Seriously, less taxes? Tax Reform's Effects over Labor Supply.

Juan Esteban Isaza Cadavid*, Juan Guillermo Salazar Duque†
School of Economics and Finance
Universidad EAFIT
May 2015

May 27, 2015

Abstract

Using a matching and a difference in differences approach we analyzed the impact of a payroll costs reduction implemented through a tax reform in Colombia on worked hours and hourly income. Existing studies have found mixed results, while increases in wages are commonly found, the results for employment and worked hours effects are a source of debate with no consensus on sight.
However, in line with earlier revisions of literature, we concluded that the effect of payroll costs reductions is positive and significant on both, worked hours and income. In this paper, using socioeconomic data from Colombia we found that those individuals affected by a payroll cost reduction between 2012-2014 exhibited significant increases in both hours and income compared to similar individuals unaffected by this framework.

Key words: Labor supply, Difference in Difference, Matching.

JEL classification: $J21, C21$

*Undergraduate Economics Student, Universidad Eafit *jisazac@eafit.edu.co*
†Undergraduate Economics Student, Universidad Eafit *jsalaz29@eafit.edu.co*

1 Introduction

Governments will always face several challenges and harbor doubts concerning the effectiveness of any tax reform and its possible effects on the whole economy. Does it seem affected? What would be the aftermath for markets? Will income levels increase or there will be any other unexpected outcomes? These issues must be considered deeply, it is true that it is better to embark on these reforms required by the nation sooner than later as it is also important to objectively analyze the way these reforms played turn out against the goals initially considered when proposing them.

We will focus on the second chapter of Law 1607 of 2012, regarding and CREE tax. This tax replaces payroll taxes that were paid by companies for employees earning less than ten minimum wages. Previously, employers made payments on health insurance (8.5%), ICBF - Colombian Institute of Family Welfare - (3%) and - SENA National Education Service - (2%), now the CREE imposes a charge of 8% on companies' income. This paper develops an estimation model for working hours and wages taking into account those decisions made by citizens when they are directly affected by a tax reform, specifically a reduction in payroll costs, in terms of choosing to work more hours or not making any changes.

This paper is divided as follows: section 2 describes the Colombian institutional features of the payroll tax reform introduced in Law 1607, section 3 presents the review if literature, section 4 describes data and methods; our results are provided in section 5; section 6 introduces a discussion of methodological considerations and robustness of our estimates, and our conclusions are presented in section 7.

2 Background

During the last decades, Colombia has been one of the countries with more tax reforms Larrea and Uribe (2012) with 19 reforms implemented between 1990-2009. The 2012 Law is different from its predecessors, while the main objectives of all the previous ones were the reduction of the deficit, increase tax collection and stronger economic stability, the three objectives of the last one were to reduce inequality though better distribution of the income tax, reducing informality and improving job creation through a payroll tax reduction and diminishing tax evasion through anti-avoidance regulations Cárdenas (2012). This reduction amounts to 8% of total labor cost and 13.5% in payroll taxes Farné and Rodríguez (2013).

Congreso de la Repúplica (2012) Law 1607 includes 10 chapters and 187 articles supporting the creation of 3 new taxes, establishing changes in the value added tax and

new repercussions for tax evaders are established. Income tax for enterprises lowed from 33% to 25%, nevertheless in the 20 article of the bill CREE, was conceived with a rate of 9% for 2013, 2014 and 2015 from 2016 rate will be 8%. It substitutes payroll taxes that were made by enterprises. Finally income tax for individuals was transformed in IMAN and IMAS taxes. In Colombia payroll tax reductions have not been common. Laws 50 of 1990 and 789 of 2002 included labor reforms, diminishing costs of hiring paid staff, but shortly after its implementation these decreases were offset by increases in social security contributionsFarné and Rodríguez (2013).

Now, after the brief overview above we will move to a specific in-detail study case. In 2012 the Colombian government carried out an ambitious tax reform which argued that payroll taxes were a major obstacle for job creation. Hence, the 20th article of Law 1607 from December 2012 creates the CREE later implemented in January 2013 with the purpose of redesigning income taxes turning them into a contribution by taxpayers bringing benefits for employees, job creation and social investment. Through the CREE law, the Colombian Congress approved a reduction in income taxes suppressing payroll taxes, mainly social security, for employees making less than ten minimum wages in businesses with two or more employees.

The reform pushed by the government is expected to reduce inequality by 1.9 points in the GINI Index, increase formal employment by 11% and increase the relative size of the formal sector between 10% and 15%. Cárdenas (2012). These tax relieves were meant to make it easier and cheaper to hire workers, especially by the formal sector. On the other hand, to guarantee the sustainability and independence of social security agencies, the Colombian Congress approved the CREE tax, which represents a whole change given the way is applied as it taxes revenues instead of payroll. This new approach would take a stake of tax burden away from economic sectors acknowledged as employment creators, such as industry and agriculture, and put it on those sectors which relies heavily on capital such as mining and financial sector Gaviria (2004).

3 Literature Review

In 1995, the unemployment rate in Colombia reached 9.0%, before the 90's crisis, reaching levels close to 20% between 1999 and 2001. By 2007 unemployment had been substantially reduced standing at 13.5%. The subsequent reduction of unemployment to one digit matched the implementation of the tax reform analyzed in this paper.

Under a partial equilibrium model on the labor market Farné and Rodríguez (2013) show that the effects of a reduction in payroll taxes depend on five main factors: labor cost demand elasticity, labor cost supply elasticity, employees' valuation of taxes - which determines the amount employees are willing to see their payment reduced for services

covered by payroll taxes - , the existence of a minimum wage and the bargaining power of workers. Under these factors it is not obvious that reducing payroll taxes drives a higher employment level, for instance if supply were perfectly inelastic there would not be any effect on employment but wages would be affected.

On the other hand, Pissarides (1998) simulates a decline in taxes on labor under different labor market models such as a competitive market, a searching model, an efficiency wages model and a model of wage bargaining between firms and unions. Results differ among models. His conclusions point that if unemployment benefits are increased in proportion to the wage rate when taxes are cut real wages are likely to absorb tax changes but if unemployment benefits are not indexed to wages (and are held fixed in real terms) the employment effects of the tax cut can be sizeable. The results found under the assumptions of union wage bargaining and search models, point to another issue which is as important in policy designing. Proportional and progressive taxes under a monopoly have much less of an impact on employment than regressive taxes. In countries where there is regressive taxation (which usually takes the form of a fixed component in the tax levied on employers, or of a ceiling on social security contributions), a revenue-neutral reform of the tax system to a proportional or progressive tax can have larger employment effects than an across-the-board reduction in the marginal tax rate.

Bennmarker et al. (2009) carry out an impact evaluation using a difference-in-differences approach, on a 10% payroll tax reduction introduced in 2002 for enterprises in Northern Sweden. Effects on employment are not found, however, they recorded a positive impact on the average wage per employee. While Betcherman and Pagés (2007) conducted a series of empirical tests for determining the impact of a reform that reduced social insurance contributions in Turkey; results suggest a null effect on employment generation.

Econpubblica (2011) develops a research on the impact of a tax reform package in 31 European countries, the United States and Japan in 1990-2008. The analysis focuses on payroll tax reduction reforms. Significant results on employment are not observed; their policy recommendations suggest that tax policy plays a weaker role on the labor market than wage bargaining arrangements, monetary and in-kind transfers, job placement services, training programs, and support to geographical mobility. And in order to have some impact on the labor market, tax policies should not be across-the-board but targeted on particular groups of workers.

Gruber (1995) studied the effect of a 25% payroll tax reduction during 6 years in Chile, years after the privatization of the social security system. He uses a panel data from 1979 to 1986 for manufacturing firms. His results indicate an increase on wages with no effects on employment.

Different papers have discussed this topic locally. Kugler and Kugler (2008) used a

panel data from 1980 to 1990 for Colombian manufacturing industry, evaluating an increase in payroll taxes; their results indicate that a 10% increase in these taxes decreased formal employment between 4% and 5% over the period of study. This suggests that a reduction in social security taxes, in Latin America countries, would increase demand for skilled workers as in these countries benefits linked to the minimum wage are often not directly linked to contributions. Results could be asymmetric hence the loss in employment because an increase in these contributions might be higher that the employment gain from decreasing these taxes.

Meanwhile, Gaviria (2004) evaluated the labor reform of 2002, which diminished firing costs and increased the ordinary labor journey, among other things. Through indirect evidence of household surveys and direct evidence from a corporate survey using a difference-in-differences approach, found that the reform had no significant effects on employment generation or on formalizing it.

Botero (2012) through a computable general equilibrium model, evaluates the total elimination of payroll taxes with and without compensation in other fiscal income relative to a baseline scenario. In the first scenario, 351,000 employments are created, reducing unemployment rate by 0.6% with an increase of 3.5 points over GPD in public deficit. Under a less adverse scenario for public finances, offsetting the reduction of payroll taxes with an increase in indirect taxes, the outcome is modest: 173,000 employments are created.

The importance of this study relies on the evidence of the positive effects on the labor market it provides. The results we found in our study should provide a useful guide for more policies concerning tax reforms and be used as reference in further studies. This study enriches former literature adding evidence of not only positive effects on hours but also on hourly income.

4 Data and Methods

4.1 Data

We use demographic data from 2012 to 2014 from Great Integrated Household Survey *Gran Encuesta Integrada de Hogares*, GEIH), a nation-wide survey that allows to analyze results by rural and urban areas, regions and departments.

The 2012, 2013 and 2014 surveyed data were obtained from the National Statistics Agency (*Departamento de Administración Nacional de Estadísticas*, DANE), which allows us to review socio-economic indicators at urban level, such as education, social

security, employment, housing and informality.

4.1.1 Unit of analysis

The major entity studied in this work as unit of analysis are the individuals who belong to one of the four following groups: The treated group, conformed by those directly affected by the reform; a first control group, individuals whose salary is equal or higher than ten minimum wages, and a second control group, workers labeled as informal following the DANE's definition regarding the size of the company and their occupation, and a fourth group of those employees that don't pay for any kind of social security or insurance.
All of the individuals analyzed reside in one of the 13 main cities used by the DANE for labor market data gathering, defining main city as a continuous urban area legally recognized by the Colombian state. The cities included in this study are large (Bogota, Medellin, Cali and Barranquilla), mid-size (Bucaramanga, Cucuta, Cartagena and Pereira) and small (Ibague, Manizales, Pasto, Villavicencio and Monteria).

4.1.2 Treatment assignment

In this study we will consider the implementation of the CREE as a treatment which could have an impact on those individuals directly affected by the tax reform. This method implies the definition of both a control group and a treated group.
Usually, a tax reform is planned to redesign a whole set of taxpayers making it difficult to select a specific set of characteristics for a treatment group, however, the CREE Law is specific enough in terms of observable variables which allows to develop a well-defined treatment group.
In this way, our treatment group will be conformed by those individuals who have a formal job and earn less than ten minimum wages, excluding those employees that work in ONGs and non-national firms due to their different tax framework.

4.1.3 Covariates

In order to isolate the causal effect of the tax reform on worked hours and salary, we compile a set of observable covariates that jointly affect the treatment assignment with worked hours and salary outcomes (and assuming that all unobservable variables do not exhibit influence). These covariates are used in our analysis to control for the observable differences between treated and untreated individuals in the observation units,

Table 1: Summary statistics for outcome variables and covariates of interest in the baseline period, December 2012

Variable	Description	Status	Mean	St. Dev.	Min	Median	Max
Hours	Total worked hours in a week	Treated	50.624	12.781	2	48	126
		Control 1	49.519	18.025	4	48	122
		Control 2	47.214	21.961	1	48	130
		Control 3	37.681	27.896	0	45	126
Income-H	monetary hourly income per hour	Treated	5114.056	4640.005	0	3472.222	104166.7
		Control 1	13077.93	20670.21	0	5000	236842.1
		Control 2	3921.968	6443.394	0	2580	120192.3
		Control 3	3525.198	4401.383	0	2604.167	80882.35
Exper	Potential experience	Treated	16.835	11.939	0	14	71
	= Age-Educa-6	Control 1	20.004	13.470	0	19	70
		Control 2	27.206	15.775	0	27	74
		Control 3	19.086	14.441	0	16.00	70.00
Educa	years of education	Treated	11.513	3.527	0	11	17
		Control 1	12.072	4.125	0	13	17
		Control 2	8.555	4.149	0	9	17
		Control 3	9.525	3.835	0	11.00	17.00
Male	=1 if Male	Treated	0.562	0.496	0	1	1
	0 Otherwise	Control 1	0.548	0.497	0	1	1
		Control 2	0.530	0.499	0	1	1
		Control 3	0.612	0.487	0	1	1
Head	=1 if Head of the house	Treated	0.562	0.496	0	1	1
	0 Otherwise	Control 1	0.404	0.490	0	0	1
		Control 2	0.482	0.499	0	0	1
		Control 3	0.366	0.481	0	0	1
Age	Age	Treated	34.346	10.958	14.00	32.00	80.00
		Control 1	38.129	12.874	12.000	37.000	78.000
		Control 2	41.760	14.088	12.00	42.00	80.00
		Control 3	34.602	12.920	12.00	32.00	79.00
Unem-house	% of unemployed in the household	Treated	0.044	0.115	0	0	0.75
		Control 1	0.029	0.093	0	0	0.600
		Control 2	0.041	0.110	0	0	0.750
		Control 3	0.133	0.205	0	0	1.00
Nonlabincome	Non labor income	Treated	152148.7	384777.1	2000	74750	7500000
		Control 1	668763.52	3338609	5833.333	100000	3.43e+07
		Control 2	161641.5	358220.9	2083.333	61666.67	5000000
		Control 3	220267.7	1083155	1666.667	70833.34	1.60e+07

therefore, isolating the impact of the tax reform.

Potential Experience: Total experience is an unobservable variable in the survey, we built potential experience as age minus years of education minus 6, which is supposed to be the age of school entry, this covariate is a determinant of the wage.

Head of the household: in the Colombian Integrated Survey of Households: Individuals self-select as head of household given their sociocultural believes, one may expect that this type of people are the main workforce of the household and participate more actively in the labor market.

% of unemployed in the household: We use these variable as a measure of pressure for offering the workforce into the labor market. Calculated as the number of people who reported being unemployed divided by the members of the household that are part of the labor force.

Non labor income: We built this variable as the sum of rental income, government subsidies, interest income and transfers of money from other households, among others. This variable could measure an income effect on the individual labor supply and the decision of choose more leisure.

4.2 Methods

To reduce the bias associated with our estimates of the effects on worked hours and hourly wage after a tax reform, we use matching as a strategy to control for those heterogeneities which could affect the treatment effect estimate. We use a model of difference-in-differences with additional regressors and repetitive cross section as an ancillary strategy not only to check the robustness of our matching specification but also to control pre-existent differences.

To measure the impact of the cut in payroll costs brought by the CREE Law on working hours and salary in the chosen cities we used difference-in-differences and matching methodologies to estimate the average treatment on the treated (ATT). Estimation of the ATT is implied in our research question: What would have the outcome in wage and worked hours for employees benefitted by the reform been hadn't they been covered by the reform?

$$\tau_{att} = E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 0]$$

$$\hat{\tau_{att}} = (\bar{Y}|D = 1) - (\bar{Y}|D = 0)$$

4.2.1 Differences in differences

The first issue to consider before the estimation is the fact that we are not dealing with a randomized assignment to treatment, instead it is determined by a set of characteristics defined by the tax reform making an individual a treated or an untreated observation in the model, making of this assignment a quasi-experiment Salzberg (1999).

Using the previous assignment, it is now necessary to define the characteristics inside the control group, however, those individuals outside the reform present several heterogeneities among them making it necessary to define not just one control group but three.
Following this statement, our first control group ($D1$) will be those individuals whose salary is equal to or higher than ten minimum wages. The second group ($D2$) includes those workers labeled as informal following the DANE's definition regarding the size of the company and their occupation, and finally, the third group ($D3$) gathers those employees that don't pay for any kind of social security or insurance.

The basic model is a difference-in-differences estimator with additional regressors and repetitive cross section. The inclusion of an auxiliary set of variables responds to the fact that the tax reform implementation acts as a natural experiment and not a controlled one, it is important to highlight that this study faces a context where there is not a perfect randomness or even any of it facing the possibility of systematic differences between the treated and the untreated Bertrand et al. (2002).

$$Y_i = \beta_0 + \beta_1 D_i + \beta_2 I[t = 2] + \beta_3 (D_i * I[t = 2]) + \beta_4 X_{1i} + ... + \beta_{k+3} X_{ki} + u_{it}$$

here D_i equals 1 if the observation belongs to treatment and 0 if not, $I[t = 2]$ is an indicator equals to 1 if $t = 2$ (A time of period after tax reform implementation) and 0 otherwise.
Hence, the basic model includes a sort of variables which measures individual characteristics non directly affected by the treatment. Another bunch of advantages in words of Bernal and Peña (2011) for this specification is not only the chance to control pre-existent differences but also improve the estimator's efficiency, analyze the alleged parallel tendencies and an adjusted estimator when the assignment was carried out conditionally to some observable characteristics of the participant individuals Rosenbaum and Rubin (1983).

This last advantage is very relevant in our study because the design of our analysis group followed a a conditional assignment based on the reform's framework Bernal and Peña (2011).

Using a repetitive cross-section methodology proves itself the fittest one due to the implicit nature of the GEIH survey, because in its development does not include any kind of follow-up for the observations over time. This implies that between months the homes surveyed vary, so in order to develop the estimation its necessary for the data to be randomly collected inside the same population using observations from the first cut as substitutes for the controls and treated of the later cut Stock and Watson (2012).

4.2.2 Matching

The second methodology used is matching as a strategy to estimate ATT by accounting for the covariates that predict receiving the treatment. Here every observation receives a propensity score understood as the probability of being assigned to treatment given a set of covariates Rosenbaum and Rubin (1983).

$$P(X) = P(D = 1|X)$$

If we wanted to estimate the impact of a treatment comparing potential outcomes $Y(0)$and $Y(1)$ of both groups we would have.

$$E(Y_i(1)|D_i = 1) - E(Y_i(0)|D_i = 0)$$

Then including the contrafactual term, $E(Y_i(0)|D = 0)$, we have

$$= E(Y_i(1)|D_i = 1) - E(Y_i(0)|D_i = 1) + E(Y_i(0)|D_i = 1) - E(Y_i(0)|D_i = 0)$$

That way treatment is assigned independent of potential outcomes $Y(1)$ and $Y(0)$ under treatment $(D_i = 1)$ and control $(D_i = 0)$. In the absence of the treatment, one would expect similar average outcomes from both groups. By the other hand, if both groups were to receive treatment, one would expect similar average outcomes from both groups. Formally, if the program assignment is completely based in observable features, we could say that

$$Y(0), Y(1) \perp D|X, \forall X$$

Known as the conditional independence condition, allowing us to assure that $E(Y(0)|D = 1, X) = E(Y(0)|D = 0, X)$ making equal to zero the selection bias. Assuming that these conditions are accomplished , the ATT using matching could be expressed as

$$\tau^{att}_{psm} = E_{P(X)|D=1}\left\{E[Y(1)|D = 1, P(X)] - E[Y(0)|D = 0, P(X)]\right\}$$

5 Results

Table 2 and Table 3 present the results from our primary analyses (table 2 provides results from for our primary specifications in 2013 and Table 3 does it for 2014). In this subsection we focus on the table 2 and table 3, which present the ATT estimates by each of the methodologies.

Table 2 shows that the results for the analysis in 2013 are relatively robust to the introduction of calipers, there is not a significant change in the point estimate of ATT. No observation was dropped from the analysis. The results show an increase in hourly wage and worked hours outcomes across matched treated and untreated individuals (0.253 and 4.69, respectively), and the point estimate is significant at 5% for the hourly wage and at any conventional level for the worked hours. Results for Diff-in-Diff specification showed that the resulting ATT estimate is 0.018 for the hourly wage and 313 for the worked hours, however, these are insignificant at conventional levels and only significant by 5%, respectively.

Table 3 shows that the results for the analysis in 2014 are relatively robust to the introduction of calipers, there is not a significant change in the point estimate of ATT. No observation was dropped from the analysis. The results show an increase in hourly wage and worked hours outcomes across matched treated and untreated individuals (0.313 and 5.26 respectively), and the point estimate is significant at 5% for the hourly wage and at any conventional level for the worked hours. Results for Diff-in-Diff specification showed that the resulting ATT estimate is 0.079 for the hourly wage and 2.81 for the worked hours, however, these are significant at conventional levels and only significant by 10%, respectively.

For both our Matching and Diff in Diff approaches we find evidence to suggest that the implementation of a tax reform over payroll costs reduction had positive effects over hourly incomes and worked hours for those which are covered by the reform.

In other words, after controlling for covariates that jointly influence the participation of individuals inside the reforms, we find that there was differentially greater increases in worked hours and hourly income between 2012 and 2014 for those individuals which were covered by the payroll costs reduction implemented by the government. The point estimates from the primary specification are statistically significant (at 10% significance level) for 2014 when worked hours hours and hourly incomes are used as the outcome of interest for the second control group, but is insignificant when the remaining control groups are used for estimating the outcomes of interest.

Table 2: Results from Primary Analyzes, 2013

	Ln(gross hourly wage)			*Worked Hours*		
	Treated Y(1)	Control Y(0)	ATT	Treated Y(1)	Control Y(0)	ATT
Naive Difference	8.324	7.807	0.516***	50.900	48.002	2.898***
in Means	[9551]	[13443]	{0.00}	[9569]	[14974]	{0.000}
Diff-in-Diff	N.A.	N.A.	0.018	N.A.	N.A.	3.135**
	[9551]	[11946]	{0.500}	[9569]	[13357]	{0.045}
Matching	8.400	8.147	0.253**	49.471	44.773	4.698***
	[1698]	[2684]	(0.044)	[1702]	[2865]	(0.98)
Matching with	8.400	8.147	0.253**	49.471	44.773	4.698***
Calipers[a]	[1698]	[2684]	(0,044)	[1702]	[2865]	(0,98)

[a] Caliper = 0.5 standard deviations.
[Number of observations].
(Standard Error).
{P-value}.
***,** represent significance at the 1% and 5% level, respectively.

Table 3: Results from Primary Analyzes, 2014

	Ln(gross hourly wage)			*Worked Hours*		
	Treated Y(1)	Control Y(0)	ATT	Treated Y(1)	Control Y(0)	ATT
Naive Difference	8.327	7.823	0.504***	50.720	47.847	2.873***
in Means	[9941]	[13220]	{0.00}	[9962]	[14874]	{0.00}
Diff-in-Diff	N.A.	N.A.	0.079***	N.A.	N.A.	2.813*
	[9941]	[11814]	{0.005}	[9962]	[14874]	{0.073}
Matching	8,379	8.066	0.313***	49.532	44.271	5.261***
	[1653]	[2638]	(0.047)	[1659]	[2841]	(0.99)
Matching with	8.379	8.066	0.313***	49.532	44.271	5.261***
Calipers[a]	[1653]	[2638]	(0.047)	[1659]	[2841]	(0.99)

[a] Caliper = 0.5 standard deviations.
[Number of observations].
(Standard Error).
{P-value}.
***,** represent significance at the 1% and 5% level, respectively.

6 Robustness

We test the robustness of our primary estimates in several ways. First, we test the sensitivity of our matching estimator to unobserved heterogeneity between treated and untreated individuals. The objective is to identify by how much the groups would have to differ in order to nullify the results of statistically significant worked hours and hourly income increases.
Second, we test the robustness of our diff-in-diff and matching specifications by comparing our main specification to a number of matching- and regression-based econometric specifications.

Matching specifications inside every observational study must take into account the fact that the elimination of any bias associated with non-randomized selection is settled by the understanding of the selection process **?**, hence, by an array of pertinent features actually observable and available for the researcher. If the selection process and outcomes were determined only by observable features then a treatment effect estimate resulting from a matching procedure properly balanced should be unbiased and consistent Rosenbaum et al. (2002).

However, in case there are unobservable characteristics uncorrelated with the observable characteristics for which we control but also contribute to determining selection and outcomes, then a hidden bias might arise to which matching estimators are not robust. We believe the tax reform´s framework provides not only sufficient and clearly observable covariates with which to control, therefore mitigating unobserved heterogeneity.

Nevertheless, we tested the sensitivity of the ATT estimates for worked hours to unobserved heterogeneity and bias using the Rosenbaum bounds Rosenbaum et al. (2002) The Rosenbaum bounds methodology resolves if the unobserved factors can alter inferences about treatment effects, so it helps to determine how strongly an unmeasured variable must influence the selection process to undermine the implications of the analysis.

If only a little amount of unobserved heterogeneity is necessary to weaken the significance of the results then these are sensitive, by the other hand, if a great amount is required we are obtaining relatively robust results. Table reftable4 indicates the level of unobserved heterogeneity (unaccounted in our matching process for worked hours) that would be required to nullify our results of statistically significant worked hours.

Our results are robust (at the 10% level) to unobserved heterogeneity that affects the odds of selection into treatment (tax reform payroll costs reduction) by a factor of 2.0.
Additionally we carry on some checks on covariate balance in the matching estimates

Table 4: Rosenbaum Upper Bound on P-Value at Given Levels of Γ for Matching Analysis

	Upper Bound
Γ	P-value
1	0.144
1.1	0.126
1.2	0.110
1.3	0.097
1.4	0.085
1.5	0.075
1.6	0.067
1.7	0.059
1.8	0.052
1.9	0.046
2	0.041

as some kind of informal diagnostics, not unlike residuals in a regression, ?, as a helping tool to think about whether our treated and control groups overlap sufficiently to be matched in first place, and whether the match currently under consideration has achieved reasonable balance.

The previous tables showed a slightly imbalance once the primary matching analysis

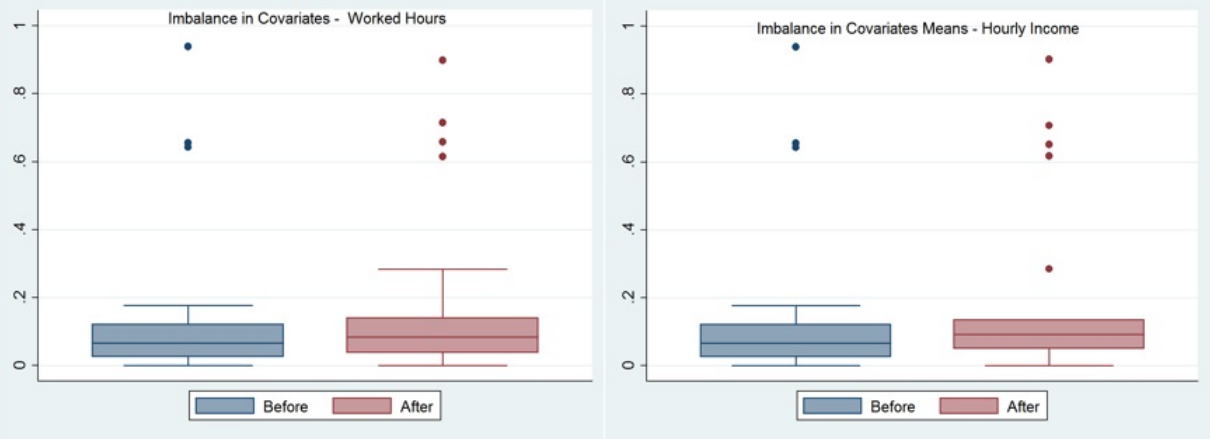

Figure 1: Balance checks for 21 covariates in the study of tax reform effects over labor supply. The box-plots show that the imbalances in covariates were slightly increased by the matching process, in both worked ours and hourly income. Balance on observed covariates does not imply balance on covariates that were not observed.

was solved, this may indicate the necessary inclusion of some penalties in the weight

matrix.

6.1 Robustness of matching specification to alternative econometric specifications

To ensure that our results are not sensitive to the choice of econometric specification, we conduct a series of ancillary matching- and regression-based analyzes. The results of these analyzes can be found in Tables 3 and 4.

Matching with calipers as penalties In our primary econometric specification we use one-to one nearest neighbor matching. Although there is an acceptable level of balance, there could be a number of treated observations that do not obtain a well-matched control.
To ensure that a set of poorly matched observations are not biasing the results, we impose calipers, equal to 0.5 standard deviations, on our primary matching specification. In other words, we intend to remove from the sample any matched pair that differ by more than half standard deviation across covariates values.

Tables 3 and 4 show that the results for the worked hours and hourly income are robust to the introduction of calipers, i.e., there is not any kind of absolute increase or decrease in the point estimates of ATT. On the other hand, the variance of the resulting ATT changes little and the estimates are significant at 10% level for 2014.

6.1.1 Regression-based specifications

We run several matching specifications to ensure that our results are not driven by the choice of an econometric specification. The results of these specifications are found in appendix for full regression results.

Although there is some heterogeneity in the specifications for which worked hours or hourly income are the outcomes, the main results from these specifications are: (1) Treated individuals are associated with significant increases in both worked hours and hourly income and; (2) the results do not differ significantly from the primary matching-based estimates.

6.2 Robustness of matching specification to alternative econometric specifications and control groups design

As in the former section, we also run several Diff-in Diff specifications which differ in the design of those control groups used to estimate the treatment effects over the outputs of interest to ensure that our results are not driven by the choice of an econometric specification or a control group. The results of these specifications are found in tables 5 and 6 for 2013 and tables 7 and 8 for 2014 for full regression results.

7 Conclusions

The central finding in our results is that the tax reform carried out by the Colombian government generated an increase in worked hours and hourly income for those covered by the change in the payroll costs framework. We used two separate methodologies in our analysis and found that the point estimates of worked hours and hourly wage increases are robust across our econometric specifications. The implications of our results are many. First, our results add to a body of literature on the impacts of tax reforms over labor markets. Besides, our findings also support findings in literature that payroll costs reduction brings a positive effect on wages and worked hours. Second, we found several differences from previous studies which indicate that evidence from a single country is likely not generalizable across countries. Further studies in Colombia and elsewhere should strive to identify and quantify the mechanisms through which a payroll costs reduction affect the unemployment and informality levels. Although studies such as ours are important for building a first understanding of the general impacts on worked hours and wages inside the labor markets, only by understanding how a reduction in payroll costs affects the labor market (especially in terms of informality and jobs creation) can policies be designed to enhance the positive impacts of this tax reform.

Bibliography

Bennmarker, H., Mellander, E., and Öckert, B. (2009). Do regional payroll tax reductions boost employment? *labour economics*, 16(5):480–489.

Bernal, R. and Peña, X. (2011). *Guía práctica para la evaluación de impacto*. Universidad de los Andes, Facultad de Economía, Centro de Estudios sobre Desarrollo Económico.

Bertrand, M., Duflo, E., and Mullainathan, S. (2002). How much should we trust differences-in-differences estimates? Technical report, National Bureau of Economic Research.

Betcherman, G. and Pagés, C. (2007). Estimating the impact of labor taxes on employment and the balances of the social insurance funds in turkey. *Synthesis Report. Washington, DC: Banco Mundial*.

Botero, J. (2012). Desempleo e informalidad en colombia: un análisis de equilibrio general computable. In Arango, L. E. and Hamann, F., editors, *El mercado de trabajo en Colombia. Hechos, tendencias e instituciones*. Banco de la Republica.

Cárdenas, M. (2012). Reforma tributaria 2012. Technical report, Ministro de Hacienda y Crédito Público.

Congreso de la Repúplica, C. (2012). *Ley 1605*. Por la cual se expiden normas en materia tributaria y se dictan otras disposiciones En: Diario Oficial No. 48.655. Bogotá: Imprenta Nacional.

Econpubblica (2011). The role and impact of labour taxation policies. Technical report, Universita Bocconi.

Farné, S. and Rodríguez, D. (2013). Bajar los impuestos al trabajo genera empleo? *Cuaderno de Trabajo*, 14.

Gaviria, A. (2004). Ley 789 de 2002:¿ funcionó o no? *Documentos CEDE*, 3140.

Gruber, J. (1995). The incidence of payroll taxation: evidence from chile. Technical report, National Bureau of Economic Research.

Kugler, A. and Kugler, M. (2008). Labor market effects of payroll taxes in developing countries: evidence from colombia. Technical report, National Bureau of Economic Research.

Larrea, C. A. M. and Uribe, J. C. L. (2012). Han sido eficientes y exitosas las reformas tributarias en colombia en el período 1990-2009? *Perfil de Coyuntura Económica*, 20:1–49.

Pissarides, C. A. (1998). The impact of employment tax cuts on unemployment and wages; the role of unemployment benefits and tax structure. *European Economic Review*, 42(1):155–183.

Rosenbaum, P. R. et al. (2002). Covariance adjustment in randomized experiments and observational studies. *Statistical Science*, 17(3):286–327.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.

Salzberg, A. J. (1999). Removable selection bias in quasi-experiments. *The American Statistician*, 53(2):103–107.

Stock, J. H. and Watson, M. W. (2012). *Introduction to Econometrics: Global Edition*. Pearson Education.

8 Appendix

Table 5: Balance Results for nearest neighbor Specification Wage

Covariate	Status	Mean Treated	Mean Non-Treated	Diff. In Means	Nom. Diff
Head of household	Unmatched	0.3913	0.623	-0.231	-0.233
	Matched	0.526	0.623	-0.096	-0.097
Age	Unmatched	30.080	45.27	-15.190	-3.463
	Matched	36.420	45.270	-8.85	-1.689
Stake of unemployed in household	Unmatched	0.079	0.046	0.033	0.064
	Matched	0.043	0.046	-0.003	-0.006
Non-Labour Income	Unmatched	194634.1	164789.6	29844.5	40.251
	Matched	165881.8	164789.6	1092.199	1.343
Years of education	Unmatched	16.47	8.45	8.02	3.517
	Matched	11.77	8.45	3.32	1.162
Potential experience	Unmatched	7.65	30.81	-23.16	-5.059
	Matched	18.66	30.81	-12.15	-2.204
Male	Unmatched	0.95	0.357	0.593	0.720
	Matched	0.464	0.357	0.107	0.108

Table 6: Balance Results for nearest neighbor Specification Hours

Covariate	Status	Mean Treated	Mean Non-Treated	Diff. In Means	Nom. Diff
Head of household	Unmatched	0.416	0.624	-0.208	-0.210
	Matched	0.526	0.624	-0.097	-0.098
Age	Unmatched	30	45.49	-15.49	-3.532
	Matched	36.490	45.49	-9.0	-1.718
Stake of unemployed in household	Unmatched	0.076	0.043	0.033	0.064
	Matched	0.043	0.043	0.0	0.0
Non-Labour Income	Unmatched	170031	168298.1	1732.899	2.337
	Matched	165494.3	168298.1	-2803.8	-3.448
Years of education	Unmatched	16.45	8.47	7.98	3.499
	Matched	11.709	8.47	3.239	1.134
Potential experience	Unmatched	7.56	31.01	-23.45	-5.123
	Matched	18.79	31.01	-12.22	-2.217
Male	Unmatched	0.916	0.35	0.566	0.687
	Matched	0.458	0.35	0.108	0.109

Table 7: Diff-in-Diff regression specifications for Hourly Income 2013

	(1)	(2)	(3)
	Ln(gross hourly wage)	Ln(gross hourly wage)	Ln(gross hourly wage)
I	0.113	0.0639**	0.00614
	(0.0874)	(0.0226)	(0.0475)
DxI	-0.0407	0.0186	0.0674
	(0.0889)	(0.0277)	(0.0501)
D	-0.247***	0.249***	0.254***
	(0.0665)	(0.0209)	(0.0366)
Exper	0.0314***	0.0296***	0.0283***
	(0.00273)	(0.00159)	(0.00201)
Exper2	-0.000296***	-0.000335***	-0.000265***
	(0.0000704)	(0.0000291)	(0.0000448)
Educa	0.132***	0.101***	0.112***
	(0.00334)	(0.00251)	(0.00294)
Male	0.174***	0.235***	0.153***
	(0.0172)	(0.0142)	(0.0163)
Cons	6.565***	6.468***	6.351***
	(0.0839)	(0.0402)	(0.0523)
N	10703	21493	11416
$Cities's Dummies$	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Diff-in-Diff regression specifications for Worked Hours 2013

	(1) Hours	(2) Hours	(3) Hours
I	4.126	-2.202	-1.465
	(3.604)	(1.367)	(3.732)
DxI	-3.566	2.799	1.953
	(3.662)	(1.569)	(3.792)
D	-0.0813	7.168***	4.312
	(2.064)	(1.176)	(2.244)
Ln(Income-H)	-6.397***	-5.042***	-6.866***
	(1.215)	(0.737)	(1.168)
Head	1.999*	3.680***	2.747**
	(0.849)	(0.875)	(0.893)
Age	0.636**	0.890***	0.679**
	(0.231)	(0.179)	(0.260)
Age2	-0.00760**	-0.0109***	-0.00901**
	(0.00286)	(0.00206)	(0.00328)
Unem-house	-0.320	0.908	0.582
	(3.816)	(3.442)	(4.119)
Ln(Nonlabincome)	0.618	-0.289	0.0317
	(0.404)	(0.380)	(0.411)
Educa	0.268	0.137	0.358*
	(0.168)	(0.126)	(0.172)
Cons	80.91***	68.17***	86.19***
	(9.126)	(7.122)	(9.684)
N	1907	4073	1952
$Cities's Dummies$	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: Diff-in-Diff regression specifications for Hourly Income 2014

	(1)	(2)	(3)
	Ln(gross hourly wage)	Ln(gross hourly wage)	Ln(gross hourly wage)
I	0.221*	0.0281	0.0308
	(0.0957)	(0.0241)	(0.0642)
DxI	-0.123	0.0791**	0.0659
	(0.0968)	(0.0283)	(0.0657)
D	-0.246***	0.259***	0.261***
	(0.0668)	(0.0209)	(0.0365)
Exper	0.0258***	0.0271***	0.0267***
	(0.00220)	(0.00158)	(0.00193)
Exper2	-0.000168**	-0.000291***	-0.000254***
	(0.0000513)	(0.0000305)	(0.0000433)
Educa	0.129***	0.0992***	0.108***
	(0.00306)	(0.00240)	(0.00285)
Male	0.147***	0.223***	0.143***
	(0.0168)	(0.0143)	(0.0161)
Cons	6.653***	6.491***	6.421***
	(0.0829)	(0.0383)	(0.0509)
N	11022	21748	11704
$Cities's Dummies$	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: Diff-in-Diff regression specifications for Worked Hours 2014

	(1)	(2)	(3)
	hstrt	hstrt	hstrt
I	4.126	-2.202	-1.465
	(3.604)	(1.367)	(3.732)
DxI	-3.566	2.799	1.953
	(3.662)	(1.569)	(3.792)
D	-0.0813	7.168***	4.312
	(2.064)	(1.176)	(2.244)
Ln(Income-H)	-6.397***	-5.042***	-6.866***
	(1.215)	(0.737)	(1.168)
Head	1.999*	3.680***	2.747**
	(0.849)	(0.875)	(0.893)
Age	0.636**	0.890***	0.679**
	(0.231)	(0.179)	(0.260)
Age2	-0.00760**	-0.0109***	-0.00901**
	(0.00286)	(0.00206)	(0.00328)
Unem-house	-0.320	0.908	0.582
	(3.816)	(3.442)	(4.119)
Ln(Nonlabincome)	0.618	-0.289	0.0317
	(0.404)	(0.380)	(0.411)
Educa	0.268	0.137	0.358*
	(0.168)	(0.126)	(0.172)
Cons	80.91***	68.17***	86.19***
	(9.126)	(7.122)	(9.684)
N	1907	4073	1952
$Cities's Dummies$	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11: Matching specifications 2013

	(1) Treatment	(2) Treatment	(3) Treatment
Head	0.0146	0.248***	0.284***
	(0.0475)	(0.0211)	(0.0825)
Age	0.0891	0.199*	1.089
	(0.279)	(0.0986)	(0.725)
Age^2	-0.00109***	-0.00144***	-0.00182***
	(0.000235)	(0.000113)	(0.000423)
Unem-house	0.263	0.481***	0.372
	(0.186)	(0.0817)	(0.266)
Ln(Nonlabincome)	-0.0429*		
	(0.0181)		
Educa	0.0851	-0.00601	-0.902
	(0.278)	(0.0985)	(0.724)
Exper	-0.0440	-0.145	-1.014
	(0.279)	(0.0988)	(0.725)
Exper2	0.000505**	0.000829***	0.000973**
	(0.000188)	(0.0000914)	(0.000308)
Male	0.307***	0.0889***	0.431***
	(0.0443)	(0.0192)	(0.0708)
lnitf_m		0.289***	0.415***
		(0.0134)	(0.0511)
lnitran_m			-0.0864**
			(0.0302)
# of Children			-0.0224
			(0.0322)
Cons	-1.591	-6.998***	-12.89**
	(1.698)	(0.624)	(4.420)
N	4571	23399	2278

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12: Matching specifications 2014

	(1) Treatment	(2) Treatment	(3) Treatment
Head	-0.00976	0.240***	0.268**
	(0.0475)	(0.0209)	(0.0860)
Age	0.119	0.180*	0.436
	(0.258)	(0.0889)	(0.366)
Age2	-0.00133***	-0.00153***	-0.00218***
	(0.000239)	(0.000113)	(0.000471)
Unem-house	0.440*	0.625***	0.538*
	(0.185)	(0.0816)	(0.267)
Ln(Nonlabincome)	-0.0495**		
	(0.0189)		
Educa	0.0733	0.0179	-0.218
	(0.258)	(0.0887)	(0.364)
Exper	-0.0615	-0.120	-0.345
	(0.258)	(0.0891)	(0.365)
Exper2	0.000643***	0.000852***	0.00117***
	(0.000190)	(0.0000918)	(0.000335)
Male	0.277***	0.0993***	0.195**
	(0.0446)	(0.0190)	(0.0735)
lnitf_m		0.314***	0.426***
		(0.0134)	(0.0533)
lnitran_m			-0.124***
			(0.0315)
# of Children			-0.0312
			(0.0336)
Cons	-1.819	-7.230***	-8.852***
	(1.569)	(0.571)	(2.335)
N	4524	23578	2188

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$