al. Quanto custa o Plano Nacional de Educação?: uma estimativa

orientada pelo Custo Aluno Qualidade (CAQ). 2016.

ASHENFELTER, O. Estimating the effect of training programs on earnings. **The Review of Economics and Statistics**, p. 47–57, 1978.

ASHENFELTER, O. C.; CARD, D. **Using the longitudinal structure of earnings to estimate the effect of training programs**. [s.l.] National Bureau of Economic Research Cambridge, Mass., USA, 1984.

ATTANASIO, O. et al. Vocational training for disadvantaged youth in colombia: A long-term follow-up. **American Economic Journal: Applied Economics**, v. 9, n. 2, p. 131–43, 2017.

ATTANASIO, O.; KUGLER, A.; MEGHIR, C. Subsidizing vocational training for disadvantaged youth in Colombia: evidence from a randomized trial. **American Economic Journal: Applied Economics**, v. 3, n. 3, p. 188–220, 2011.

BALTAGI, B. **Econometric analysis of panel data**. [s.l.] John Wiley & Sons, 2008.

BANERJEE, A. et al. Do informed voters make better choices? Experimental evidence from urban India. **Unpublished manuscript**, 2011.

BARNOW, B. S. Thirty years of changing federal, state, and local relationships in employment and training programs. **Publius: The Journal of Federalism**, v. 23, n. 3, p. 75–94, 1993.

BARNOW, B. S. Vouchers in US vocational training programs: an overview of what we have learned. **Zeitschrift für Arbeitsmarktforschung**, v. 42, n. 1, p. 71–84, 2009.

BARROS, R. et al. Educação técnica e distribuição de renda no Espírito Santo. **Revista Brasileira de Monitoramento e Avaliação**, n. 1, p. 104–135, 2011.

BASTO, J. B. T. et al. **Can program design improve the effectiveness of worker retraining? An evaluation of Brazil's Pronatec-MDIC**. PET 16-Rio. **Anais**...2016.

BECKER, G. S. Investment in human capital: A theoretical analysis. **The journal of political economy**, p. 9–49, 1962.

BELL, A.; FAIRBROTHER, M.; JONES, K. Fixed and random effects models: making an informed choice. **Quality & Quantity**, v. 53, n. 2, p. 1051–1074, 2019.

BELLUZO, W.; ANUATTI-NETO, F.; PAZELLO, E. T. Distribuição de Salários e o Diferencial Público-Privado no Brasil. v. 4, n. 59, p. 511–53, 2005.

BENNELL, P. General versus vocational secondary education in developing countries: a review of the rates of return evidence. **The Journal of Development Studies**, v. 33, n. 2, p. 230–247, 1996.

BEN-PORATH, Y. The production of human capital and the life cycle of earnings. **Journal of political economy**, v. 75, n. 4, Part 1, p. 352–365, 1967.

BESLEY, T.; PRAT, A. Handcuffs for the grabbing hand? Media capture and government accountability. **American economic review**, v. 96, n. 3, p. 720–736, 2006.

BIAVASCHI, C. et al. Youth unemployment and vocational training. 2012.

BIMBER, B. The Internet and political mobilization: Research note on the 1996 election season. **Social Science Computer Review**, v. 16, n. 4, p. 391–401, 1998.

BIMBER, B. **Information and American democracy: Technology in the evolution of political power**. [s.l.] Cambridge University Press, 2003.

BISHOP, J. H.; MANE, F. The impacts of career-technical education on high school labor market success. **Economics of education Review**, v. 23, n. 4, p. 381–402, 2004.

BLASCO, S.; CRÉPON, B.; KAMIONKA, T. The effects of on-the-job and out-of-employment training programmes on labour market histories. **Documents de travail du Centre pour la Recherche Économique et ses Applications (Cepremap)**, 2012.

BLASCO, S.; ROSHOLM, M. The impact of active labour market policy on post-unemployment outcomes: evidence from a social experiment in Denmark. **IZA Discussion Paper No. 5631**, 2011.

BOUERI, R.; ROCHA, F.; RODOPOULOS, F. Avaliação da qualidade do gasto público

e mensuração da eficiência. **Rogério Boueri, Fabiana Rocha, Fabiana Rodopoulos (Organizadores)-Brasília: Secretaria do Tesouro Nacional**, 2015.

BOULDING, C.; BROWN, D. S. Do political parties matter for turnout? Number of parties, electoral rules and local elections in Brazil and Bolivia. **Party Politics**, v. 21, n. 3, p. 404–416, 2015.

BROWN, A. J.; KOETTL, J. Active labor market programs-employment gain or fiscal drain? **IZA Journal of Labor Economics**, v. 4, n. 1, p. 12, 2015.

BRUNELLO, G.; COMI, S. L.; SONEDDA, D. Training subsidies and the wage returns to continuing vocational training: Evidence from Italian regions. **Labour Economics**, v. 19, n. 3, p. 361–372, 2012.

CAHUC, P.; LE BARBANCHON, T. Labor market policy evaluation in equilibrium: Some lessons of the job search and matching model. **Labour Economics**, v. 17, n. 1, p. 196–205, 2010.

CALONICO, S. et al. rdrobust: Software for regression-discontinuity designs. **The Stata Journal**, v. 17, n. 2, p. 372–404, 2017.

CALONICO, S.; CATTANEO, M. D.; FARRELL, M. H. On the effect of bias estimation on coverage accuracy in nonparametric inference. **Journal of the American Statistical Association**, v. 113, n. 522, p. 767–779, 2018.

CALONICO, S.; CATTANEO, M. D.; TITIUNIK, R. Robust nonparametric confidence intervals for regression-discontinuity designs. **Econometrica**, v. 82, n. 6, p. 2295–2326, 2014.

CAMPANTE, F.; DURANTE, R.; SOBBRIO, F. Politics 2.0: The multifaceted effect of broadband internet on political participation. **Journal of the European Economic Association**, v. 16, n. 4, p. 1094–1136, 2017.

CANCELA, J.; GEYS, B. Explaining voter turnout: A meta-analysis of national and subnational elections. **Electoral Studies**, v. 42, p. 264–275, 2016.

CARD, D.; KLUVE, J.; WEBER, A. Active labour market policy evaluations: A meta-analysis. **The economic journal**, v. 120, n. 548, 2010.

CARD, D.; KLUVE, J.; WEBER, A. What works? A meta analysis of recent active labor market program evaluations. **Journal of the European Economic Association**, v. 16, n. 3, p. 894–931, 2018.

CARRUTHERS, C. K.; SANFORD, T. Way station or launching pad? Unpacking the returns to adult technical education. **Journal of Public Economics**, v. 165, p. 146–159, 2018.

CATTANEO, M. D.; JANSSON, M.; MA, X. Simple local polynomial density estimators. **Journal of the American Statistical Association**, p. 1–7, 2019.

CATTANEO, M. D.; JANSSON, M.; MA, X. Local regression distribution estimators. **Journal of Econometrics**, 2021.

CATTANEO, M. D.; TITIUNIK, R.; VAZQUEZ-BARE, G. Analysis of Regression Discontinuity Designs with Multiple Cutoffs or Multiple Scores. p. 18, 2018.

CORSEUIL, C. H.; FOGUEL, M. N.; GONZAGA, G. Apprenticeship as a stepping stone to better jobs: Evidence from Brazilian matched employer-employee data. **Labour Economics**, v. 57, p. 177–194, 2019.

CRÉPON, B.; FERRACCI, M.; FOUGÈRE, D. Training the unemployed in France: how does it affect unemployment duration and recurrence? **Annals of Economics and Statistics/ANNALES D'ÉCONOMIE ET DE STATISTIQUE**, p. 175–199, 2012.

CRÉPON, B.; VAN DEN BERG, G. J. Active labor market policies. **Annual Review of Economics**, v. 8, p. 521–546, 2016.

CROISSANT, Y.; MILLO, G. Panel data econometrics in R: The plm package. **Journal of statistical software**, v. 27, n. 2, 2008.

CZERNICH, N. Broadband Internet and Political Participation: Evidence for G ermany. **Kyklos**, v. 65, n. 1, p. 31–52, 2012.

DAVIDSON, C.; WOODBURY, S. A. The displacement effect of reemployment bonus programs. **Journal of Labor Economics**, v. 11, n. 4, p. 575–605, 1993.

DELLAVIGNA, S.; KAPLAN, E. The Fox News effect: Media bias and voting. **The Quarterly Journal of Economics**, v. 122, n. 3, p. 1187–1234, 2007.

DOERR, A.; STRITTMATTER, A. Assignment mechanisms, selection criteria, and the effectiveness of training programmes. 2014.

DOLTON, P.; O'NEILL, D. Unemployment duration and the restart effect: some experimental evidence. **The Economic Journal**, v. 106, n. 435, p. 387–400, 1996.

DOUGHERTY, S. M. The effect of career and technical education on human capital accumulation: Causal evidence from Massachusetts. **Education Finance and Policy**, v. 13, n. 2, p. 119–148, 2018.

DOWNS, A. An economic theory of political action in a democracy. **Journal of political economy**, v. 65, n. 2, p. 135–150, 1957.

DRAGO, F.; NANNICINI, T.; SOBBRIO, F. Meet the press: How voters and politicians respond to newspaper entry and exit. **American Economic Journal: Applied Economics**, v. 6, n. 3, p. 159–88, 2014.

DURANTE, R.; KNIGHT, B. Partisan control, media bias, and viewer responses: Evidence from Berlusconi's Italy. **Journal of the European Economic Association**, v. 10, n. 3, p. 451–481, 2012.

ENIKOLOPOV, R.; PETROVA, M.; ZHURAVSKAYA, E. Media and political persuasion: Evidence from Russia. **The American Economic Review**, v. 101, n. 7, p. 3253–3285, 2011.

FALCK, O.; GOLD, R.; HEBLICH, S. E-lections: Voting Behavior and the Internet. **American Economic Review**, v. 104, n. 7, p. 2238–65, 2014.

FEREJOHN, J. A.; FIORINA, M. P. The paradox of not voting: A decision theoretic analysis. **American political science review**, v. 68, n. 2, p. 525–536, 1974.

FERRACCI, M.; JOLIVET, G.; BERG, G. J. VAN DEN. Evidence of treatment spillovers within markets. **Review of Economics and Statistics**, v. 96, n. 5, p. 812–823, 2014.

FERRAZ, C.; FINAN, F. Exposing corrupt politicians: the effects of Brazil's publicly released audits on electoral outcomes. **The Quarterly journal of economics**, v. 123, n. 2, p. 703–745, 2008.

FERSTERER, J.; PISCHKE, J.-S.; WINTER-EBMER, R. Returns to apprenticeship training in Austria: Evidence from failed firms. **Scandinavian journal of economics**, v. 110, n. 4, p. 733–753, 2008.

FRAKER, T.; MAYNARD, R. The adequacy of comparison group designs for evaluations of employment-related programs. **Journal of Human Resources**, p. 194–227, 1987.

FRIGOTTO, G. Concepções e mudanças no mundo do trabalho e o ensino médio. **Ensino médio integrado: concepção e contradições. São Paulo: Cortez**, p. 57–82, 2005.

FUJIWARA, T.; MENG, K.; VOGL, T. Habit formation in voting: Evidence from rainy elections. **American Economic Journal: Applied Economics**, v. 8, n. 4, p. 160–88, 2016.

GAVAZZA, A.; NARDOTTO, M.; VALLETTI, T. Internet and politics: Evidence from UK local elections and local government policies. **The Review of Economic Studies**, v. 86, n. 5, p. 2092–2135, 2019.

GELMAN, A.; IMBENS, G. Why high-order polynomials should not be used in regression discontinuity designs. **Journal of Business & Economic Statistics**, v. 37, n. 3, p. 447–456, 2019.

GENTZKOW, M. Television and voter turnout. **The Quarterly Journal of Economics**, v. 121, n. 3, p. 931–972, 2006.

GENTZKOW, M.; SHAPIRO, J. M.; SINKINSON, M. The effect of newspaper entry and exit on electoral politics. **American Economic Review**, v. 101, n. 7, p. 2980–3018, 2011.

GERBER, A. S.; KARLAN, D.; BERGAN, D. Does the media matter? A field experiment

measuring the effect of newspapers on voting behavior and political opinions. **American Economic Journal: Applied Economics**, v. 1, n. 2, p. 35–52, 2009.

GÖRG, H.; STROBL, E. Do government subsidies stimulate training expenditure? Microeconometric evidence from plant-level data. **Southern Economic Journal**, p. 860–876, 2006.

GREENBERG, D. H.; MICHALOPOULOS, C.; ROBINS, P. K. A meta-analysis of government-sponsored training programs. **ILR Review**, v. 57, n. 1, p. 31–53, 2003.

GREENE, W. H. **Econometric analysis**. [s.l.] Pearson Education India, 2003.

GRUBB, W. N. **Working in the Middle: Strengthening Education and Training for the Mid-Skilled Labor Force.** [s.l.] ERIC, 1996.

HANUSHEK, E. A. et al. Education and economic growth. **Education Next**, v. 8, n. 2, 2008.

HANUSHEK, E. A.; MACHIN, S. J.; WOESSMANN, L. **Handbook of the Economics of Education**. [s.l.] Elsevier, 2016.

HECKMAN, J. J. **Sample selection bias as a specification error (with an application to the estimation of labor supply functions)**. [s.l.] National Bureau of Economic Research Cambridge, Mass., USA, 1977.

HECKMAN, J. J.; ICHIMURA, H.; TODD, P. Matching as an econometric evaluation estimator. **The Review of Economic Studies**, v. 65, n. 2, p. 261–294, 1998.

HECKMAN, J. J.; ICHIMURA, H.; TODD, P. E. Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. **The review of economic studies**, v. 64, n. 4, p. 605–654, 1997.

HECKMAN, J. J.; LALONDE, R. J.; SMITH, J. A. The economics and econometrics of active labor market programs. In: **Handbook of labor economics**. [s.l.] Elsevier, 1999. v. 3p. 1865–2097.

HECKMAN, J. J.; SMITH, J. A. The determinants of participation in a social program: Evidence from a prototypical job training program. **Journal of Labor Economics**, v. 22, n. 2, p. 243–298, 2004.

HIPP, L.; WARNER, M. E. Market forces for the unemployed? Training vouchers in Germany and the USA. **Social Policy & Administration**, v. 42, n. 1, p. 77–101, 2008.

HO, D. et al. **Package 'MatchIt'**. [s.l.] Version, 2018.

HOLANDA, A. L. N. Diferencial de salários entre os setores público e privado: uma resenha da literatura. v. 1457, n. IPEA, 2009.

HOLANDA BARBOSA, A. L. N.; HOLANDA BARBOSA FILHO, F. Diferencial de salários entre os setores público e privado no Brasil: um modelo de escolha endógena. n. IPEA, 2012.

HORACIO, A.; MONTEIRO, J. C. Media Networks and Political Accountability: Evidence from Radio Networks in Brazil. 2014.

JABER, A. Broadband Internet and Political Behavior: Evidence from the United States. **Available at SSRN 2353549**, 2013.

JESPERSEN, S. T.; MUNCH, J. R.; SKIPPER, L. Costs and benefits of Danish active labour market programmes. **Labour economics**, v. 15, n. 5, p. 859–884, 2008.

JOHNES, G.; JOHNES, J. **International handbook on the economics of education**. [s.l.] Edward Elgar Publishing, 2007.

KAHYARARA, G.; TEAL, F. The returns to vocational training and academic education: Evidence from Tanzania. **World Development**, v. 36, n. 11, p. 2223–2242, 2008.

KLUVE, J. The effectiveness of European active labor market programs. **Labour economics**, v. 17, n. 6, p. 904–918, 2010.

KRAFFT, C. Is school the best route to skills? Returns to vocational school and vocational skills in Egypt. **The Journal of Development Studies**, v. 54, n. 7, p. 1100–1120, 2018.

LALONDE, R. J. Evaluating the econometric evaluations of training programs with experimental data. **The American economic review**, p. 604–620, 1986.

LARSON, K. **The Internet and political participation: the effect of Internet use on voter turnout**. {PhD} {Thesis}—[s.l: s.n.].

LAZER, D. M. et al. The science of fake news. **Science**, v. 359, n. 6380, p. 1094–1096, 2018.

LECHNER, M.; MIQUEL, R.; WUNSCH, C. The curse and blessing of training the unemployed in a changing economy: The case of East Germany after unification. **German Economic Review**, v. 8, n. 4, p. 468–509, 2007.

LELKES, Y.; SOOD, G.; IYENGAR, S. The hostile audience: The effect of access to broadband internet on partisan affect. **American Journal of Political Science**, v. 61, n. 1, p. 5–20, 2017.

LEON, F. L. L. DE; MENEZES-FILHO, N. A. Reprovação, avanço e evasão escolar no Brasil. **Pesquisa e Planejamento Econômico (PPE) - Artigos**, 2002.

LEON, L. DE; LEITE, F.; RIZZI, R. A test for the rational ignorance hypothesis: Evidence from a natural experiment in Brazil. **American Economic Journal: Economic Policy**, v. 6, n. 4, p. 380–98, 2014.

LIJPHART, A. Unequal participation: Democracy's unresolved dilemma presidential address, American Political Science Association, 1996. **American political science review**, v. 91, n. 1, p. 1–14, 1997.

MADDUX, C. D.; JOHNSON, D. L. The World Wide Web: History, cultural context, and a manual for developers of educational information-based Web sites. **Educational Technology**, v. 37, n. 5, p. 5–12, 1997.

MCCALL, B.; SMITH, J.; WUNSCH, C. Government-sponsored vocational education for adults. In: **Handbook of the Economics of Education**. [s.l.] Elsevier, 2016. v. 5p. 479–652.

MCCRARY, J. Manipulation of the running variable in the regression discontinuity design: A density test. **Journal of econometrics**, v. 142, n. 2, p. 698–714, 2008.

MEER, J. Evidence on the returns to secondary vocational education. **Economics of education review**, v. 26, n. 5, p. 559–573, 2007.

MENEZES, A. Internet Availability, Political Information, and Voting: Evidence from Brazil. 2015.

MENEZES-FILHO, N. Avaliação econômica de projetos sociais. In: MENEZES-FILHO, N. (Ed.). **Avaliação econômica de projetos sociais**. 1. ed. São Paulo: Dinâmica Gráfica e Editora Ltda, 2012. p. 9–11.

MINARD, P.; LANDRIAULT, M. Internet Use, Voter Turnout, and Party Preference across Four Regime Types in East Asia. **Taiwan Journal of Democracy**, v. 11, n. 2, p. 75–108, 2015.

MINCER, J. Investment in human capital and personal income distribution. **The journal of political economy**, p. 281–302, 1958.

MINCER, J. A. Schooling and earnings. In: **Schooling, experience, and earnings**. [s.l.] NBER, 1974. p. 41–63.

MINER, L. The unintended consequences of Internet diffusion: Evidence from Malaysia. **Journal of Public Economics**, v. 132, p. 66–78, 2015.

MONASTERIO, L. **O FPM ea estranha distribuição da população dos pequenos municípios brasileiros**. [s.l.] Texto para Discussão, 2013.

NERI, M. Motivos da evasão escolar. **Pesquisa Todos pela educação**, 2015.

NETO, L. G. Desigualdades e políticas regionais no Brasil: caminhos e descaminhos. **Planejamento e políticas públicas**, n. 15, 2009.

O'CONNELL, S. D. et al. Can Business Input Improve the Effectiveness of Worker Training? **WB: Policy Research Working Papers**, 2017.

OBERHOLZER-GEE, F.; WALDFOGEL, J. Media markets and localism: Does local news en Espanol boost Hispanic voter turnout? **American Economic Review**, v. 99, n. 5, p. 2120–28, 2009.

OLIVEIRA, A. M. H. C. DE; RIOS-NETO, E. L. G. Uma avaliação experimental dos impactos da política de qualificação profissional no Brasil. **Revista Brasileira de Economia**, v. 61, n. 3, p. 353–378, 2007.

PETTERSSON-LIDBOM, P. Do parties matter for economic outcomes? A regression-discontinuity approach. **Journal of the European Economic Association**, v. 6, n. 5, p. 1037–1056, 2008.

POWER, T. J.; ZUCCO JR, C. Elite preferences in a consolidating democracy: the Brazilian legislative surveys, 1990–2009. **Latin American Politics and Society**, v. 54, n. 4, p. 1–27, 2012.

POY, S.; SCHÜLLER, S. Internet and voting in the Web 2.0 era: evidence from a local broadband policy. **Working Paper**, 2016.

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

QUINTANA, R. E.; CRAVO, T. **Demand-driven training and job turnover: The effects of Brazil's Pronate-MDIC at firm and worker level**. [s.l.] IDB Working Paper Series, 2019. Disponível em: <https://www.econstor.eu/handle/10419/208180>.

REIS, M. Vocational training and labor market outcomes in brazil. **The BE Journal of Economic Analysis & Policy**, v. 15, n. 1, p. 377–405, 2015.

RICHARDSON, K.; VAN DEN BERG, G. J. The effect of vocational employment training on the individual transition rate from unemployment to work. **IFAU-Institute for Labour Market Policy Evaluation, Working Paper**, n. 2002, p. 8, 2002.

RIKER, W. H.; ORDESHOOK, P. C. A Theory of the Calculus of Voting. **American political science review**, v. 62, n. 1, p. 25–42, 1968.

ROSHOLM, M. Experimental evidence on the nature of the Danish employment miracle. n. Discussion Paper, 2008.

ROSHOLM, M.; SKIPPER, L. Is labour market training a curse for the unemployed? Evidence from a social experiment. **Journal of Applied Econometrics**, v. 24, n. 2, p. 338–365, 2009.

RUBIN, D. B. Estimating causal effects of treatments in randomized and nonrandomized studies. **Journal of educational Psychology**, v. 66, n. 5, p. 688, 1974.

SAKELLARIOU, C. Rates of return to investments in formal and technical/vocational education in Singapore. **Education Economics**, v. 11, n. 1, p. 73–87, 2003.

SCHULTZ, T. W. Investment in human capital. **The American economic review**, p. 1–17, 1961.

SCHWERDT, G. et al. The impact of an adult education voucher program: Evidence from a randomized field experiment. **Journal of Public Economics**, v. 96, n. 7-8, p. 569–583, 2012.

SENADO. **Avaliação de políticas públicas: programa nacional de acesso ao ensino técnico e emprego (PRONATEC)**Relatório nº 6 , de 2017, 2017. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento?dm=7319702&disposition=inline>

SILVA, P. L. N.; PESSOA, D. G. C.; LILA, M. F. Análise estatística de dados da PNAD: incorporando a estrutura do plano amostral. **Ciência & Saúde Coletiva**, v. 7, n. 4, p. 659–670, 2002.

SILVA, S. P. **Financiamento das políticas públicas de trabalho e renda no Brasil: uma análise a partir da trajetória operacional do Fundo de Amparo ao Trabalhador (FAT)**. [s.l.] Texto para Discussão, 2018.

SOUZA, C. Federalismo, desenho contitucional e instituições federativas no Brasil pós-1988. **Revista de sociologia e política**, n. 24, p. 105–121, 2005.

SOUZA, P.; MEDEIROS, M. Diferencial Salarial Público-Privado e Desigualdade de Renda

per capita no Brasil. v. 43, p. 5–28, 2013.

SPENCE, M. Job market signaling. In: **Uncertainty in economics**. [s.l.] Elsevier, 1978. p. 281–306.

STEVENS, A. H.; KURLAENDER, M.; GROSZ, M. Career Technical Education and Labor Market Outcomes Evidence from California Community Colleges. **Journal of Human Resources**, v. 54, n. 4, p. 986–1036, 2019.

STRÖMBERG, D. Radio's impact on public spending. **The Quarterly Journal of Economics**, v. 119, n. 1, p. 189–221, 2004.

SWAMY, P.; ARORA, S. S. The exact finite sample properties of the estimators of coefficients in the error components regression models. **Econometrica: Journal of the Econometric Society**, p. 261–275, 1972.

TCU. **Relatório: auditoria operacional. Bolsa formação. Programa Nacional de Acesso ao Ensino Técnico e Emprego - Pronatec. Avaliação da impantação e da efetividade. Recomendações. Determinações.**, a2015.

TCU. **Relatório: levantamento. Programa Nacional de Acesso ao Ensino Técnico e Emprego – PRONATEC. Determinação à SEGECEX. Arquivamento.**, b2015.

TOLBERT, C. J.; MCNEAL, R. S. Unraveling the effects of the Internet on political participation? **Political research quarterly**, v. 56, n. 2, p. 175–185, 2003.

TRAIN, K. E. **Discrete choice methods with simulation**. [s.l.] Cambridge university press, 2009.

TRAVASSOS, C.; VIACAVA, F.; LAGUARDIA, J. Os suplementos saúde na Pesquisa Nacional por Amostra de Domicílios (PNAD) no Brasil. **Revista Brasileira de Epidemiologia**, v. 11, p. 98–112, 2008.

UHLANER, C. J. Rational turnout: The neglected role of groups. **American Journal of Political Science**, p. 390–422, 1989.

VASCONCELLOS, L.; LIMA, F. C.; MENEZES-FILHO, N. Avaliação Econômica do Ensino Médio Profissional. **Fundação Itaú Social, Brasilia**, 2010.

VOOREN, M. et al. The effectiveness of active labor market policies: a meta-analysis. **Journal of Economic Surveys**, v. 33, n. 1, p. 125–149, 2019.

WOOLDRIDGE, J. M. **Econometric analysis of cross section and panel data**. [s.l.] MIT press, 2010.

Appendix

Table 61: First stage of Fuzzy-RDD for all outcomes

Year	Cutoff	Outcome	(Intercept)	Tr	Xl	Xr	F.stat
2008	20,000	Blank or Null votes	-4.407	-1.414	0.001**	0.001	13.09***
		Left wing vote share	3.071	-2.095	0.001*	0.000	5.83***
		N. cand.	-4.140	-1.380	0.001**	0.001	13.25***
		PSOL budget	-17.673	-2.821	0.000	0.000	1.35
		PSOL vote share	-10.418	-4.467	0.000	0.001**	2.42***
		Turnout	-8.690	-1.710	0.002**	0.001	11.10***
		Young budget	-3.631	-1.049	0.001***	0.001*	14.47***
		Young votes	-3.965	-1.131	0.001***	0.001*	13.96***
2008	40,000	Blank or Null votes	16.382	4.026	0.000*	0.000	10.70***
		Left wing vote share	15.618	-0.761	0.001*	0.000	7.45***
		N. cand.	16.405	4.025	0.000*	0.000	10.70***
		PSOL budget	-242.491*	24.409*	-0.001	0.001	3.51**
		PSOL vote share	-75.356	-9.385	0.001	0.002***	5.76***
		Turnout	17.322	3.934	0.001*	0.000	10.03***
		Young budget	4.718	5.284**	0.000	0.000	18.23***
		Young votes	9.598	5.335**	0.000*	0.000	13.65***
2008	60,000	Blank or Null votes	142.193**	16.646**	0.000	-0.001	9.43***
		Left wing vote share	103.102	20.470**	0.001	-0.001*	7.92***
		N. cand.	89.264**	15.927***	0.000	-0.001	16.31***
		PSOL budget	171.406	-34.691	0.002	-0.001	1.92
		PSOL vote share	-27.075	23.145**	0.001	-0.001**	3.68***
		Turnout	171.414**	17.917**	0.000	-0.001	8.34***
		Young budget	244.117**	25.441**	-0.001	-0.003	4.72***
		Young votes	249.056**	27.166**	-0.002	-0.003	4.74***
2008	Pooled	Blank or Null votes	1.800	3.883***	0.000	0.000	59.68***
		Left wing vote share	3.057	3.004**	0.000	0.000**	27.44***
		N. cand.	20.586***	3.626***	0.000***	0.000*	59.84***
		PSOL budget	-48.025	11.061*	-0.001	-0.001	2.95***
		PSOL vote share	7.967	13.170**	-0.001	-0.001	3.73***
		Turnout	34.252***	1.207	0.000	0.000*	38.57***
		Young budget	1.751	3.356***	0.000**	0.000**	56.94***
		Young votes	1.846	3.214***	0.000**	0.000**	57.08***
2010	20,000	Blank or Null votes	-7.007	2.495***	0.001***	0.001***	23.79***
		Left wing vote share	-6.890	2.892***	0.001***	0.001***	29.57***
		PSOL vote share	-8.461**	3.591***	0.001***	0.001***	46.85***
		Turnout	-6.873	2.970***	0.001***	0.001***	30.33***
		Young votes	-7.555	1.891**	0.002***	0.001**	17.33***

Table 61: First stage of Fuzzy-RDD for all outcomes *(continued)*

Year	Cutoff	Outcome	(Intercept)	Tr	Xl	Xr	F.stat
2010	40,000	Blank or Null votes	-71.169***	3.481**	0.001***	0.001**	15.22***
		Left wing vote share	-82.716***	1.954	0.001**	0.002***	11.55***
		PSOL vote share	-59.078***	5.131***	0.001***	0.001**	21.05***
		Turnout	-63.840***	4.361**	0.001***	0.001**	17.97***
		Young votes	-65.281***	4.164**	0.001***	0.001**	16.65***
2010	60,000	Blank or Null votes	-68.369	19.281***	0.001**	0.000	10.36***
		Left wing vote share	-55.390	18.911***	0.001**	0.000	12.87***
		PSOL vote share	-57.071	18.786***	0.001**	0.000	11.99***
		Turnout	-48.205	18.167***	0.001***	0.000	14.43***
		Young votes	-128.045**	6.629	0.010***	-0.003	8.07***
2010	Pooled	Blank or Null votes	22.296***	0.938	0.000	0.000	52.48***
		Left wing vote share	18.773***	1.647**	0.000	0.000	65.08***
		PSOL vote share	-0.946	3.991***	0.000***	0.001***	60.34***
		Turnout	19.819***	1.500**	0.000	0.000	59.57***
		Young votes	-9.476**	3.015***	0.000**	0.001***	67.44***
2012	20,000	Blank or Null votes	-5.575	1.615**	0.002***	0.001**	21.47***
		Left wing vote share	-8.703	0.960	0.002***	0.001**	10.85***
		N. cand.	-4.608**	4.804***	0.000***	0.000***	117.72***
		PSOL budget	-48.958**	2.497	0.001**	0.001*	6.44***
		PSOL vote share	-22.785*	3.957**	0.001**	0.000	8.29***
		Turnout	-5.077**	4.444***	0.000***	0.000***	91.36***
		Young budget	-5.402	2.168***	0.001***	0.001**	31.88***
		Young votes	-5.876	1.793**	0.001***	0.001**	26.52***
2012	40,000	Blank or Null votes	-44.942***	5.117***	0.001***	0.000**	25.09***
		Left wing vote share	-40.295***	3.609**	0.001***	0.001**	15.00***
		N. cand.	-60.061***	0.091	0.002***	0.001*	9.24***
		PSOL budget	-84.435**	5.518	0.001**	0.000	7.08***
		PSOL vote share	-40.603	5.733*	0.001**	0.000	6.59***
		Turnout	-60.151***	0.070	0.002***	0.001*	9.24***
		Young budget	-60.978**	-1.447	0.003***	0.001	7.73***
		Young votes	-64.062**	-3.238	0.004***	0.001	6.96***

Table 61: First stage of Fuzzy-RDD for all outcomes *(continued)*

Year	Cutoff	Outcome	(Intercept)	Tr	Xl	Xr	F.stat
2012	60,000	Blank or Null votes	-79.639	-2.756	0.002**	0.003**	7.81***
		Left wing vote share	-82.973	1.088	0.002**	0.002**	5.81***
		N. cand.	-77.764	-7.264	0.002**	0.004**	6.23***
		PSOL budget	-135.674*	11.076*	0.001	0.002**	11.04***
		PSOL vote share	-174.023*	11.344	0.001	0.001	6.41***
		Turnout	-78.936	-6.264	0.002**	0.004**	6.79***
		Young budget	-33.465*	15.215***	0.001***	0.000	32.02***
		Young votes	-78.819	-4.944	0.002**	0.004**	7.44***
2012	Pooled	Blank or Null votes	-5.428**	3.752***	0.000***	0.001***	78.94***
		Left wing vote share	17.203**	2.108**	0.000***	0.001***	46.51***
		N. cand.	20.621***	1.988**	0.000**	0.000**	70.85***
		PSOL budget	-38.712	5.558	-0.001	0.001***	6.71***
		PSOL vote share	-7.134	1.610	0.000	0.002***	8.41***
		Turnout	31.327***	0.075	0.000	0.000**	46.55***
		Young budget	-6.665**	2.338***	0.000***	0.001***	69.19***
		Young votes	-6.475**	2.268***	0.000***	0.001***	70.24***

Obs.: Standard Errors (SE) are clustered by regions. Optimal bandwidth (Bw) selection by Mean Square Error following Calonico, Cattaneo and Titiunik (2014).

Triangular kernel with quadratic local-polynomial.

Turnout for the first round.

Results with controls listed in Table 10

Figure 14: Datasets used in analysis: schematic view

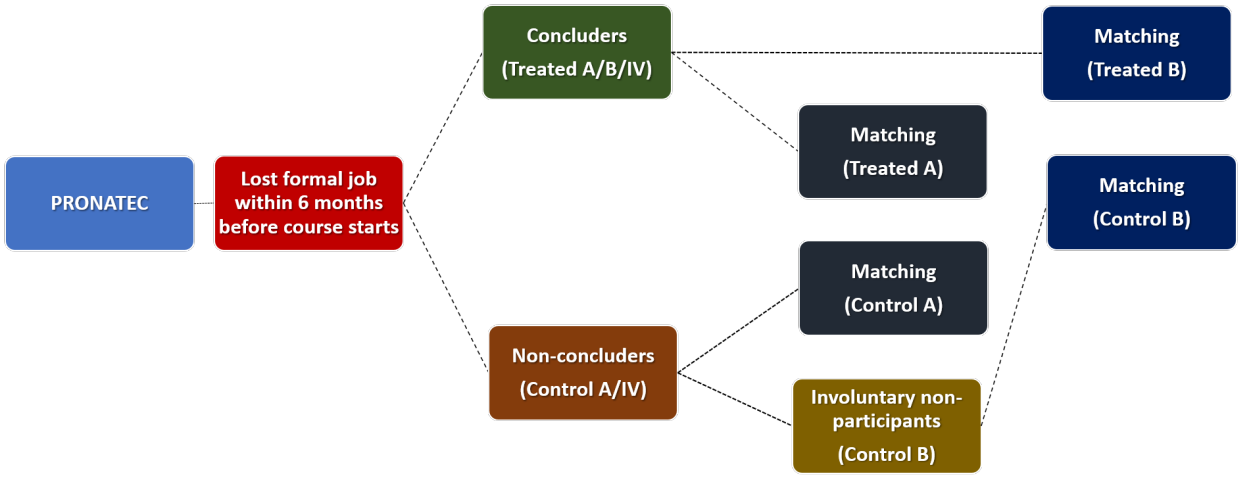


Table 62: Pronatec program – application by demandant, 2011-2015

Demandant	N	%
Ministry of Social Development (MDS)	1,944,207	49.55
On-line subscription	820,768	20.92
State and FD Education Secretaries	472,113	12.03
Ministry of Labour	393,559	10.03
Other Ministries	154,480	3.94
Ministry of Tourism	89,287	2.28
Ministry of Development, Industry and Foreign Trade (MDIC)	49,552	1.26
Total	3,923,966	100.00

Source: Sistec

Obs.: Other ministries are Justice, Defense, Agrarian Reform, Culture, Fishing and Aquaculture, Social Security, Communication, Environment, Human Rights.

Table 63: Pronatec program - application by axis, 2011-2015

Axis	N	%
Business and Management	1,025,553	26.14
Industrial Control and Processes	542,256	13.82
Information and Communication	413,051	10.53
Infrastructure	399,901	10.19
Environment and Health	361,123	9.20
Tourism, Hospitality and Leisure	318,067	8.11
Social and Education Development	236,054	6.02
Natural Resources	204,107	5.20
Industrial Production	194,345	4.95
Cultural Production and Design	125,941	3.21
Food Production	90,183	2.30
Security	13,385	0.34
Total	3,923,966	100.00

Source: Sistec

Figure 15: PSM before matching – Subset A

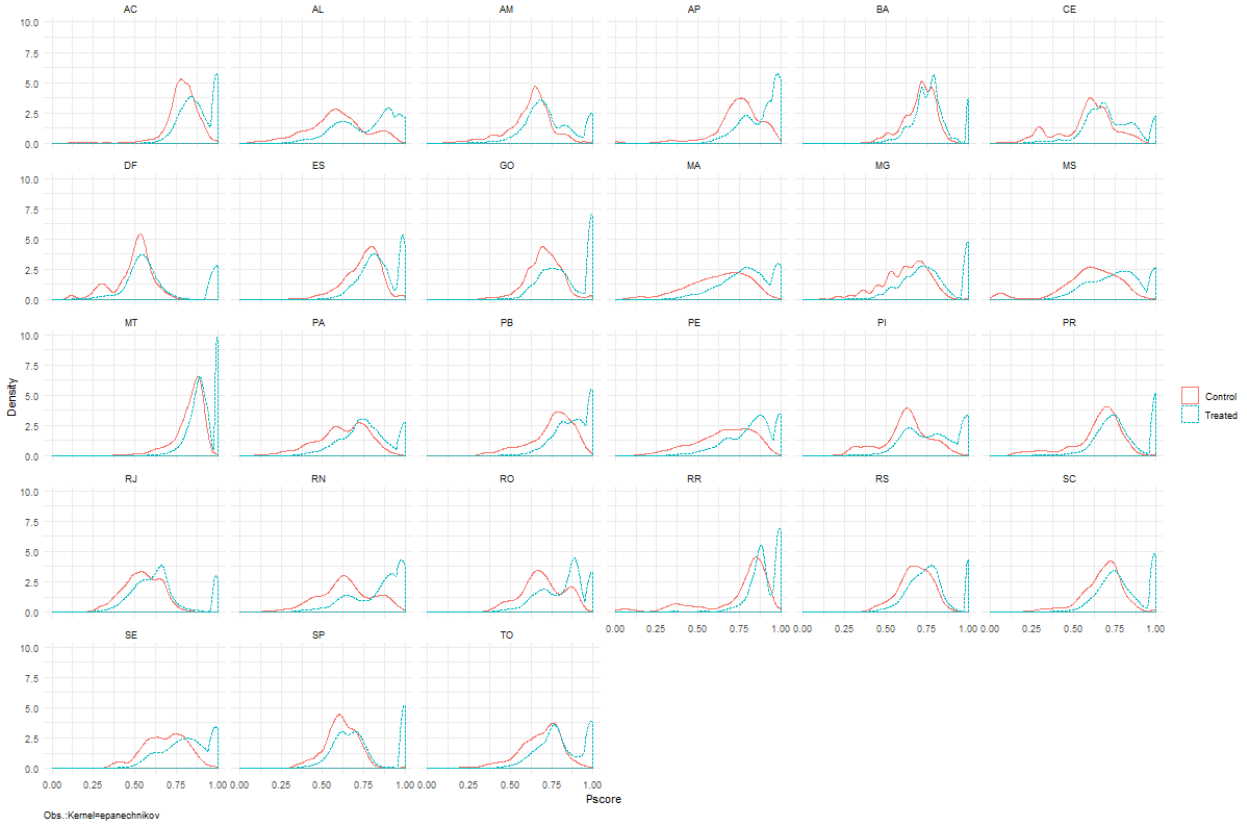


Figure 16: PSM after matching – Subset A

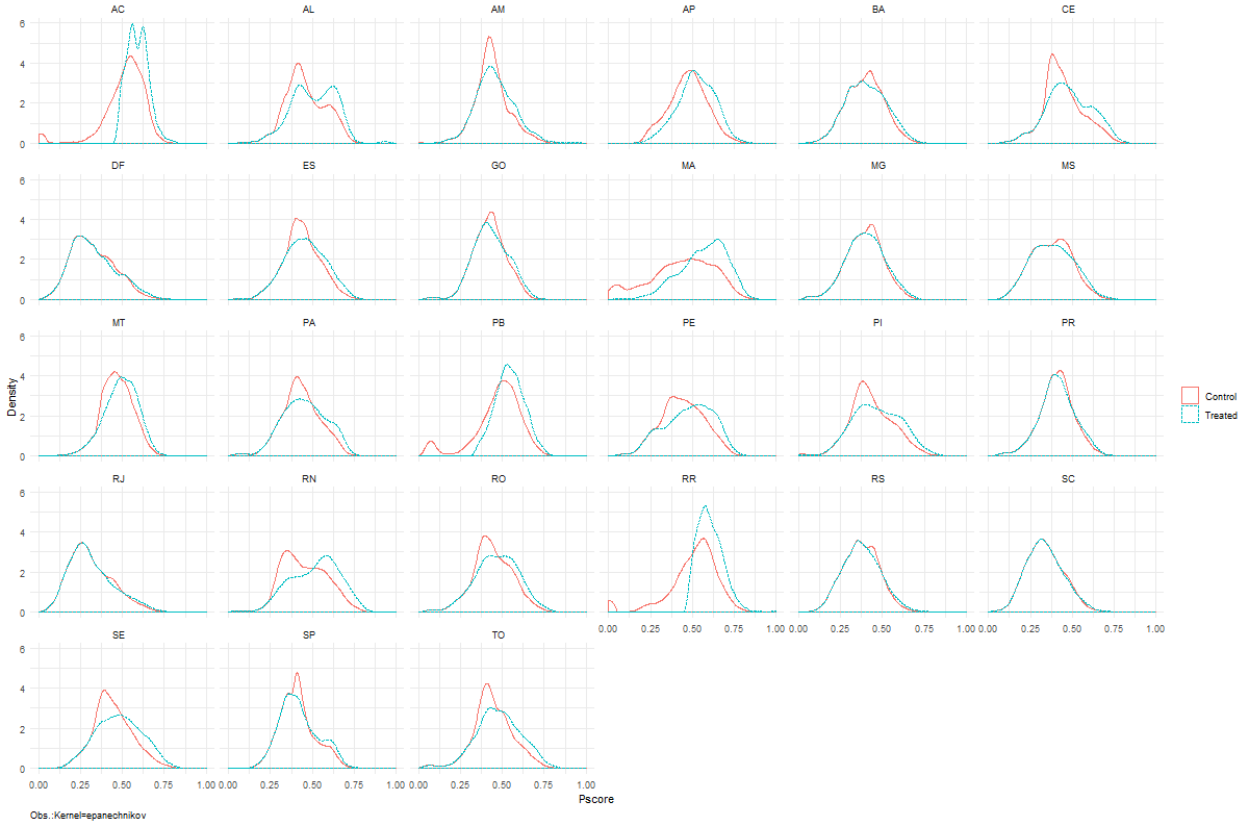


Figure 17: PSM before matching – Subset B

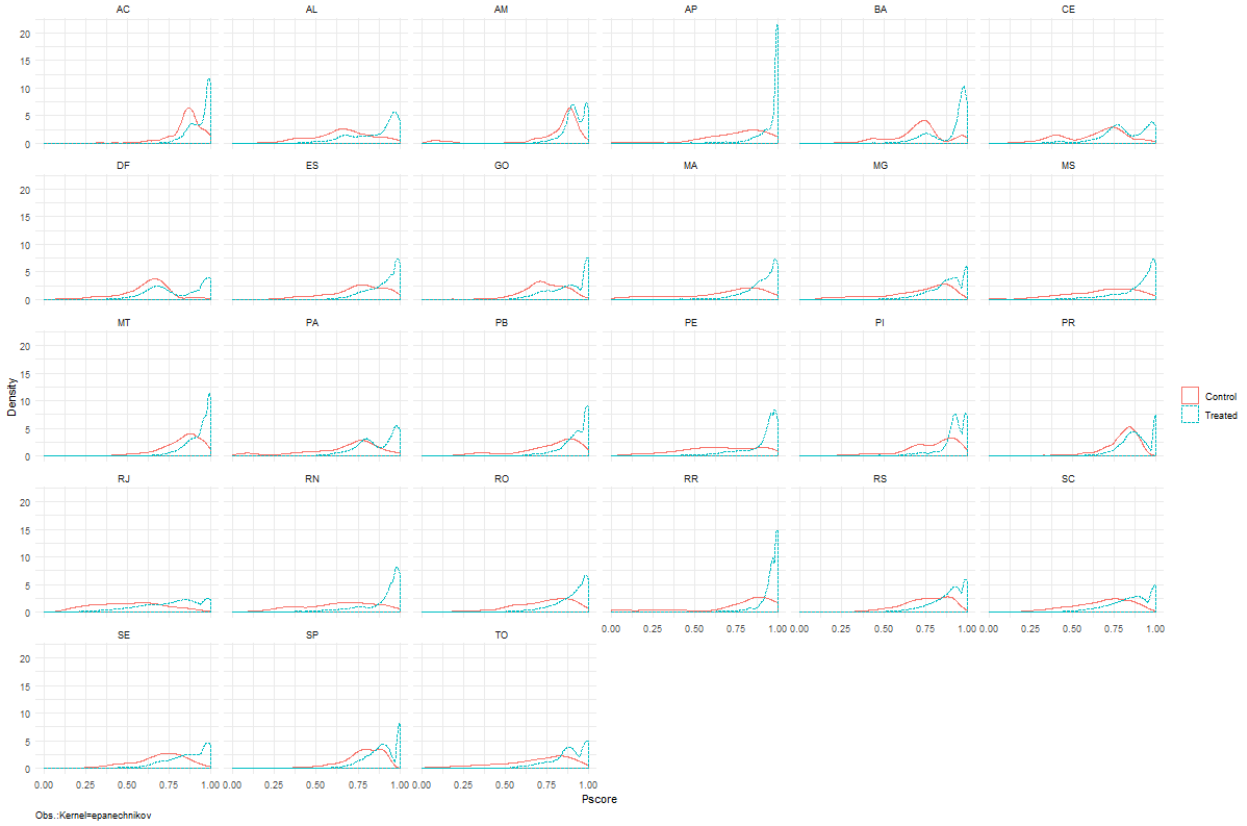


Figure 18: PSM after matching – Subset B

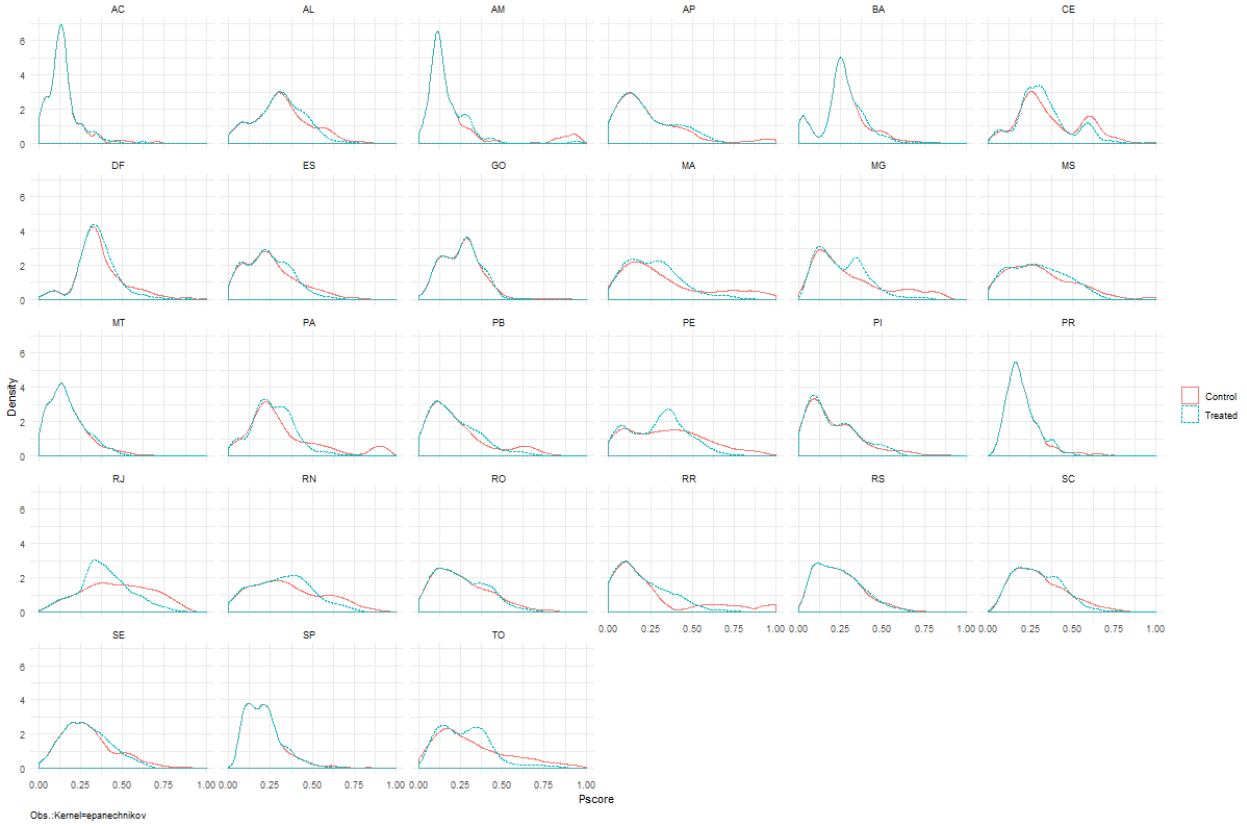


Table 64: Weak instruments and Wu-Hausman tests for employment IV regression

test	df1	df2	statistic	p.value	time
Weak instruments (Pronatec)	2	640,201	11,358.48	0.000	
Weak instruments (Proantec.share)	2	640,201	1,134.83	0.000	6 months
Wu-Hausman	2	640,199	27.86	0.000	
Weak instruments (Pronatec)	2	640,201	11,358.48	0.000	
Weak instruments (Proantec.share)	2	640,201	1,134.83	0.000	12 months
Wu-Hausman	2	640,199	4.96	0.007	
Weak instruments (Pronatec)	2	640,201	11,358.48	0.000	
Weak instruments (Proantec.share)	2	640,201	1,134.83	0.000	18 months
Wu-Hausman	2	640,199	4.87	0.008	
Weak instruments (Pronatec)	2	640,201	11,358.48	0.000	
Weak instruments (Proantec.share)	2	640,201	1,134.83	0.000	24 months
Wu-Hausman	2	640,199	4.05	0.017	
Weak instruments (Pronatec)	2	640,201	11,358.48	0.000	
Weak instruments (Proantec.share)	2	640,201	1,134.83	0.000	36 months
Wu-Hausman	2	640,199	0.16	0.852	
Weak instruments (Pronatec)	2	622,157	11,240.40	0.000	
Weak instruments (Proantec.share)	2	622,157	1,210.62	0.000	

					48 months
Wu-Hausman	2	622,155	0.39	0.676	
Weak instruments (Pronatec)	2	566,992	10,520.02	0.000	
Weak instruments (Proantec.share)	2	566,992	1,229.38	0.000	60 months
Wu-Hausman	2	566,990	0.51	0.600	
Weak instruments (Pronatec)	2	272,689	3,807.12	0.000	
Weak instruments (Proantec.share)	2	272,689	531.45	0.000	72 months
Wu-Hausman	2	272,687	0.75	0.471	
Weak instruments (Pronatec)	2	26,230	225.39	0.000	
Weak instruments (Proantec.share)	2	26,230	1.64	0.194	84 months
Wu-Hausman	2	26,228	1.14	0.320	

Source: Sistec

Obs.: df = Degrees of freedom.

Table 65: Weak instruments and Wu-Hausman tests for wages IV regression

test	df1	df2	statistic	p.value	time
Weak instruments (Pronatec)	2	640,201	11,358.48	0.000	
Weak instruments (Proantec.share)	2	640,201	1,134.83	0.000	6 months
Wu-Hausman	2	640,199	27.36	0.000	
Weak instruments (Pronatec)	2	640,201	11,358.48	0.000	
Weak instruments (Proantec.share)	2	640,201	1,134.83	0.000	12 months
Wu-Hausman	2	640,199	6.36	0.002	
Weak instruments (Pronatec)	2	640,201	11,358.48	0.000	
Weak instruments (Proantec.share)	2	640,201	1,134.83	0.000	18 months
Wu-Hausman	2	640,199	6.36	0.002	
Weak instruments (Pronatec)	2	640,201	11,358.48	0.000	
Weak instruments (Proantec.share)	2	640,201	1,134.83	0.000	24 months
Wu-Hausman	2	640,199	3.76	0.023	
Weak instruments (Pronatec)	2	640,201	11,358.48	0.000	
Weak instruments (Proantec.share)	2	640,201	1,134.83	0.000	36 months
Wu-Hausman	2	640,199	0.58	0.559	
Weak instruments (Pronatec)	2	622,157	11,240.40	0.000	
Weak instruments (Proantec.share)	2	622,157	1,210.62	0.000	48 months
Wu-Hausman	2	622,155	0.48	0.619	
Weak instruments (Pronatec)	2	566,992	10,520.02	0.000	
Weak instruments (Proantec.share)	2	566,992	1,229.38	0.000	60 months
Wu-Hausman	2	566,990	1.32	0.266	
Weak instruments (Pronatec)	2	272,689	3,807.12	0.000	
Weak instruments (Proantec.share)	2	272,689	531.45	0.000	72 months
Wu-Hausman	2	272,687	1.62	0.198	
Weak instruments (Pronatec)	2	26,230	225.39	0.000	
Weak instruments (Proantec.share)	2	26,230	1.64	0.194	84 months
Wu-Hausman	2	26,228	1.02	0.362	

Source: Sistec

Obs.: df = Degrees of freedom.

Table 66: IV regression – first stage for Pronatec conclusion endogenous variable

term	6 months	12 months	18 months	24 months	36 months	48 months	60 months	72 months	84 months
Z1	-0.365 (0.001)***	-0.365 (0.001)***	-0.365 (0.001)***	-0.365 (0.001)***	-0.365 (0.001)***	-0.369 (0.001)***	-0.371 (0.001)***	-0.390 (0.002)***	-0.300 (0.009)***
Z2	0.007 (0.006)	0.007 (0.006)	0.007 (0.006)	0.007 (0.006)	0.007 (0.006)	-0.003 (0.006)	0.006 (0.006)	0.177 (0.010)***	0.180 (0.019)***
N. obs.	640,270	640,270	640,270	640,270	640,270	622,225	567,059	272,755	26,294

Source: Sistec

Table 67: IV regression – first stage for Pronatec share endogenous variable

term	6 months	12 months	18 months	24 months	36 months	48 months	60 months	72 months	84 months
Z1	0.001 (0.000)***	0.001 (0.000)***	0.001 (0.000)***	0.001 (0.000)***	0.001 (0.000)***	0.001 (0.000)***	0.001 (0.000)***	0.000 (0.000)	0.001 (0.001)
Z2	-0.050 (0.001)***	-0.050 (0.001)***	-0.050 (0.001)***	-0.050 (0.001)***	-0.050 (0.001)***	-0.052 (0.001)***	-0.054 (0.001)***	-0.046 (0.002)***	0.002 (0.001)
N. obs.	640,270	640,270	640,270	640,270	640,270	622,225	567,059	272,755	26,294

Source: Sistec

Figure 19: Descriptive for all workers by CBO, 2007 to 2018

Figure 20: Descriptive for all workers by Axis, 2007 to 2018

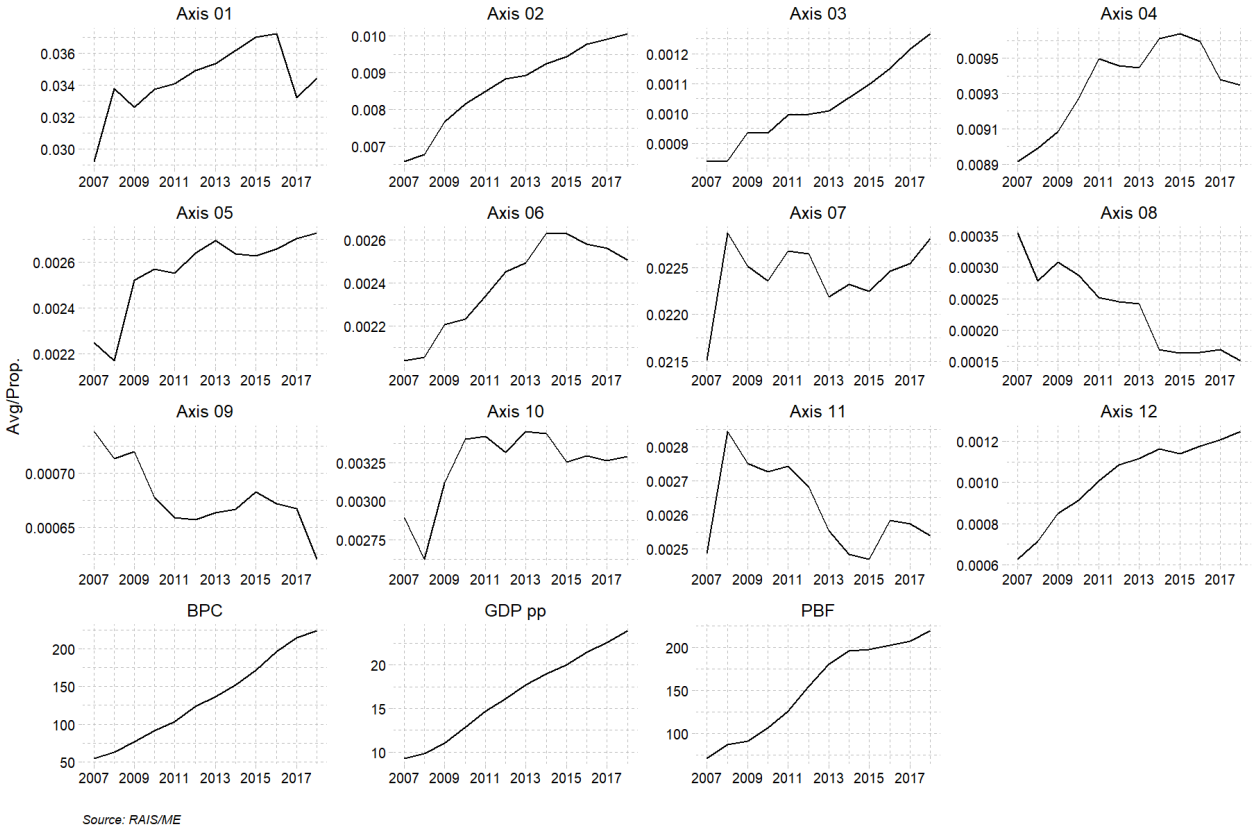


Figure 21: Descriptive for all workers by CNAE, 2007 to 2018

Figure 22: Descriptive for all workers for matching database, 2007 to 2018

Figure 23: Descriptive for young workers for matching database, 2007 to 2018

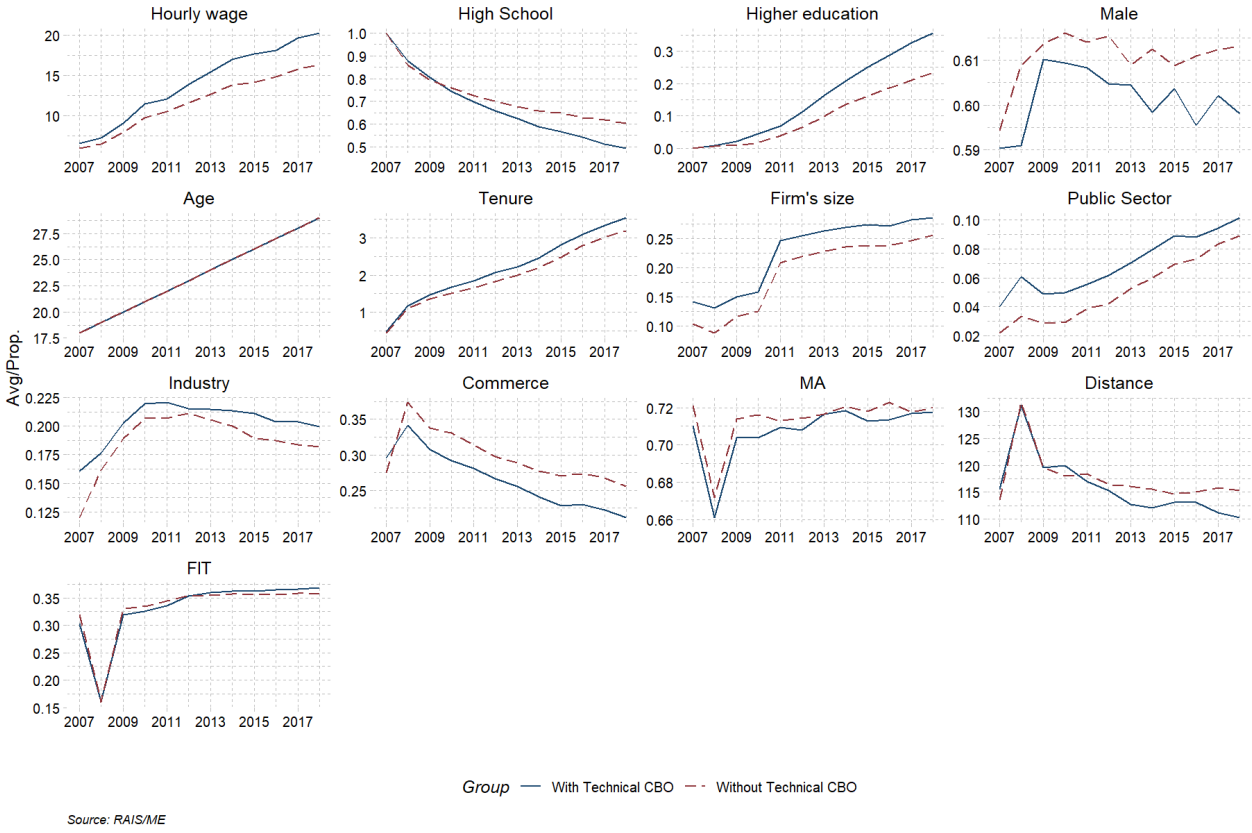


Table 68: Descriptive statistics and mean difference test for matching data (all workers), 2007 and 2014

	2007			2014		
Variable	Tech	No Tech	P.value	Tech	No Tech	P.value
Formal Job (%)	40.8	27.0	0.000	41.2	31.1	0.000
Married Couple (%)	74.5	72.9	0.001	72.8	71.5	0.007
Commerce (%)	13.7	13.1	0.064	12.7	13.0	0.506
High School (%)	65.3	31.1	0.000	60.8	35.0	0.000
Higher Education (%)	19.8	11.2	0.000	24.4	15.9	0.000
Children over 14	1.4	1.6	0.000	1.1	1.2	0.000
Children under 14	1.3	1.5	0.000	0.9	1.0	0.000
Student (%)	14.3	11.2	0.000	11.6	8.9	0.000
Public Sector (%)	13.1	5.3	0.000	10.6	5.3	0.000
Age	36.6	40.4	0.000	39.5	42.4	0.000
Industry (%)	14.2	10.4	0.000	12.6	9.0	0.000
Male (%)	48.0	46.7	0.023	50.4	46.7	0.000
Black (%)	37.2	43.9	0.000	42.1	48.0	0.000
Ref. Person (%)	41.9	40.7	0.029	46.1	42.7	0.000
Household Income (WPI)	2,065.0	1,706.8	0.000	3,303.4	2,843.7	0.000

MR (%)	39.7	36.4	0.000	40.2	39.0	0.107
Hourly wage	39.8	27.5	0.000	120.1	87.6	0.000
Tenure (%)	692.0	746.8	0.000	767.7	770.0	0.837
N# household members	3.6	3.8	0.000	3.3	3.4	0.000
Urban (%)	96.6	91.5	0.000	96.5	93.2	0.000

Source: PNAD/IBGE

Table 69: Descriptive statistics and mean difference test for matching data (18 young), 2007 and 2014

	2007			2014		
Variable	Tech	No Tech	P.value	Tech	No Tech	P.value
Formal Job (%)	38.1	16.5	0.009	29.5	26.3	0.480
Married Couple (%)	70.7	70.3	0.968	75.4	74.4	0.809
Commerce (%)	11.1	12.5	0.766	14.0	14.3	0.919
High School (%)	97.4	31.2	0.000	90.9	84.2	0.019
Higher Education (%)	0.0	0.0		0.0	0.0	
Children over 14	2.1	2.2	0.707	1.9	1.9	0.627
Children under 14	1.1	1.6	0.068	1.0	1.0	0.577
Student (%)	20.3	53.2	0.000	31.7	27.6	0.361
Public Sector (%)	0.0	0.2	0.044	0.2	0.0	0.318
Age	18.0	18.0	0.961	18.0	18.0	0.665
Industry (%)	8.4	8.6	0.954	6.4	9.8	0.208
Male (%)	59.8	51.6	0.305	48.3	52.4	0.438
Black (%)	42.8	48.4	0.485	47.4	50.9	0.471
Ref. Person (%)	3.0	2.7	0.918	2.1	2.9	0.596
Household Income (WPI)	1,958.3	1,811.2	0.621	3,273.3	3,087.0	0.493
MR (%)	32.4	34.8	0.742	33.1	35.8	0.543
Hourly wage	13.2	9.5	0.020	32.9	38.3	0.263
Tenure (%)	91.3	143.5	0.154	58.2	107.4	0.007
N# household members	4.3	4.6	0.215	4.1	4.1	0.975
Urban (%)	97.9	86.4	0.000	92.2	89.6	0.318

Source: PNAD/IBGE