Vol. 51 No. 1 (May, 2014), 1–39

LAJE
Latin American
Journal of Economics

COULD AN INCREASE IN EDUCATION RAISE INCOME INEQUALITY? EVIDENCE FOR LATIN AMERICA*

Diego Battistón**
Carolina García***
Leonardo Gasparini****

This paper explores the direct effect of an education expansion on the level of earnings inequality by carrying out microsimulations for most Latin American countries. We find that the direct effect of the increase in years of education in the region in the 1990s and 2000s was unequalizing; this result is expected to hold for future expansions if increases in education are not highly progressive. Both facts are closely linked to the convexity of returns to education in the labor market. On average, the estimated impact of the education expansion remains unequalizing when allowing for changes in returns to schooling, although the effect becomes smaller.

JEL classification: I24, I25, D31

Keywords: Education, inequality, earnings, Latin America

1. Introduction

Increasing education is one of the main ingredients in a typical recipe for development with equity. An upgrading of the human capital of a population is expected to contribute to higher productivity and hence a generalized increase in well-being, and also reduce income inequality. However, the link between education and inequality may not be that straightforward. Given that there may be convexities in returns to education, even an equalizing increase in schooling may generate an unequalizing change in the distribution of labor incomes. Bourguignon, Ferreira, and Lustig (2005) have labeled this phenomenon "the paradox of progress," a situation where educational expansion is associated with higher income inequality. In this paper we explore whether this is merely a theoretical possibility with little

* We are grateful to Jorge Paz, Nora Lustig, and seminar participants at Universidad Nacional de la Plata, Universidad de San Andrés, Fundación Vidanta, the World Bank and the annual meetings of the Asociación Argentina de Economía Política (AAEP). We would also like to thank two anonymous referees for their comments and suggestions. All errors and omissions are our responsibility.

** Centre for Economic Performance, London School of Economics; CEDLAS, *Facultad de Ciencias Económicas*, *Universidad Nacional de La Plata*; CONICET.

*** CEDLAS, *Facultad de Ciencias Económicas*, *Universidad Nacional de La Plata*.

**** CEDLAS, *Facultad de Ciencias Económicas*, *Universidad Nacional de La Plata* and CONICET.

doi 10.7764/LAJE.51.1.1

relevance in practice or a widespread phenomenon across real-world developing economies.

Towards that end, we perform microeconometric decompositions that isolate the direct effect of changes in the distribution of education on earnings inequality. In particular, we estimate the counterfactual distribution of individual earnings that would be generated in a given period t if the distribution of education took the observed values in t^* and all other variables remained at their values in t. The difference between the real earnings distribution and the counterfactual one characterizes the direct impact of the change in the distribution of education on the earnings distribution. The methodology is applied to household survey microdata for the Latin American countries in the period 1990-2009, exploiting a dataset that includes homogeneous definitions for the education, labor and income variables used in the analysis.

We find that the direct effect of the increase in education experienced by most countries in Latin America in the last two decades was unequalizing, a result that is closely linked to the convexity of returns to education. The paper includes simulations of alternative future changes in the distribution of education and concludes that even education reforms that lead to an equalizing increase in schooling may be associated with higher earnings inequality.

The paper makes two main contributions. On the one hand, it adds to the literature on education and inequality by highlighting a link between these two variables that is usually neglected, and by providing empirical evidence on its practical relevance. On the other hand, the paper contributes to the growing literature on the determinants of changes in inequality in Latin America (López Calva and Lustig, 2010; Gasparini and Lustig, 2011; Cornia, 2011) by examining a channel whose potential relevance has been recognized, but for which only scattered evidence has been available. To that aim the paper uses a unique homogenous dataset comparable across years and countries that covers all Latin American economies.

The rest of the paper is organized as follows. Using a simple model, in Section 2 we briefly illustrate the links between education and earnings and discuss the possibility of the paradox of progress. In Section 3 we explain the methodology of the microeconometric decompositions and comment on the data used. Section 4 presents the results of applying the microsimulations to characterize changes in earnings inequality in Latin America during the last two decades, while Section 5 presents projections of earnings inequality under alternative education upgrading scenarios. Section 6 extends the analysis from previous sections in order to allow for changes in returns to education. Section 7 provides concluding remarks.

2. The theoretical link

The most frequent general policy advice for a developing country is to increase the educational level of its population. Without much discussion, a reduction in inequality is often included in the list of the several positive consequences of an educational expansion.[1] However, if returns to education are convex, an increase in schooling in the population may lead to higher earnings inequality even when the upgrade is moderately biased toward less-educated groups. Bourguignon *et al.* (2005) have dubbed this phenomenon "the paradox of progress," a situation where an education expansion is accompanied by a surge in earnings inequality.

This argument refers to the first-round, partial-equilibrium impact of the increase in education on inequality, and in particular assumes no change in returns to skills. Naturally, an education expansion, by shifting the supply of skilled labor, may reduce the wage premium and contribute to a reduction in earnings inequality. Assessing the overall, long-run general equilibrium impact of an increase in schooling on income distribution is certainly a very challenging task, one we cannot fully address in this paper. For this reason, we tackle the issue in two steps. First, we estimate the size of the initial direct impact of an education expansion, assuming no changes in returns to schooling, in order to illustrate the potential for the paradox. Second, we estimate changes in returns to education following the methodology proposed by Katz and Murphy (1992), which although it falls short of a full general equilibrium model, provides good approximations of the relevant parameters and has been extensively used in the literature (Card and Lemieux, 2001; Manacorda, Manning, and Wadsworth, 2012). Using these estimates we perform a robustness exercise to determine if the paradox still holds under changing returns.

We start in this section by illustrating the possibility of an inequality-increasing expansion of education with a simple model. Consider first

1. The link between education and inequality has been addressed by Benhabib and Spiegel (1994), Krueger and Lindhal (2001), Soto (2002) and Pritchett (2001), among others.

that the logarithm of individual earnings Y_i is related to the individual level of education X_i in a linear way. Ignoring other determinants for simplicity's sake, this relationship at period t can be expressed as

$$\ln Y_{it} = \alpha_t + \beta_t X_{it} + \varepsilon_{it} \tag{1}$$

where unobservable determinants are summarized in the zero-mean term ε_i. Under the assumption of independence between X_i and ε_i parameter β is interpreted as a measure of returns to education.[2] Assume that the whole set of income earners can be divided into two groups H and L, with $X_H > X_L$, and $E(\ln Y_H) > E(\ln Y_L)$. A simple measure of earnings inequality in this two-group society is the expected proportional earnings gap G. Taking conditional expectation and rearranging,

$$G \equiv E(\ln Y_{Ht} - \ln Y_{Lt}) = \beta_t \, (X_{Ht} - X_{Lt}) \tag{2}$$

From Equation (2) the change in earnings inequality between periods 1 and 2 can be expressed as

$$\begin{aligned} \Delta G &\equiv E(\ln Y_{H2} - \ln Y_{L2}) - E(\ln Y_{H1} - \ln Y_{L1}) \\ &= (\beta_2 - \beta_1)(X_{H1} - X_{L1}) + \beta_2(dX_H - dX_L) \end{aligned} \tag{3}$$

where dX_i is the change in the level of education for earners in group $i = H,L$. Equation (3) implies that the change in inequality depends on changes in returns to education over time $(\beta_2 - \beta_1)$, the initial difference in educational levels $(X_{H1} - X_{L1})$, and the relative change in education $(dX_H - dX_L)$. If returns to education do not vary over

2. In order to keep the notation uncluttered, we assume independence between education and the unobservable component. It is straightforward to allow for other assumptions like mean independence but the main conclusions remain unchanged.

time and the growth in educational levels is similar across groups, earnings inequality remains unchanged.

These results are modified when we allow the model to include convex returns to education. Assume that the logarithm of earnings and education are related through a quadratic function:

$$\ln Y_{it} = \alpha_t + \beta_t X_{it} + \gamma_t X_{it}^2 + \varepsilon_{it} \tag{4}$$

In such a case, the expected change in the proportional gap of earnings between H and L takes the form:

$$\begin{aligned} \Delta G = {} & (\beta_2 - \beta_1)(X_{H1} - X_{L1}) + \beta_2(dX_H - dX_L) \\ & + (\gamma_2 - \gamma_1)(X_{H1}^2 - X_{L1}^2) + \gamma_2(dX_H^2 - dX_L^2) \\ & + 2\gamma_2\,(X_{H1}dX_H - X_{L1}dX_L) \end{aligned} \tag{5}$$

Notice that when returns to education remain unchanged and changes in education across groups are similar, Equation (5) becomes $\Delta G = 2\gamma_2(X_{H1} - X_{L1})dX$, which is positive under convex returns to education: inequality rises in response to an equal increase in education across the population. From (5), if returns to education do not change and returns are convex, even an unbalanced increase in education in favor of the unskilled group L may lead to a surge in earnings inequality. To see this, assume $dX_L = \lambda dX_H$ with $\lambda > 1$. Earnings inequality G increases in this case if

$$X_{H1} - \lambda X_{L1} > \frac{1}{2}(\lambda - 1)(\frac{\beta_2}{\gamma_2} + (\lambda - 1)dX_H) \tag{6}$$

which is more likely to occur with highly convex returns to education. Similarly, if the convexity is sufficiently high, earnings inequality may increase even after an education expansion that reduces returns to skills.[3]

3. Alejo (2012) explores another possibility for an increase in earnings inequality that arises due to the greater variability of earnings at higher educational levels.

3. Empirical strategy

This section presents an empirical strategy to provide evidence on the direct impact of changes in education on earnings inequality. The methodology follows Gasparini, Marchionni, and Sosa Escudero (2005), which in turn is based on Bourguignon, Ferreira, and Lustig (2005). It requires the estimation of earnings equations at the individual level and the use of the resulting coefficients to construct counterfactual distributions. Earnings are modeled as parametric functions of observable characteristics, and the residuals of the regressions are interpreted as the effect of unobservable factors. In this section we describe the methodology that we follow to estimate the counterfactual distribution of individual earnings that would be generated in a given period t (or country p) if the distribution of education took the observed values in t^* (or p^*) and the rest of the earning's determinants remained at their values in t (or p). The difference between the real distribution and the counterfactual one characterizes the direct first-round distributional impact of the change in the distribution of education.

3.1. Empirical model

Following Gasparini *et al.* (2005), we represent the individual earnings-generating process at time t as

$$\ln Y_{it} = F(X_{it}, Z_{it}, \varepsilon_{it}, \beta_{Xt}, \beta_{Zt}) \tag{7}$$

where Y_{it} are individual earnings, X_{it} is the vector of individual observable characteristics related to education, Z_{it} is the vector of observable non-educational characteristics, ε_{it} is the vector of individual unobservable characteristics, and β_{Xt} and β_{Zt} are the vectors of parameters that link X_{it} and Z_{it} with Y_{it}.

The distribution of individual earnings is a vector

$$D_t \equiv \{ Y_{1t}, \ldots, Y_{Nt} \} \tag{8}$$

where N is the number of workers in the economy. Our microsimulation strategy consists of estimating the counterfactual income distribution that would arise if the educational structure were different from the actual structure. In particular, we perform three types of exercises: (i) simulate the counterfactual earnings on year t assuming an educational structure similar to that observed in year t^*; (ii) simulate the counterfactual earnings of a country p assuming an educational structure similar to that observed in country p^*; and (iii) simulate the counterfactual earnings that would arise under different education upgrading scenarios (e.g., an increase of one year of education for each worker in the population).

The counterfactual log income for individual i in year t if X^* instead of X were observed can be defined as

$$\ln Y_{it}(X_{it}^*) = F(X_{it}^*, Z_{it}, \varepsilon_{it}, \beta_{Xt}, \beta_{Zt}) \tag{9}$$

Notice that we are measuring only the direct impact of a change in X, and then in (9) we keep all other factors in the income-generating function fixed. The counterfactual earnings distribution is then

$$D_t(X^*) = \{ Y_{1t}(X_{it}^*), \ldots, Y_{Nt}(X_{Nt}^*) \} \tag{10}$$

Therefore, if we measure inequality by means of an index $I[D]$, the direct impact of the change in the educational structure X on earnings inequality is

$$I[D_t(X^*)] - I[D_t] \tag{11}$$

3.2. Estimation strategy

In order to calculate (11), we need to obtain estimations of the vectors of parameters β_{Xt} and β_{Zt} and the vector of unobservable characteristics ε_{it}. Moreover, given that no panel data is available for our purpose, we need a device to replicate the educational structure of one year (or country) into the population of another year (or country).

The estimations of β_{Xt}, β_{Zt} and ε_{it} are obtained from standard Mincer equations (Mincer, 1974), in which we model the logarithm of individual monthly earnings as a linear function of observable individual characteristics:

$$\ln Y_{it} = \alpha_t + X_{it}\beta_{Xt} + Z_{it}\beta_{Zt} + \varepsilon_{it} \tag{12}$$

Education-related characteristics X_{it} are alternatively measured by a set of dummies for the highest educational level completed and by the number of years of formal education and its square, while observable non-educational characteristics Z_{it} include age, age squared, a gender dummy, a dummy for youths less than 18 years old, regional dummies, and an urban/rural dummy. There are well-known limitations derived from the econometric specification of this model. In particular, it is difficult to identify returns to education from returns to unobservable skills given that they are potentially correlated.[4] Data limitations do not allow us to instrument educational variables (Angrist and Krueger, 1991) in order to obtain consistent estimations of returns to education.

In order to replicate the educational structure of year t^* (or country p^*) into the population of year t (or country p), we use two alternative methods. The first is adapted from Gasparini *et al.* (2005) who propose dividing the adult population into homogeneous age-gender groups (cells) and then replicating the levels of education of a certain cell in year t^* into the corresponding cell of year t. The procedure requires the selection of individuals who "move" from one level of education to another until the desired structure is replicated. This selection process is random, but we impose the restriction that individuals move sequentially across levels. For example, assume that we need to replicate the "incomplete primary school" level of a particular cell in year t^* into the population of year t and that this implies increasing the number of individuals with this level in t. We start by assigning to the counterfactual "incomplete primary school" level all the individuals in t with this level of education (and the corresponding

4. If A is the unobserved ability correlated with earnings by the parameter φ and related to education by $X = \rho A + v$, then OLS estimation of (4) results in the usual asymptotical ability bias ($\varphi\rho$) for β_t. It is straightforward to show that if ability is symmetrically distributed, OLS consistently estimates γ_t in (4). Assuming that ability is time invariant, our simulated inequality change differs from the "true change" by the term $\varphi\rho$ $(dX_H - dX_L)$ (see Equation (5)). According to Card (1999) in most empirical applications IV estimations of β_t are higher than OLS (implying a negative $\varphi\rho$ or measurement error); therefore, our simulations would be a lower bound for the simulated change when ability is observed.

age). Subsequently, since more people are needed in this level-age cell in order to achieve the counterfactual size, those individuals who have completed primary school would be chosen, followed by, those with incomplete secondary education, and so on until the share of people with "incomplete primary level" in this cell in year t^* is reached.

The second procedure closely follows Legovini, Bouillón, and Lustig (2005). The adult population of year t is also divided into homogeneous age-gender cells. For each individual i within cell j we perform the following transformation over the variable years of formal education:

$$X_{ijt}^* = (X_{ijt} - \mu_{jt})\left(\frac{\sigma_{jt}^*}{\sigma_{jt}}\right) + \mu_{jt}^* \tag{13}$$

where μ_{jt} and σ_{jt} are the sample mean and standard deviation within cell j in year t, whereas μ_{jt}^* and σ_{jt}^* are the sample mean and standard deviation estimated for the corresponding cell j in year t^*. For each cell in year t this adjustment results in a distribution of the years of education with mean and variance similar to the corresponding cell in year t^*.

As emphasized above, the approach outlined provides estimations of the partial-equilibrium, first-round impact of a change in the distribution of education on earnings inequality. Of course, if educational levels are modified, other variables that are fixed in the analysis may change, such that the final effect of a shock in education may differ from the direct impact. For instance, as the population becomes more educated, the change in the relative supply of skilled workers modifies returns to education, which can in turn compensate for the first-order unequalizing impact.[5] There are two main justifications for going ahead with the decompositions despite this important caveat: (i) estimating a full general equilibrium model that properly takes into account the movement of all the relevant variables is beyond the technical capabilities in many cases, and (ii) it is illustrative of the direction and magnitude of the direct impact of a change, which in many applications turns out to be the most important. In addition, while in the next two sections we estimate the direct impact of an education expansion, in Section 6 we estimate changes in returns to education and carry out a robustness analysis of the main results.

5. See for example Katz and Murphy (1992) and Manacorda *et al.* (2010).

3.3. Dataset and methodological decisions

The main source of data for this paper is the Socio-Economic Database for Latin America and the Caribbean (SEDLAC), jointly developed by the Center for Distributive, Labor and Social Studies (CEDLAS) at the *Universidad Nacional de La Plata* (Argentina) and the World Bank's Latin America and the Caribbean Poverty and Gender Group (LCSPP). This database contains information on more than 300 national household surveys in 25 Latin American and Caribbean (LAC) countries. All variables in SEDLAC are constructed using consistent criteria across countries and years, and identical programming routines (see sedlac.econo.unlp.edu.ar). In this paper we use microdata for 18 Latin American countries, covering the period 1990-2009.[6]

All calculations are performed using the subsample of workers aged 14 to 65 and, following a standard procedure, we exclude from the inequality measurement and Mincer estimations those individuals who do not receive any payment for their work. We define the logarithm of monthly labor income as the dependent variable in Mincer equations. Given that the structural relationship between individual characteristics and earnings could be different for heads and other members of the household, we follow Gasparini *et al.* (2005) and separately estimate models for the head of household, the spouse, and other members.

As we discussed in previous sections, a key factor in the relationship between education and inequality is the convexity of returns to education. Parametric assumptions about a particular functional form of these returns may modify the results. In our estimations we include education using two alternative definitions: (i) years of formal education and (ii) dummies for the highest educational level completed by each individual. The first definition, in which years of schooling is used as educational variable, allows us to obtain a parametric measure of the convexity of returns by means of the coefficient of the squared variable. On the other hand, the dummies for educational levels allow for a more flexible estimation of the structure of returns to education. As described above, we use a different simulation method for each type of educational variable. Notice that results from both types of simulations can substantially differ because there is not a

6. For comparison purposes, in each country we restrict the sample to the areas covered by the national household survey in the entire period of analysis. Therefore, in Argentina we restrict the sample to the 15 cities covered in the 1992 survey, in Brazil we exclude rural-northern areas that have been included since 2004, and we only use urban areas from Uruguay since rural areas were only added in 2006.

direct correspondence between a change in years of education and a change in the share of workers with different levels of schooling. For instance, an increase in years of education could have little impact on the education structure if it is insufficient to move enough people to the subsequent level. We perform non-parametrical estimations to provide evidence on the convexity of returns and the validity of the quadratic specification.

4. Results

In this section we present the results of the microsimulations in order to characterize changes in earnings inequality during the 1990s and 2000s in 13 Latin American economies. In particular, we seek to evaluate how the education expansion in these countries affected the earnings distribution. To do this, we start with a brief description of the changes in years of education during the period 1990-2009.

4.1. Changes in education

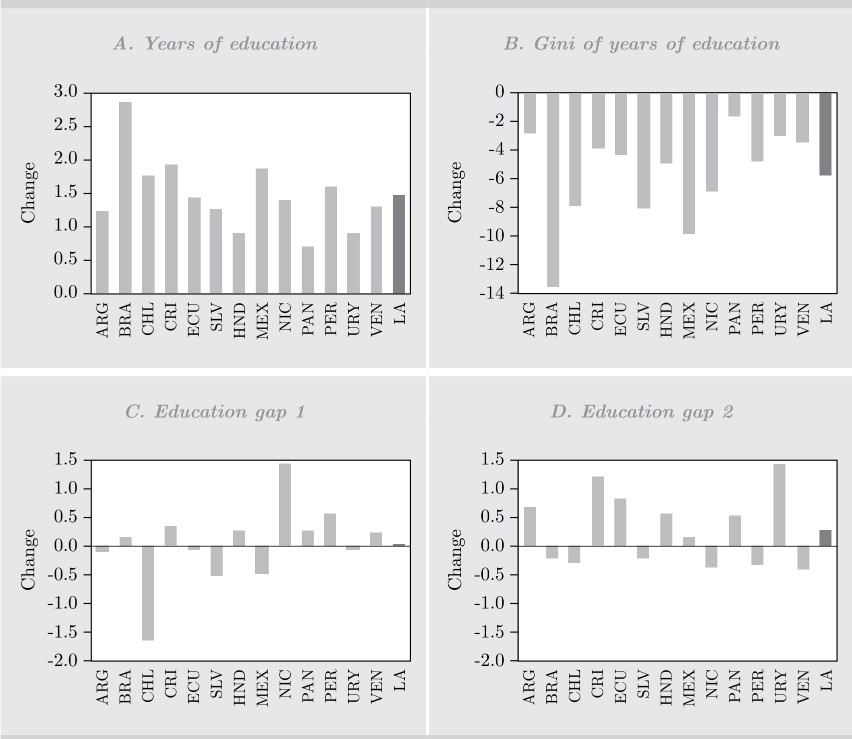
All countries in Latin America experienced a substantial education expansion during the 1990s and 2000s (Figure 1). On average, the number of years of formal education for the working population grew by 1.5 years between 1990 and 2009, with a minimum of 0.7 in Panama, and a maximum of 2.9 in Brazil.

This educational expansion was not homogeneous across population groups. To examine educational inequality we report three different measures. The educational Gini measures relative inequality in the distribution of years of schooling, independently of income, while the difference in the average years of education between the top and bottom quintiles of that distribution (Gap 1) measures absolute inequality in education, and the difference in mean years of education between the richest and poorest earnings quintiles (Gap 2) is a measure of absolute inequality in education relative to earnings.[7]

7. Whether a change in years of education should be evaluated using a relative or absolute definition of inequality is a matter of subjective assessment. Nevertheless, for non-monetary variables like education, it is sometimes more natural to evaluate changes in absolute rather than relative terms (Kolm, 1977). In the case of years of schooling, an absolute inequality measure remains constant under identical additions of years of education to all individuals, whereas a relative indicator remains unchanged under proportional increments in this variable. If we multiply every individual's years of education by a constant, the Gini coefficient, which is a relative index, remains constant, whereas an absolute indicator, such as the educational gap, increases.

Figure 1. Changes in years of education and educational inequality, 1990-2009

Working population [14, 65]

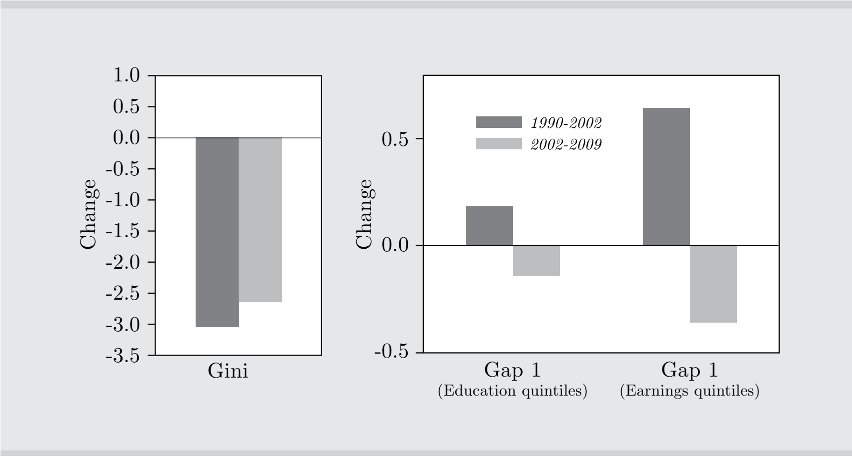


Source: Own calculations based on microdata from household surveys.
Note: LA = mean for Latin America. Gap 1: difference in mean years of education between quintiles 5 and 1 of the distribution of years of education. Gap 2: difference in mean years of education between quintiles 5 and 1 of the distribution of earnings.

During the last two decades education inequality measured by a relative index (Gini of years of education) fell in all countries, whereas results are mixed when using absolute indicators such as the educational gaps. The difference in years of education between extreme education quintiles dropped in three countries, increased in two and remained relatively unchanged in the rest. Measured by the gap between earnings quintiles, education became more unequal in seven countries, whereas in the rest inequality fell slightly. The average educational Gini coefficient in Latin America fell by 5.7 points, whereas the gap between years of education quintiles remained unchanged and the educational gap between earnings quintiles rose by 0.3 years.

Figure 2. Changes in educational inequality

Mean for Latin American countries, 1990s and 2000s



Source: Own calculations based on microdata from household surveys.

Changes in the measures differed greatly in the two decades under analysis. While the mean educational Gini dropped substantially in both periods, the average educational gaps increased between 1990 and 2002, but decreased between 2002 and 2009 (Figure 2). These results suggest that the education growth path was biased toward more educated (or wealthier) groups between 1990 and 2002, and slightly biased toward less educated (or poorer) groups between 2002 and 2009. This break in educational gaps can entail dissimilar effects on earnings inequality during each sub-period, as we will discuss later in this section.

4.2. Results from the microsimulations

For each country/period, Table 1 reports the actual change in the Gini coefficient of the earnings distribution, along with the counterfactual changes simulated by altering the educational structure. For Simulation 1 we use education levels as the relevant educational variable, whereas in Simulation 2 we use years of formal education. Given that the results are path dependent, we alternatively simulate (i) the change in the Gini coefficient if the education structure of the first year of the period is simulated on the last-year population, and (ii) the change in the Gini coefficient if the education structure of the last year is

Table 1. Effect of change in education distribution on earnings inequality

Country	Period	Observed Gini			Education effect (Δ Gini)	
		t_1	t_2	Change	Simulation 1	Simulation 2
Argentina	1992-2009	39.4	40.1	0.7	0.2 ***	1.2 ***
Brazil	1992-2009	50.4	51.1	0.7	1.0 ***	1.6 ***
Chile	1990-2009	52.5	50.2	-2.3	0.6 ***	0.7 ***
Costa Rica	1990-2009	40.0	45.4	5.4	0.9 ***	3.2 ***
Ecuador	1994-2009	53.3	45.5	-7.8	0.4 ***	2.1 ***
El Salvador	1995-2008	45.6	44.6	-1.0	2.5 ***	1.5 ***
Honduras	1995-2009	52.4	52.0	-0.4	1.7 ***	1.0 ***
Mexico	1989-2008	48.1	49.3	1.2	0.6 ***	1.1 ***
Nicaragua	1993-2005	53.6	49.4	-4.2	0.9 ***	1.3 ***
Panama	1991-2009	47.0	47.4	0.4	0.2 ***	2.0 ***
Peru	1997-2009	50.4	50.5	0.1	-0.0	1.7 ***
Uruguay	1992-2009	44.9	47.7	2.8	-0.9 ***	0.5 ***
Venezuela	1992-2006	36.7	37.8	1.1	0.6 ***	0.7 ***
Average				*-0.2*	*0.7*	*1.4*

Source: Own calculations based on microdata from household surveys.
Note: Simulation 1 follows Gasparini *et al.* (2005) to change the educational structure, while Simulation 2 follows Legovini *et al.* (2005). Workers aged 14 to 65. Significance levels obtained using 200 bootstrap repetitions.

simulated on the first-year population. We report the average of the results obtained from each procedure.[8]

The interpretation of Table 1 is straightforward. For example, in the case of Brazil the Gini coefficient for the earnings distribution increased 0.7 points between 1992 and 2009. The first simulation reveals that the education expansion had a direct, first-round unequalizing impact on the earnings distribution of approximately 1 Gini point. If only the educational structure had changed between 1992 and 2009, the Gini coefficient for the earnings distribution would have increased by 1 point.

Under Simulation 1 the education expansion had an unequalizing impact on earnings in 11 countries, while it was equalizing only in Uruguay.[9] As mentioned before, Simulation 2 uses years of education

8. The estimated coefficients of Mincer equations for each country are available upon request.

9. Similarly, Bourguignon *et al.* (2005) find that in five of the seven countries studied (Argentina, Colombia, Mexico, Indonesia, and Malaysia) the effect of educational expansions was to increase inequality. Other authors also report similar results for other countries (Langoni, 1973; Almeida dos Reis and Paes de Barros, 1991; Knight and Sabot, 1983; Reyes, 1988; Lam, 1999).

instead of levels of schooling in order to measure changes in education. In this case, the estimated effects are always unequalizing. In addition, for most countries increases in inequality are more pronounced than those estimated under Simulation 1.

Table 2 splits the results from Table 1 into two sub-periods: 1990-2002 and 2002-2009. The outcomes from Simulation 1 indicate that during the 1990s changes in education in Latin American countries had, on average, a direct unequalizing impact on the earnings distribution of 0.6 Gini points, whereas in the 2000s the average estimated increase was 0.2. Simulation 2 reveals a similar pattern: the average simulated increase in the earnings Gini was 1.3 between 1990 and 2002, and only 0.4 between 2002 and 2009. The difference in the magnitude of the unequalizing impacts of the educational expansions in the 1990s and 2000s is consistent with the dissimilar patterns in the educational gaps documented above. In the 1990s the combination of convex returns (as we will see below) and educational improvements biased toward the most educated (or wealthier) groups resulted in a larger unequalizing effect on the earnings distribution. In contrast, during the 2000s educational changes seemed to be slightly biased toward less educated (or poorer) groups, a fact that resulted in a smaller unequalizing effect on earnings. In fact, for some countries (Argentina, Chile, Honduras, Peru, and Uruguay) the educational expansion had a direct equalizing impact on earnings.

Convexity of returns to education

As discussed in Section 2, the way in which an education expansion affects earnings inequality critically depends on the convexity of returns to education. Convexity has been widely documented for Latin American labor markets (e.g., Gasparini *et al.* (2005) for Argentina; Legovini, Bouillón, and Lustig (2005) for Mexico; and Blom, Holm-Nielsen, and Verner (2001) for Brazil).[10] [11] In addition, there is literature that documents and discusses determinants of the increase in the degree of convexity of returns in Latin America, mainly in the

10. For Chile, Rau (2013), uses a general modeling framework for nonparametric models with endogenous regressors and heterogeneity to estimate the returns to education. The author finds that the local average returns to schooling are highly nonlinear.

11. Using quantile regression estimates of the returns to schooling over a sample of male workers in 16 developed countries during the mid-1990s, Martins and Pereira (2004) find that returns to education increase along the wage distribution.

Table 2. Effect of change in education distribution on earnings inequality (Gini index)

Results of microeconometric decomposition

Country	1990s				2000s			
	Period	Δ Gini	Education effect		Period	Δ Gini	Education effect	
			Simulation 1	Simulation 2			Simulation 1	Simulation 2
Argentina	1992-2004	5,5	0,4 ***	1,3 ***	2004-2009	-4,8	-0,3 ***	0,1 ***
Brazil	1992-2002	5,1	0,5 ***	1,6 ***	2002-2009	-4,4	0,8 ***	1,7 ***
Chile	1990-2003	0,4	0,4 ***	0,6 ***	2003-2009	-2,8	-0,3 ***	-0,1 **
Costa Rica	1990-2002	5,9	1,0 ***	2,1 ***	2002-2009	-0,6	0,1	0,7 ***
Ecuador	1994-2003	-2,6	-0,6 ***	2,2 ***	2003-2009	-5,2	1,5 ***	0,1 *
El Salvador	1995-2002	1,9	2,3 ***	0,9 ***	2002-2008	-3,0	0,0	1,0 ***
Honduras	1995-2002	1,0	0,5 ***	1,4 ***	2002-2009	-1,4	1,0 ***	-0,6 ***
Mexico	1989-2002	2,4	0,3 ***	1,1 ***	2002-2008	-1,2	0,6 ***	0,4 ***
Nicaragua	1993-2001	3,7	1,0 ***	1,5 ***	2001-2005	-7,9	0,5 ***	1,3 ***
Panama	1991-2002	5,7	0,2 ***	1,8 ***	2002-2009	-5,3	-0,0	0,4 ***
Peru	1997-2003	4,8	0,6 ***	1,8 ***	2003-2009	-4,7	-0,1 *	0,3 ***
Uruguay	1992-2002	3,2	0,9 ***	0,5 ***	2002-2009	-0,4	-1,6 ***	-0,1 ***
Venezuela	1992-2002	6,6	0,6 ***	0,7 ***	2002-2006	-5,5	0,1 ***	0,3 ***
Average		*3,4*	*0,6*	*1,3*		*-3,6*	*0,2*	*0,4*

Source: Own calculations based on microdata from household surveys.
Note: Simulation 1 follows Gasparini *et al.* (2005) to change the educational structure, while Simulation 2 follows Legovini *et al.* (2005). Workers aged 14 to 65. Significance levels obtained using 200 bootstrap repetitions.

1990s (Manacorda *et al.*, 2010; Attanasio *et al.*, 2004; Pavcnik *et al.*, 2005; Binelli, 2008; Gasparini *et al.*, 2011).

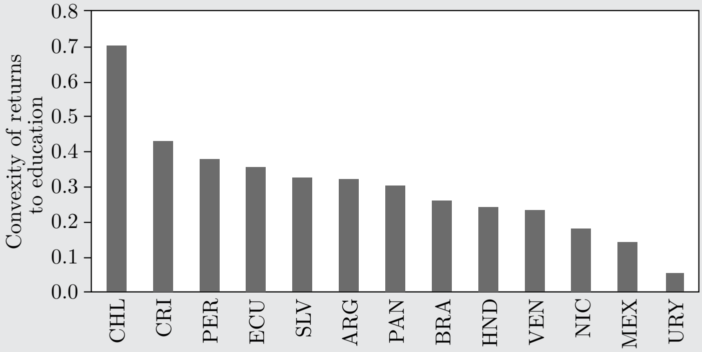
The convexification of returns to education has been previously discussed for different countries. Mincer (1998) and Deschênes (2002) highlight this phenomenon for the United States, and Lemieux (2006) also shows that wages have become a much more convex function of education in the mid-2000s compared to the mid-1970s. Mehta *et al.* (2013) find evidence that the expansion of the service sector drove the convexification of returns in India, the Philippines, and Thailand. Savanti and Patrinos (2005) show evidence for Argentina in the period 1992-2002. Other examples are Esquivel and Rodriguez-Lopez (2003) and Binelli (2012) for Mexico, and Soderbom *et al.* (2003) for Kenya and Tanzania.

Different explanations have been proposed to account for the convexification of returns to education. As pointed out by Binelli (2012), most of them are demand-driven explanations, such as the Lemiux (2006) model with heterogeneous returns to schooling, or the task-based technical change model of Autor, Katz and Kearney (2006) where new technologies have varying complementarity with skilled and unskilled labor.[12]

In the context of LAC countries, there is an extensive literature linking the changes in returns to education with trade and market reforms (e.g., Attanasio *et al.*, 2004; Pavcnik *et al.*, 2005; Revenga, 1997; and Galiani and Sanguinetti, 2003). In a more comprehensive study Behrman, Birdsall, and Székely (2006) analyze the effects of trade liberalization policies during the 1990s for 18 Latin American countries. They find that in most of these economies returns to higher education dramatically increased, whereas returns to secondary and primary school decreased. The authors discuss some potential mechanisms behind this convexification process. First, trade liberalization itself could have shifted demand toward industries intensive in natural resources or land, displacing the production of unskilled-intensive industries to China and other Asian economies (as implied by Spilimbergo *et al.*, 1999). Trade liberalization could also have increased demand for intermediate goods that are intensive in skilled labor in developing countries (Feenstra and Hanson, 2001). Second, inflows of external capital, particularly investment in capital equipment, can be skilled-biased due to a high complementarity between these factors. Third, financial sector liberalization could have benefited more large firms

12. Binelli (2012) is among the few, if not the only, explanation relying on the supply side.

Figure 3. Convexity of returns to education



Source: Own calculations based on microdata from household surveys.
Note: Mean of coefficients of squared years of education in Mincer equations, over the period 1990-2009. All coefficients are significant at 1% level.

that demand skilled labor with more intensity. Finally, other potential mechanisms behind this convexification process include tax reforms, labor market reforms, and privatizations.

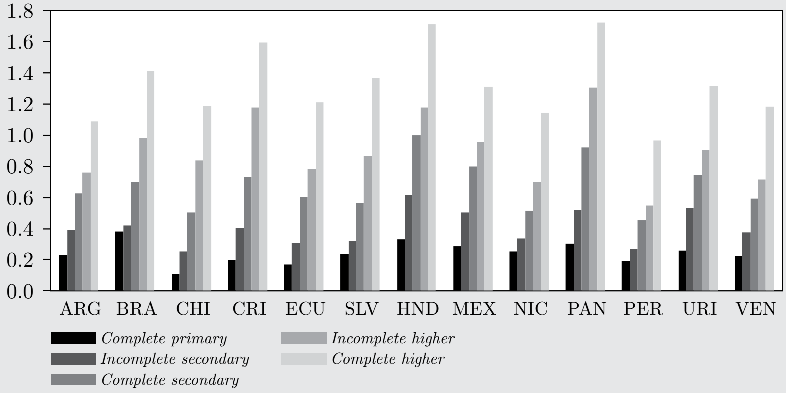
Figure 3 reports the estimated average coefficients of the variable *years of education squared* over the period under analysis. That coefficient is usually taken as a measure of the convexity of returns to education.[13] In all the countries in the sample that coefficient is positive and significant: Convexity in returns to education is a common feature of Latin American labor markets. We come to the same conclusion using the alternative Mincer equation (with dummies for levels of education instead of years of schooling). This is a much more flexible specification and also captures the convexity of returns (see Figure 4).

Non-parametric estimations provide further evidence of the convexity of returns to education.[14] The quadratic approximation is much closer to

13. Given that we separately estimate Mincer equations for head of household, spouse, and other members of the household, we average the coefficients of these regressions for all household members and all periods of analysis. The coefficients are comparable since dependent variables in all Mincer equations are expressed in 2005 PPP dollars and independent variables are homogeneously constructed using SEDLAC definitions.

14. All the non-parametric estimations are available upon request.

Figure 4. Returns to education (in levels)



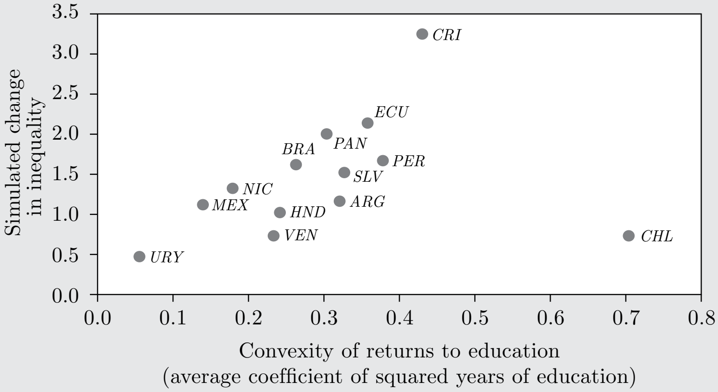
Source: Own calculations based on microdata from household surveys.
Note: Mean of coefficients of dummies for levels of education in Mincer equations. The omitted category is primary incomplete. The mean is over the period 1990-2009 and the average is across head, spouse, and other members of the household.

the non-parametric estimation than the linear approximation. For low years of education, the non-parametric and quadratic approximation estimate flatter returns than the profile estimated under the linear specification. Also, in most cases the returns at the upper tail of the distribution are steeper for the non-parametric/quadratic specification than for the linear one. These facts support the convex-returns hypothesis and also make the quadratic specification a reasonable simplification assumption.

Convexity makes it harder for educational improvements to reduce earnings inequality, according on the theoretical model described in Section 2.[15] Under convex returns, even an unbalanced increase in education in favor of the less educated (or poorer) groups may lead to a surge in earnings inequality. Moreover, the higher the convexity of

15. The convexity condition in our model is defined as the "logarithmic convexity" of returns to education. As noted by Bourguignon *et al.* (2005), if we proportionally increase the years of education of every worker, a stronger condition is required to keep inequality unchanged. In terms of our model, this condition can be denoted as a "strong convexity" of returns to education with respect to earnings (instead of log earnings). Our estimations suggest that in all countries, returns are strongly convex with respect to earnings, meaning that education inequality must drop by a significant amount in order to reduce earnings inequality.

Figure 5. Convexity in returns to education and simulated changes in earnings inequality



Source: Own calculations based on microdata from household surveys.
Note: The vertical axis shows changes in the Gini coefficients as reported in Table 1 for Simulation 2.

the returns, the larger the bias toward the more disadvantaged groups in the education expansion should be in order to reduce inequality in the distribution of earnings (Equation (6)). There is a clear, positive relationship between the convexity of returns to education and the counterfactual changes in earnings inequality, with Chile as the only outlier (Figure 5). The linear correlation coefficient is 0.27 with Chile, and 0.79 without that observation (significant at 1%). This positive relationship suggests that the education expansion during the last two decades brought about a stronger unequalizing effect on the earnings distribution in those countries with higher convexity in returns to education.

As expected from the model in Section 2 and the evidence on convexity, the simulated changes in earnings inequality associated with changes in years of education (Simulation 2) are positively correlated with the change in mean years of education (coefficient equal to 0.29), and the change in education inequality, as measured alternatively by the Gini (0.32), educational gap 1 (0.41), and gap 2 (0.61). The initial gap in years of education is also positively correlated with the simulated change in earnings inequality (coefficients of 0.45 for gap 1 and 0.44 for gap 2).

The theoretical model in Section 2 also shows that changes in returns to education affect earnings inequality. In all countries except Panama, Argentina, and Uruguay, returns to education declined during the period 1990-2009, implying an equalizing impact on the earnings distribution. The effect of the change in structural parameters on inequality is usually defined by the literature as the *parameter* or *price* effect (Bourguignon *et al.*, 2005). Although its estimation is straightforward from the methodology described in Section 3, measuring and discussing this effect is beyond the scope of this paper.

4.3. Characterizing differences across countries

Alternatively, the methodology described above could be applied to assess the extent to which differences in the distribution of years of education across countries can account for the observed differences in labor income inequality. Tables 4.3 and 4.4 report, for each country, the counterfactual change in the Gini coefficient for the earnings distribution simulated by replicating the educational structure of the country in a given column. For instance, if in Argentina we simulate an educational structure similar to that observed in Bolivia, the Gini coefficient for the earnings distribution would be 2.7 points higher than the one that is actually observed. Conversely, inequality would be lower in Argentina if the educational structure were similar to that observed in Costa Rica, Panama, or Uruguay. Similarly to previous simulations, in Table 3 we use completed levels of education as the relevant educational variables, whereas in Table 4 we use years of formal education and its square.

Two opposing patterns are evident in Tables 3 and 4. On the one hand, if the relatively more unequal educational structures of Bolivia or Peru were imposed on other economies, other things being equal, earnings inequality would rise. On the other hand, if the relatively less unequal educational structure of Costa Rica or Uruguay were imposed on other countries, earnings inequality would drop. For the rest of the countries, the results depend on the simulated structure and the definition of the education variable (years or levels).

Table 3. Change in Gini from microeconometric cross-country decompositions

Country	Education structure of country...																
	Arg	**Bol**	**Bra**	**Chl**	**Col**	**Cri**	**Dom**	**Ecu**	**Slv**	**Gtm**	**Mex**	**Nic**	**Pan**	**Pry**	**Per**	**Ury**	**Ven**
Arg	-	2.7	1.8	0.2	0.5	-0.8	2.1	0.7	2.0	0.8	1.2	0.9	-0.4	0.7	1.5	-0.7	0.3
Bol	-0.6	-	-0.3	-1.0	-1.4	-2.1	0.2	-1.2	-0.3	-1.3	-1.3	-1.2	-1.4	-1.4	0.2	-2.2	-1.1
Bra	-1.3	1.0	-	-1.3	-1.0	-3.1	0.1	-1.0	-0.4	-2.0	-0.4	-1.6	-1.9	-1.4	0.1	-2.1	-1.3
Chl	0.4	1.2	0.2	-	-0.1	-2.6	0.5	0.4	-0.9	-2.5	0.4	-1.7	-1.0	-0.4	1.4	-1.4	-0.0
Col	0.0	1.8	1.2	-0.2	-	-4.0	1.9	0.2	-0.1	-2.4	0.7	-0.6	-1.3	-0.5	1.3	-1.9	0.1
Cri	3.0	4.8	3.7	2.7	2.9	-	4.5	3.1	2.7	0.0	3.5	1.7	1.5	2.3	4.6	1.2	3.1
Dom	-0.6	0.2	-0.3	-0.4	-0.9	-4.0	-	-1.0	-1.5	-2.7	-0.7	-1.5	-1.6	-1.3	0.6	-2.7	-0.7
Ecu	0.4	1.1	0.7	0.1	-0.1	-2.0	0.8	-	-0.3	-1.8	-0.0	-1.0	-0.7	-0.4	1.3	-1.3	-0.0
Slv	0.7	0.8	0.3	0.1	-0.2	-3.2	0.9	-0.4	-	-3.7	-0.3	-2.2	-0.8	-1.0	1.7	-1.4	-0.1
Gtm	1.1	3.0	2.3	1.4	1.2	-1.7	2.7	1.0	1.9	-	1.8	0.7	-0.1	1.0	2.3	-0.8	1.1
Mex	-0.2	2.4	1.7	-0.1	-0.2	-1.8	1.9	0.1	1.7	0.1	-	0.4	-1.0	-0.2	0.9	-1.8	-0.2
Nic	0.7	1.4	0.9	0.5	0.4	-1.8	1.3	0.3	0.5	-1.2	0.7	-	-0.5	-0.4	1.3	-1.0	0.4
Pan	0.9	5.2	4.3	0.6	1.8	-0.2	4.7	2.2	4.6	2.7	2.2	3.2	-	2.3	2.4	-0.1	1.4
Pry	0.2	1.6	1.2	0.1	0.1	-1.2	1.5	0.2	0.7	-0.4	0.3	0.3	-0.4	-	1.3	-1.1	0.2
Per	-0.7	0.6	-0.0	-0.9	-0.9	-1.8	0.2	-0.7	-0.2	-1.2	-0.5	-1.0	-1.3	-0.9	-	-1.7	-1.0
Ury	1.5	5.3	4.4	1.7	2.3	0.5	4.6	2.6	4.8	2.7	2.7	2.8	1.0	2.5	3.4	-	1.9
Ven	0.2	1.8	1.2	0.3	0.1	-1.7	1.4	0.1	0.9	-0.5	0.3	0.1	-0.6	-0.0	1.3	-1.2	-

Source: Own calculations based on microdata from household surveys.
Note: Table reports simulated changes in Gini index. Mincer equations estimated using levels of education and Gasparini, Marchionni, and Sosa Escudero (2005) procedure for changing educational structure. Workers between 14 and 65.

Table 4. Change in Gini from microeconometric cross-country decompositions

Country	Education structure of country...																
	Arg	Bol	Bra	Chl	Col	Cri	Dom	Ecu	Slv	Gtm	Mex	Nic	Pan	Pry	Per	Ury	Ven
Arg	-	0.8	-1.0	-0.7	-0.4	-0.6	0.0	0.6	-0.1	-1.8	0.0	-2.1	1.0	-1.0	0.9	-1.0	-0.6
Bol	-1.3	-	-1.2	-2.2	-0.8	-1.2	-0.4	-0.2	-0.3	-1.6	-0.7	-1.9	-0.0	-1.7	-0.0	-1.8	-1.2
Bra	-0.2	2.3	-	-1.2	0.8	0.4	1.1	1.6	1.8	0.3	1.0	-0.7	1.6	-0.3	2.0	-0.9	0.0
Chl	1.4	0.5	-3.0	-	-1.6	-1.7	-0.3	1.3	-1.4	-5.0	-0.5	-5.4	2.4	-2.9	1.8	-2.0	-1.5
Col	0.4	1.8	-1.2	-1.6	-	-0.9	1.3	2.1	0.3	-4.5	0.4	-4.3	3.1	-1.6	2.4	-1.6	-0.4
Cri	1.7	3.2	-0.4	-0.1	1.0	-	2.6	3.2	1.5	-3.9	1.7	-3.5	4.1	-0.4	3.6	-0.6	0.6
Dom	0.8	0.9	-1.2	-0.2	-0.3	-0.5	-	1.2	-0.4	-3.1	0.1	-3.3	2.0	-1.2	1.5	-0.9	-0.5
Ecu	-0.4	-0.2	-1.8	-1.2	-1.1	-1.2	-0.8	-	-1.0	-3.2	-0.8	-3.3	0.5	-1.8	0.3	-1.6	-1.3
Slv	0.8	1.1	-0.7	-0.9	0.0	-0.5	0.8	1.6	-	-3.4	0.0	-3.1	2.7	-1.2	1.8	-0.8	-0.0
Gtm	-0.4	1.8	0.2	-1.3	0.8	-0.0	1.1	1.4	1.6	-	0.7	-0.4	1.2	-0.5	1.6	-1.0	-0.0
Mex	-0.6	1.1	-0.9	-1.5	-0.2	-0.7	0.3	0.7	0.5	-1.3	-	-1.9	0.9	-1.2	1.1	-1.6	-0.8
Nic	2.6	4.1	1.9	1.4	2.7	2.1	3.4	4.0	3.3	0.2	2.9	-	4.6	1.5	4.3	1.3	2.2
Pan	-1.8	0.8	-1.4	-2.8	-0.7	-1.4	-0.2	0.2	0.3	-1.2	-0.5	-2.1	-	-1.8	0.4	-2.6	-1.5
Pry	0.9	1.9	0.4	-0.0	0.9	0.6	1.4	1.7	1.4	-0.3	1.0	-0.6	2.3	-	2.1	0.1	0.6
Per	-0.8	-0.1	-1.6	-1.4	-1.1	-1.2	-0.9	-0.3	-0.8	-2.1	-0.8	-2.4	-0.1	-1.7	-	-1.7	-1.4
Ury	1.0	4.4	1.6	-0.0	2.5	1.5	3.1	3.5	3.8	1.8	2.9	0.8	3.1	1.0	4.0	-	1.4
Ven	0.3	1.5	-0.2	-0.4	0.3	-0.1	0.7	1.2	0.8	-0.9	0.5	-1.1	1.5	-0.2	1.5	-0.5	-

Source: Own calculations based on microdata from household surveys.
Note: Table reports simulated changes in Gini index. Mincer equations estimated using years of education and Legovini, Bouillón, and Lustig (2005) procedure for changing educational structure. Workers between 14 and 65.

5. Projecting the future

In the previous section we discussed how past educational changes influence levels of earnings inequality. It is also interesting to assess how future changes in education could affect inequality. In this section we use microeconometric decompositions to simulate the impact of alternative education expansions on earnings inequality measures. The results again are estimates of the direct, first-round effect of the education expansion.

5.1. Results from microsimulations

We simulate changes in earnings inequality driven by two counterfactual changes in education: an increase of one year of formal education for each worker in the sample (Simulation 3) and a proportional change that raises the average years of education by one year (Simulation 4).

If we assume that returns to education remain constant, the effect of one year more of education for every worker (Simulation 3) is undoubtedly unequalizing in all countries (Table 5). Since the change in education is assumed to be balanced across less- and more-skilled groups, this example illustrates the outstanding role of the convexity of the returns to education. Unsurprisingly, a change in education biased toward more educated groups, like the proportional increase in years of education assumed in Simulation 4, raises earnings inequality in all countries even more than Simulation 3.

Under both simulations the highest counterfactual increases in inequality occur in Chile, Colombia, and Costa Rica, whereas the lowest changes occur in Guatemala, Uruguay, and Paraguay. In fact, there is high positive correlation between the simulated changes in earnings inequality and the estimated convexity of returns to education (Figure 6). The linear correlation coefficient is 0.93. Once more, this shows that the higher the convexity of the returns to education, the greater the unequalizing effect of an education expansion.

5.2. Inequality-reducing education expansions

The results of Simulations 3 and 4 are consistent with the theoretical model: A proportional increase in years of education or even a uniform increase for all workers would result in higher earnings inequality under convex returns. We now examine the conditions under which an increase in education would produce a decline in inequality. With

Table 5. Effect of an extra year of education on earnings inequality

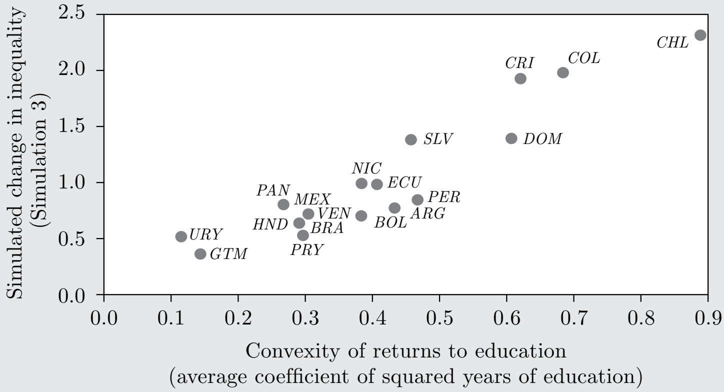
Results from microeconometric decomposition

Country	Year	Observed Gini	Effect of one year more of formal education				Convexity of returns
			Simulation 3		Simulation 4		
			Gini	Δ Gini	Gini	Δ Gini	
Argentina	2009	40.1	40.9	0.8 ***	41.9	1.8 ***	0.44
Bolivia	2005	53.9	54.6	0.7 ***	56.1	2.2 ***	0.38
Brazil	2009	51.1	51.8	0.6 ***	53.4	2.3 ***	0.29
Chile	2009	50.2	52.5	2.3 ***	54.4	4.3 ***	0.89
Colombia	2007	53.4	55.4	2.0 ***	58.6	5.2 ***	0.69
Costa Rica	2009	45.4	47.3	1.9 ***	50.1	4.7 ***	0.62
Dominican Rep.	2009	46.6	48.0	1.4 ***	49.9	3.3 ***	0.61
Ecuador	2009	45.5	46.5	1.0 ***	48.1	2.6 ***	0.41
El Salvador	2008	44.6	45.9	1.4 ***	48.6	4.0 ***	0.46
Guatemala	2006	52.3	52.7	0.4 ***	55.3	3.0 ***	0.15
Honduras	2009	52.0	52.7	0.6 ***	55.1	3.1 ***	0.29
Mexico	2008	49.3	50.1	0.8 ***	51.8	2.5 ***	0.27
Nicaragua	2005	49.4	50.4	1.0 ***	52.8	3.4 ***	0.39
Panama	2009	47.4	48.3	0.8 ***	50.0	2.6 ***	0.27
Paraguay	2009	48.7	49.2	0.5 ***	50.4	1.7 ***	0.30
Peru	2009	50.5	51.4	0.8 ***	52.4	1.9 ***	0.47
Uruguay	2009	47.7	48.3	0.5 ***	49.7	2.0 ***	0.12
Venezuela	2006	37.8	38.5	0.7 ***	39.8	2.0 ***	0.31

Source: Own calculations based on microdata from household surveys.
Note: Simulation 3: One year more of education for each individual in the sample. Simulation 4: Proportional change increasing by one year the mean years of education in the sample. Convexity of returns estimated as the average coefficient (x100) of squared years of education for “head of household”, “spouse” and “other members” Mincer equations. Workers between 14 and 65. Significance levels obtained using 200 bootstrap repetitions.

Figure 6. Convexity of returns to education and simulated change in earnings inequality

(Simulation 3)



Source: Own calculations based on microdata from household surveys.

this aim, we define the following transformation that will be used to simulate an average increase of one year of education (X) under different educational growth paths:

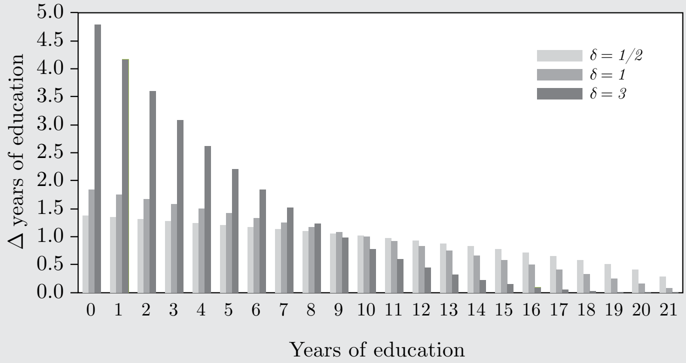
$$X_i^* = X_i + \theta\left(1 - \frac{X_i}{X_{\max}}\right)^{\delta}, \theta > 0 \tag{14}$$

Equation (14) defines the transformation as a function of two exogenous parameters δ and θ. $X_{\max}$ is the highest value of the years of education variable in the sample. The higher the value of parameter δ, the more intense the increase in education for the less educated relative to the more educated.[16] We impose the following restriction:

$$\frac{1}{N}\sum_i X_i + 1 = \frac{1}{N}\sum_i \left[X_i + \theta\left(1 - \frac{X_i}{X_{\max}}\right)^{\delta}\right] \tag{15}$$

16. In particular, values of $\delta > 0$ imply that in absolute terms the increase in years of education is biased toward the less-educated population. Negative values of δ result in a change biased toward the more-educated population.

Figure 7. Changes in years of education using different values of δ, Uruguay



Source: Own calculations based on microdata from household surveys.

Equation (15) restricts the transformation to simulate an average increase of one year of education. When $\delta = 0$, then $\theta = 1$ and the change in educational structure matches Simulation 3. Figure 7 shows the underlying changes in years of education for different values of δ for Uruguay.[17] A value of $\delta = 3$ implies an extremely biased change toward the less educated, whereas $\delta = 1$ and $\delta = 1/2$ are still changes biased toward the less educated population.

Table 6 reports the simulated changes in earnings inequality when the average years of education is increased by one year, assuming different values for δ. Additionally, in order to illustrate how significant is the change produced in the educational structure, for each value of δ we report the change in education inequality by means of the educational Gini and the educational gap between extreme earnings quintiles.

The simulations suggest that in 12 of the 18 countries a value of $\delta > 1/2$ is required to yield an educational expansion that lowers earnings inequality. In some cases, such as the Dominican Republic or El Salvador, a value of $\delta > 1$ is required for this to happen. The requirement of $\delta > 1/2$ is strong, taking into account that $\delta = 0$ implies a uniform change. Therefore, our estimations suggest that even when

17. The rest of the countries present figures very similar to this.

Table 6. Effect of an extra year of education on earnings inequality

Results from microeconometric decomposition

Country	Year	Observed Gini earnings	Effect of one year more of formal education								
			$\delta = 1/2$			$\delta = 1$			$\delta = 3$		
			Earnings	Education		Earnings	Education		Earnings	Education	
			Δ Gini	Δ Gini	Δ Gap	Δ Gini	Δ Gini	Δ Gap	Δ Gini	Δ Gini	Δ Gap
Argentina	2009	40.1	0.0	-2.6	-0.1	-0.4 ***	-3.5	-0.8	-1.1 ***	-6.0	-1.7
Bolivia	2005	53.9	-0.2 **	-5.2	-0.5	-0.4 ***	-6.4	-0.9	-0.7 ***	-9.2	-1.7
Brazil	2009	51.1	-0.1 ***	-5.6	-0.3	-0.6 ***	-6.8	-0.8	-1.4 ***	-10.0	-1.9
Chile	2009	50.2	0.7 ***	-2.4	0.2	-0.1	-3.2	-0.3	-1.0 ***	-5.8	-1.4
Colombia	2007	53.4	0.4 ***	-4.5	0.0	-0.2 ***	-5.6	-0.5	-0.9 ***	-8.5	-1.4
Costa Rica	2009	45.4	0.3 ***	-3.8	-0.2	-0.4 ***	-4.9	-0.6	-1.3 ***	-7.4	-1.4
Dominican Rep.	2009	46.6	0.6 ***	-4.4	0.3	0.2	-5.4	-0.1	-0.3 ***	-8.5	-1.1
Ecuador	2009	45.6	0.2 ***	-4.6	-0.1	-0.2 ***	-5.7	-0.6	-0.6 ***	-8.4	-1.4
El Salvador	2008	44.6	0.5 ***	-5.4	-0.1	0.1	-6.4	-0.4	-0.6 ***	-9.4	-1.3
Guatemala	2006	52.3	-0.2 ***	-8.9	-0.4	-0.6 ***	-10.1	-0.7	-1.1 ***	-13.4	-1.5
Honduras	2009	52.0	0.1 **	-5.6	-0.1	-0.3 ***	-6.6	-0.5	-0.7 ***	-9.6	-1.2
Mexico	2008	49.3	0.1	-4.0	-0.3	-0.4 ***	-4.9	-0.7	-1.0 ***	-7.7	-1.6
Nicaragua	2005	49.4	0.0	-7.5	0.0	-0.2 **	-8.9	-0.4	-0.4 ***	-12.3	-0.9
Panama	2009	47.4	-0.2 ***	-3.3	-0.3	-0.8 ***	-4.2	-0.9	-1.5 ***	-6.8	-1.9
Paraguay	2009	48.7	-0.2 ***	-4.6	-0.6	-0.4 ***	-5.7	-1.0	-0.7 ***	-8.4	-1.7
Peru	2009	50.5	0.0	-6.2	-0.2	-0.2 ***	-7.5	-0.6	-0.4 ***	-10.6	-1.4
Uruguay	2009	47.7	-0.2 ***	-2.8	-0.3	-0.6 ***	-3.6	-0.7	-1.4 ***	-5.8	-1.5
Venezuela	2006	37.8	0.0 *	-4.0	-0.3	-0.3 ***	-5.1	-0.7	-0.8 ***	-7.9	-1.5

Source: Own calculations based on microdata from household surveys.
Note: Gap: Educational gap between earnings quintiles. Workers between 14 and 65. Significance levels obtained using 200 bootstrap repetitions.

the educational expansion is assumed to be biased toward those less-skilled, earnings inequality will rise if the increase in education is not progressive enough.

6. Including the effect of changes in the returns to education

In the previous sections we assumed constant returns to education. The first-order impact results obtained with this assumption can be complemented with a more robust counterfactual that includes the effect of the shift in labor supply composition on returns. At this point, it is difficult to go further without imposing some theoretical structure on the data. Our setting closely follows the approach in Katz and Murphy (1992), Card and Lemieux (2001) and Manacorda, Manning, and Wadsworth (2012).

We assume that firms produce using a typical neoclassical production function that combines labor and capital,

$$Y_t = A_t K_t^{1-\alpha} L_t^{\alpha} \tag{16}$$

Capital is assumed to be exogenous to the firms' hiring decisions. Labor is a composite input that aggregates E different skills or education groups (indexed by e) using a CES technology[18]:

$$L_t = \left[\sum_{e=1}^{E} \theta_{et} L_{et}^{\rho}\right]^{\frac{1}{\rho}} \tag{17}$$

where $\theta_{1t} = 1$ is a normalization for the relative efficiency parameters. We allow this parameter to change across periods to capture differences in education quality of different cohorts or skill-biased technological change. Substitution between different education groups is measured by the elasticity of substitution $\sigma = 1/(1 - \rho)$. For simplicity, we assume that different age/experience groups are perfect substitutes,

18. Similarly to Manacorda, Sanchez-Paramo, and Schady (2010), we assume that there are more than two educational or skill levels but we do not nest multiple levels within the unskilled factor.

although they can differ in efficiency/quality units.[19] The main reason to keep this assumption is to avoid cross-effects between changes in education and returns to experience. Therefore, L_{et} is composed of different experience or age groups indexed by a, that is

$$L_{et} = \sum_{a=1}^{A} \lambda_{ea} L_{eat} \tag{18}$$

The relative efficiency parameter λ_{ea} is assumed to be time invariant and λ_{e1} is normalized to one. Under the assumption of competitive markets and normalizing the price of output to one, log wages are given by:

$$\begin{aligned} \log w_{eat} = {} & \log\alpha + \log Y_t + \log\theta_{et} + \log\lambda_{ea} - \rho\log L_t \\ & + (\rho - 1)\log L_{et} \end{aligned} \tag{19}$$

Equation (19) cannot be directly estimated by OLS since, for instance, we need an estimation of ρ and θ_{et} in order to measure L_t. Nevertheless, consider the following specification based on (19) where d_t denotes time fixed effects, $(d_t \times d_e)$denotes interactions between time and education dummies, and $(d_t \times d_a)$ denotes interactions between education and age dummies:

$$\log w_{eat} = const + d_t + (d_t \times d_e) + (d_e \times d_a) + \in_{eat} \tag{20}$$

The time fixed effects absorb $\log Y_t - \rho\log L_t$, the time-education interactions absorb $\log\theta_{et} + (\rho - 1)\log L_{et}$ and education-age interactions identify $\log\lambda_{ea}$.[20] Consider also the wage gap relative to the lowest educational group:

$$\log \frac{w_{eat}}{w_{1at}} = \log\theta_{et} + \log\lambda_{ea} + (\rho - 1)\log\frac{L_{et}}{L_{1t}} \tag{21}$$

19. This could be extended in the same direction as Card and Lemieux (2001) to account for more interesting cohort effects. With a different approach, Sapelli (2007) performs an interesting analysis using synthetic cohorts, constructed from successive cross-section surveys, to study the evolution of income distribution in Chile. In particular, the author analyzes whether the pattern of cohort effects can be explained by trends in the mean and dispersion of years of education and returns to education within the cohort.

20. Identification of λ_{ea} allows us to estimate L_{et} from (18).

Equation (21) cannot be implemented empirically without further assumptions because L_{et} varies at the same level as the unobserved $\log\theta_{et}$. Instead of imposing a linear trend like Card and Lemieux (2001) we allow $\log\theta_{et}$ to vary additively in e and t, i.e., $\theta_{et} = f_e + f_t$ where f_e and f_t are education and time dummy variables. Then, the estimable version of equation (21) is given by[21]:

$$\log\frac{w_{eat}}{w_{1at}} = d_e + d_t + \left(d_e \times d_a\right) + (\rho - 1)\log\frac{L_{et}}{L_{1t}} + \in_{at} \tag{22}$$

Changes in the educational composition of the population affect returns to schooling according to the elasticities of substitution embedded in labor demand. Given the assumptions of the model, it is easy to see that the elasticity of $\log(w_{eat}/w_{1at})$with respect to the relative labor supply L_{et}/L_{1t} is given by:

$$\eta_{eat} = \frac{(\rho - 1)}{\log\left({}^{w_{eat}}\!/\!_{w_{1at}}\right)} \tag{23}$$

If we focus on the Mincer equation specified on educational levels (where the estimated returns are interpreted as log-differences with the lowest level), we can use (21) to simulate the percentage variation of the Mincerian coefficients in response to the counterfactual educational changes.[22] Since the elasticity is not constant over time, we use the baseline year elasticity in order to simulate the changes during the subsequent period. For example, we use $\eta_{eat=1990}$ to simulate the changes in the Mincerian coefficients of educational levels ($e = 1,...,6$) when the educational structure of the year 2009 is replicated in the baseline surveys of 1990.[23] Table A.1 in the appendix shows the estimated ρ,

21. Identification of ρ relies on the exogenous variation of L_{et}/L_{1t}. Since labor participation decisions are endogenous to the wage level, this ratio can change in response to demand shocks. To overcome this drawback, we instrument the ratio with N_{et}/N_{1t} where N_{et} is the size of the population at the cell (e,t) regardless of the activity status. The estimations use the hourly wage instead of total earnings, to hold to a closer proxy of worker productivity. Finally, we weight regressions by the number of workers in each cell (e,a,t) and cluster standard errors at the (e,a) cell level.

22. Note that from equation (21) it follows that $(\log w_{eat} - \log w_{1at})$ does not depend on the supply of levels of education other than e and 1. This is because the CES technology implies that changes in L_{jt} with $j \neq e \neq 1$ proportionally affect w_{eat} and w_{1at}.

23. In order to calculate the percentage change in returns we need to use the structural parameters to estimate L_{et}/L_{1t} and its change.

Table 7. Simulated inequality change after accounting for changes in returns to education

Country	Period	No changes in returns		Changes in returns	
		Simulation 1	Simulation 2	Simulation 1	Simulation 2
Argentina	1992-2009	0,2 ***	1,2 ***	0,0	1,0 ***
Brazil	1992-2009	1,0 ***	1,6 ***	0,0	0,3 ***
Chile	1990-2009	0,6 ***	0,7 ***	-0,2 ***	-0,2 ***
Costa Rica	1990-2009	0,9 ***	3,2 ***	0,8 ***	3,2 ***
Ecuador	1994-2009	0,4 ***	2,1 ***	-0,6 ***	1,5 ***
El Salvador	1995-2008	2,5 ***	1,5 ***	1,5 ***	0,2 ***
Honduras	1995-2009	1,7 ***	1,0 ***	1,7 ***	0,9 ***
Mexico	1989-2008	0,6 ***	1,1 ***	0,4 ***	0,9 ***
Nicaragua	1993-2005	0,9 ***	1,3 ***	0,6 ***	1,1 ***
Panama	1991-2009	0,2 ***	2,0 ***	0,2 ***	2,0 ***
Peru	1997-2009	-0,0	1,7 ***	-0,2 ***	1,4 ***
Uruguay	1992-2009	-0,9 ***	0,5 ***	-0,9 ***	0,5 ***
Venezuela	1992-2006	0,6 ***	0,7 ***	-0,7 ***	-1,0 ***
Average		*0,7*	*1,4*	*0,2*	*0,9*

Source: Own calculations based on microdata from household surveys.
Note: Simulation 1 follows Gasparini *et al.* (2005) to change the educational structure, while Simulation 2 follows Legovini *et al.* (2005). Returns to education are adjusted according to the elasticities estimated from the structural demand model. (*) Significance levels obtained using 200 bootstrap repetitions.

the average (across age groups) η_{et} and the average change in returns to each level of education relative to the incomplete primary level.

The procedure described above allows simulation of the impact of an expansion in education measured in levels (as in Simulation 1 in Table 1). To implement the procedure when education is measured in years of schooling (as in Simulation 2 in Table 1), we proceed as follows: We generate counterfactual earnings using returns to levels of education adjusted by the changes predicted by the CES model and we keep other variables and residuals unchanged for each individual. Then, using these counterfactual earnings we re-estimate the quadratic-Mincer equation, and substitute the constant linear and quadratic coefficients from this regression in the original Mincer equation. Figure A.1 in the appendix illustrates how the quadratic profile of returns to education is adjusted in this simulation for a subset of countries.

Table 7 reports the results of the microsimulations in Table 1 along with those obtained when taking into account changes in returns resulting from the estimated elasticities, for the two simulations discussed above (years of schooling and educational levels). Table 7 suggests that, on average, the estimated impact of the education expansion remains unequalizing when allowing for changes in the returns to schooling, but it becomes smaller. For instance, under Simulation 2, while the mean increase in the Gini is 1.4 points when ignoring changes in returns to schooling, the estimated impact is reduced to 0.9 points when including this effect. For all countries except Chile and Venezuela, inequality rises in response to the simulated changes in years of education, although less than in the partial equilibrium estimates.[24] For Simulation 1 the mean impact is reduced from 0.7 to 0.2 Gini points, but it is still positive (i.e., unequalizing). Under this simulation in some countries changes in returns compensate the initial unequalizing effect of education, keeping inequality unchanged (Argentina and Brazil), while for others the overall impact of the education expansion becomes equalizing (Chile, Ecuador, Peru, and Venezuela).

7. Concluding remarks

We find that the direct effect of the increase in education experienced by Latin American countries in the last two decades was unequalizing, and that according to the projected scenarios, this result is expected to hold for future improvements in education if they are not strongly biased toward the less-educated population. Both facts are closely linked to the convexity of returns to education. With convex returns, even a progressive change in education may lead to a more unequal distribution of earnings and hence to a more unequal income distribution. This paper shows that this is not merely a theoretical possibility with little relevance in practice, but that it is a widespread phenomenon across Latin American economies. We also find that, on average, the estimated impact of the education expansion remains unequalizing when allowing for changes in returns to schooling, although the effect becomes smaller.

24. There is a concern regarding the ability of the quadratic specification to accurately fit the counterfactual earnings profile if returns to level of education change in a non-trivial way. For this reason we recalculated Simulation 2 using a third-order polynomial (not reported) but we found no significant differences with results obtained under the quadratic specification. Indeed, the education effect is unequalizing for the same set of countries.

Of course, showing that under certain circumstances an increase in education may be linked in the short run to a growth in income inequality does not lead to the conclusion that investment in education should be reduced, as an education expansion has many positive implications for growth, equality of opportunity, mobility, and poverty reduction, among others. Moreover, if improvements in schooling are progressive enough, earnings inequality will fall after an education expansion.

REFERENCES

Alejo, J. (2012), "La paradoja del progreso: ¿convexidad o heterogeneidad de los retornos a la educación?" Working Paper, CEDLAS-UNLP.

Almeida dos Reis, J. and R. Paes de Barros (1991), "Wage inequality and the distribution of education: A study of the evolution of regional differences in inequality in metropolitan Brazil," *Journal of Development Economics* **36**: 117–43.

Angrist, J.D. and A.B. Krueger (1991), "Does compulsory school attendance affect schooling and earnings?" *Quarterly Journal of Economics* **106**(4): 979–1014.

Attanasio, O., P.K. Goldberg, and N. Pavcnik (2004), "Trade reforms and wage inequality in Colombia," *Journal of Development Economics*, **74**(2): 331–66.

Behrman, J., N. Birdsall and M. Székely (2003), "Economic policy and wage differentials in Latin America," Working Paper No. 29, Center for Global Development, Washington, D.C.

Benhabib, J. and M. Spiegel (1994), "The role of human capital in economic development: Evidence from aggregate cross-country data," *Journal of Monetary Economics* **34**: 143–79.

Binelli, C. (2008), "Returns to education and increasing wage inequality in Latin America," Working Paper Series 30-08, Rimini Centre for Economic Analysis, revised Jan. 2008.

Binelli, C. (2012), "How the wage-education profile got more convex: Evidence from Mexico," unpublished.

Blom, A., L. Holm-Nielsen, and D. Verner (2001), "Education, earnings and inequality in Brazil, 1982-1998," Policy Research Working Paper No. 2686. World Bank.

Bourguignon, F. and F. Ferreira (2005), "Decomposing changes in the distribution of household incomes: Methodological aspects," in Bourguignon, F., F. Ferreira and N. Lustig, eds., *The microeconomics of income distribution dynamics in East Asia and Latin America.* Washington, D.C.: Oxford University Press and the World Bank.

Bourguignon, F., F. Ferreira, and N. Lustig (2005), *The microeconomics of income distribution dynamics in East Asia and Latin America.* Washington, D.C.: Oxford University Press and the World Bank.

Card, D. (1999), "The causal effect of education on earnings," in: Ashenfelter, O. and D. Card, eds., *Handbook of Labor Economics*, Edition 1, Volume 3,Chapter 30: 1801–63.

Card, D. and T. Lemieux (2001), "Can falling supply explain the rising return to college for younger men? A cohort-based analysis," *The Quarterly Journal of Economics* **116**(2): 705–46.

Cornia, G. (2011), "Economic integration, inequality and growth: Latin America vs. the European economies in transition," *Review of Economics and Institutions* **2**(2).

Deschênes, O. (2001), "Unobserved ability, comparative advantage and the rising return to education in the United States: A cohort-based approach," Working Paper No. 465, Princeton University Industrial Relations Section.

Esquivel, G. and J.A. Rodríguez-López (2003), "Technology, trade, and wage inequality in Mexico before and after NAFTA," *Journal of Development Economics* **72**: 543–65.

Feenstra, R. and G. Hanson (2001), "Global production sharing and rising inequality: A survey of trade and wages," NBER Working Paper 8372.

Gasparini, L., M. Marchionni, and W. Sosa Escudero (2005), "Characterization of inequality changes through microeconometric decompositions: The case of Greater Buenos Aires," in Bourguignon, F., F. Ferreira and N. Lustig, eds., *The microeconomics of income distribution dynamics in East Asia and Latin America*. Washington, D.C.: Oxford University Press and the World Bank.

Gasparini, L., S. Galiani, G. Cruces, and P. Acosta (2011), "Educational upgrading and returns to skills in Latin America: Evidence from a supply-demand framework, 1990-2010," IZA Discussion Paper 6244 and Policy Research Working Paper WPS 5921, the World Bank.

Gasparini, L. and N. Lustig (2011), "The rise and fall of income inequality in Latin America," In Ocampo, J.A. and J. Ros, eds., *The Oxford Handbook of Latin American Economics*. Oxford: Oxford University Press.

Katz, L. F. and K.M. Murphy (1992), "Changes in relative wages, 1963–87: Supply and demand factors," *Quarterly Journal of Economics* **107**: 35–78.

Knight, J. B. and R.H. Sabot (1983), "Educational expansion and the Kuznets effect," *American Economic Review* **73**(5): 1132–36.

Kolm, S. (1977), "Multidimensional equalitarianisms," *The Quarterly Journal of Economics* **91**: 1–13.

Krueger, A.B. and M. Lindahl (2001), "Education for growth: Why and for whom," *Journal of Economic Literature* **39**(4): 1101–36.

Lam, D. (1999), "Generating extreme inequality: Schooling, earnings, and intergenerational transmission of human capital in South Africa and Brazil," Research Report 99-439, Population Studies Center, University of Michigan, Ann Arbor.

Langoni, C.G. (1973), "Distribuição da renda e desenvolvimento econômico do Brasil," Expressão e Cultura, Rio de Janeiro.

Legovini, A., C. Bouillón, and N. Lustig (2005), "Can education explain changes in income inequality in Mexico?" In Bourguignon, F., F. Ferreira and N. Lustig, eds., *The microeconomics of income distribution dynamics in East Asia and Latin America*. Washington, D.C.: Oxford University Press and the World Bank.

Lemieux, T. (2006), "Postsecondary education and increasing wage inequality," *American Economic Review* **96**(2): 195–9.

López Calva, L.N. and Lustig, eds., (2010), *Declining inequality in Latin America: A decade of progress?* Brookings Institution and UNDP.

Manacorda, M., A. Manning, and J. Wadsworth (2012), "The impact of immigration on the structure of wages: Theory and evidence from Britain," *Journal of the European Economic Association* **10**(1): 120–51.

Manacorda, M., C. Sanchez-Paramo, and N.R. Schady (2010), "Changes in returns to education in Latin America: The role of demand and supply of skills," *Industrial and Labour Relations Review* **63**(2): 307–26.

Martins, P.S. and P.T. Pereira (2004), "Does education reduce wage inequality? Quantile regression evidence for 16 countries," *Labour Economics* **11**(3), 355–71.

Mehta, A., J. Felipe, P. Quising, and S. Camingue (2013), "Where have all the educated workers gone? Education and structural transformation in three Asian economies," *Metroeconomica* **64**(3): 401–572.

Mincer, J.A. (1974), *Schooling, experience, and earnings*. Cambridge: NBER Books.

Mincer, J. (1997), "Changes in wage inequality, 1970-1990," *Research in Labor Economics* **16**: 1–18.

Pavcnik, N., A. Blom, P. Goldberg, and N. Schady (2005), "Trade liberalization and industry wage structure: Evidence from Brazil," *World Bank Economic Review* **18**(3): 319–44.

Pritchett, L. (2001), "Where has all the education gone?" *World Bank Economic Review* **15**: 367–91.

Rau, T. (2013), "Modeling structural equations with endogenous regressors and heterogeneity through derivative constraints," *Quantitative Economics* **4**(1): 125–48.

Revenga, A. (1997), "Employment and wage effects of trade liberalization: The case of Mexican manufacturing," *Journal of Labor Economics* **15**(3): S20–S43.

Reyes, A. (1988), "Evolución de la distribución del ingreso en Colombia," *Desarrollo y Sociedad* **21**: 39–51.

Sapelli, C. (2007), "A cohort analysis of the income distribution in Chile," Workign Paper Nº 290, Instituto de Economia, Pontificia Universidad Católica de Chile.

Savanti, M. P. and H. Patrinos (2005), "Rising returns to schooling in Argentina, 1992–2002: Productivity or credentialism?" World Bank Policy Research Working Paper.

SEDLAC (CEDLAS and the World Bank) (2011), Socio-Economic Database for Latin America and the Caribbean. CEDLAS-Universidad Nacional de La Plata and the World Bank (LCSPP). Retrieved from sedlac.econo.unlp.edu.ar.

Soderbom, M., F. Teal, A. Wambugu, and G. Kahyarara (2006), "The dynamics of returns to education in Kenyan and Tanzanian manufacturing," *Oxford Bulletin of Economics and Statistics* **68**(3): 261–88.

Soto, M. (2002), "Rediscovering education in growth regressions," OECD Development Centre Working Paper 202.

Spilimbergo, A., J.L. Londoño, and M. Székely (1999), "Income distribution, factor endowments, and trade openness," *Journal of Development Economics* **59**: 77–101.

APPENDIX

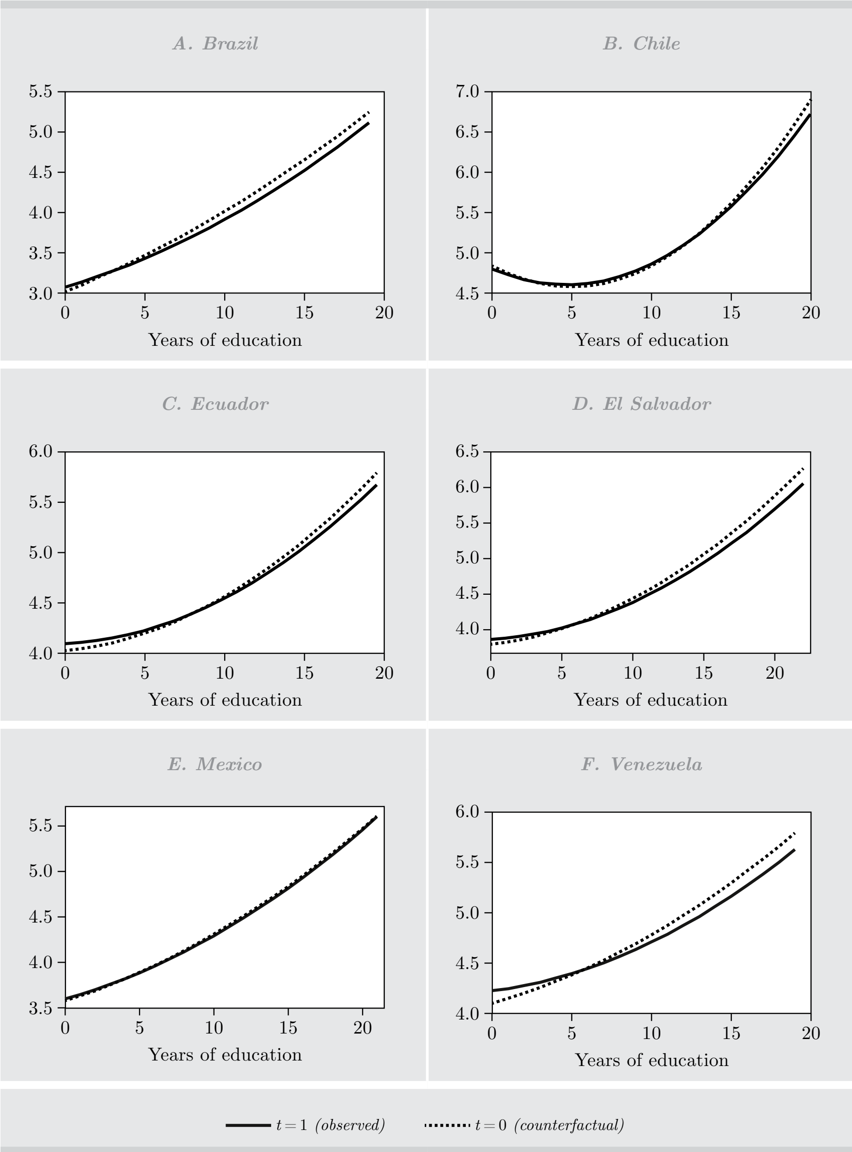
Table A1. Estimated substitution parameter, average elasticities and average change (%) in returns to each educational level predicted by the CES model

	Arg	Bra	Chi	Cri	Ecu	Slv	Hnd	Mex	Nic	Pan	Per	Ury	Ven
ρ	0.98	0.83	0.86	0.98	0.83	0.69	0.92	0.97	0.67	1.00	0.85	1.00	0.72
$se(\rho)$	0.06	0.02	0.04	0.04	0.07	0.04	0.07	0.05	0.12	0.09	0.09	0.04	0.03
$\eta_{e\,=\,\text{pric},t\,=\,1}$	-0.27	-0.29	0.06	-0.07	0.63	-0.65	-0.44	-0.15	-0.85	0.00	-0.44	-0.03	-1.41
$\eta_{e\,=\,\text{pric},t\,=\,3}$	-0.06	-0.57	-0.39	-0.01	-1.30	-1.23	-0.35	0.12	-2.72	0.00	-1.43	-0.06	-0.32
$\eta_{e\,=\,\text{seci},t\,=\,1}$	-0.10	-0.24	-0.19	-0.05	-0.38	-0.57	-0.05	-0.07	-0.51	0.00	-0.55	-0.01	-0.77
$\eta_{e\,=\,\text{seci},t\,=\,3}$	-0.06	-0.55	-0.69	-0.08	0.71	-0.75	-0.20	-0.09	-0.63	0.00	-0.51	-0.01	-0.97
$\eta_{e\,=\,\text{secc},t\,=\,1}$	-0.05	-0.19	-0.31	-0.03	-0.32	0.42	-0.12	-0.04	-0.42	0.00	-0.28	-0.02	-0.44
$\eta_{e\,=\,\text{secc},t\,=\,3}$	0.03	-0.27	-0.21	-0.03	-0.45	-0.57	-0.13	-0.04	-0.61	0.00	-0.32	-0.01	-0.69
$\eta_{e\,=\,\text{supi},t\,=\,1}$	-0.03	-0.15	-0.14	-0.01	-0.23	-0.40	-0.09	-0.03	-0.29	0.00	0.01	-0.01	-0.34
$\eta_{e\,=\,\text{supi},t\,=\,3}$	-0.04	-0.16	-0.19	-0.02	-0.30	-0.31	-0.10	-0.03	-0.48	0.00	-0.18	-0.00	-0.58
$\eta_{e\,=\,\text{supc},t\,=\,1}$	-0.03	-0.12	-0.10	-0.01	-0.12	-0.22	-0.06	-0.02	-0.41	0.00	-0.15	-0.00	-0.23
$\eta_{e\,=\,\text{supc},t\,=\,3}$	-0.02	-0.13	-0.20	-0.01	0.17	-0.19	-0.06	-0.02	-0.22	0.00	-0.15	-0.00	-0.31
$\Delta\%\ \beta$pric	10.7	-10.4	-45.8	1.2	-54.7	-19.8	-3.8	9.1	-9.4	0.0	-7.0	0.8	25.0
$\Delta\%\ \beta$seci	0.8	-35.7	-5.9	-2.7	31.7	-35.6	-5.4	-3.8	2.7	0.0	-7.8	0.2	3.8
$\Delta\%\ \beta$secc	2.3	-20.1	-3.5	-0.1	-22.9	-17.4	1.2	-0.7	-6.6	0.0	-4.7	0.6	-19.0
$\Delta\%\ \beta$supi	-1.6	-13.5	-13.1	-0.7	-22.0	-11.1	2.1	-1.4	-10.9	0.0	-5.7	0.1	-10.5
$\Delta\%\ \beta$supc	-0.8	-5.7	-6.7	-0.2	-5.7	-8.3	-2.1	-0.6	-1.8	0.0	-3.0	0.1	-10.8

Source: Own calculations based on microdata from household surveys.
Note: Changes in coefficients are relative to the incomplete primary level. Elasticities are averaged across age groups. Changes in returns to education are averaged across simulation and averaged across heads and other members of the household. Period $t = 1$ corresponds to surveys before 1995 and period $t = 3$ to surveys after 2004. Education levels are: pric = complete primary, seci = incomplete secondary, secc = complete secondary, supi = incomplete higher, supc = complete higher. Estimation of ρ follows the IV procedure described in the text; each cell is weighted by its population size and standard errors are clustered at the age-education group level.

Figure A1. Simulated changes in returns to education

(Simulation 2)



Source: Own calculations based on microdata from household surveys.
Note: Figure shows the simulated changes in returns to education measured by the linear and quadratic coefficient of the Mincer equation. Simulated changes follow from re-estimating the Mincer equation with counterfactual earnings obtained after adjusting returns to levels of education according to CES model predictions. The counterfactual scenario t=0 follows from simulating the initial period educational structure on the population of the final period (including the coefficient adjustment).

Vol. 51 No. 1 (May, 2014), 41–83
LAJE LATIN AMERICAN JOURNAL OF ECONOMICS

INFLATION TARGETING AND AN OPTIMAL TAYLOR RULE FOR AN OPEN ECONOMY: EVIDENCE FOR COLOMBIA 1990-2011*

Edgar Villa**
Martha A. Misas***
Andrés F. Giraldo****

An optimal monetary policy Taylor rule is developed for an open economy, which we then estimate following a Markov regime-switching model for quarterly data from Colombia during 1990-2011. We find two opposite monetary regimes characterized by different policy rules: until October 2000 the Central Bank of Colombia reacted only statistically to output gap changes while after October 2000, when inflation targeting was officially adopted, monetary policy reacted only statistically to changes in the inflation rate. The latter regime is consistent with the Taylor principle as shown analytically and verified empirically by a unit root test for a Markov regime-switching model.

JEL classification: C24, E42, E52, E58, N16

Keywords: Optimal Taylor rule, inflation targeting, Taylor principle, Markov switching

1. Introduction

The econometric assessment of monetary rules such as Taylor's rule (1993) has become an important field of study in the modern monetary policy literature, as that paper stimulated a series of theoretical and empirical studies with varying objectives.[1] Issues such as the optimality, robustness, performance, and efficiency of monetary rules are the main focus of this research agenda. Conducting monetary policy according to simple rules that help achieve policy objectives, such as inflation targeting, has become

* The authors are grateful for comments made on an early version of this paper by Hernando Vargas, Andrés González, Franz Hamann and Andrés Velasco.

** Corresponding author. Associate Professor, International School of Economic and Administrative Sciences at Universidad de La Sabana, Campus del Puente Común, Chía-Colombia. Email: edgar.villa@unisabana.edu.co.

*** Associate Professor in the Economics Department of Pontificia Universidad Javeriana Bogotá, Edificio Gabriel Giraldo, piso 7mo, Bogotá-Colombia.

**** Assistant Professor in the Economics Department of Pontificia Universidad Javeriana Bogotá, Edificio Gabriel Giraldo, piso 7mo, Bogotá-Colombia.

1. Taylor (1999) presents various studies that provide empirical evidence on the efficiency and robustness of different monetary rules when they are used to guide the decisions of the monetary authority.

doi 10.7764/LAJE.51.1.41

a crucial issue for price and macroeconomic stability, especially in developing economies, as it increases the transparency and credibility of the monetary authority (Woodford, 2003a). This prevents time inconsistencies in monetary policy rules (Kydland and Prescott, 1977; Barro and Gordon, 1983) and also provides transparency for the public (Svensson, 2000, 2008 and 2011).

Empirical assessment of monetary rules is useful from both the operational and academic points of view. In terms of operations, it provides a method for describing the behavior of central banks "as if" they use monetary rules, while from an academic point of view, it helps in analyzing the optimal behavior of these monetary rules. A key concept that emerges in the literature starting with Taylor (1993) is the so-called Taylor principle. According to this principle, the monetary authority increases the intervention interest rate more than proportionally to inflationary pressures, in order to lower inflation expectations and reach a long-term inflation target.[2] Murray *et al.* (2008) have argued that a central bank that systematically satisfies the Taylor principle generates a stationary inflation rate as a time series; in other words, it should not exhibit a stochastic trend or a unit root behavior. Thus, the literature has linked inflation targeting with simple monetary rules in which a central bank changes the intervention interest rate of an economy according to inflation deviations from a target as an appropriate monetary intervention mechanism (Clarida *et al.*, 1998, 1999, 2000)[3] which should satisfy the so-called Taylor principle.

It is important to note that Taylor (1993) presents his rule as a simple ad hoc rule, in the sense that it is not derived from a macroeconomic optimization model. However, subsequent papers deduce monetary rules in the style of Taylor based on the development of a general model in which agents and the monetary authority are optimizing[4]. The purpose of this paper is to find a monetary rule, in the style of Taylor, that is derived from an optimization problem of a central bank in an open economy, where it takes into account deviations of inflation

2. See Walsh (2010) and Woodford (2003a).

3. According to Mishkin (1999), central banks have followed different monetary strategies whose common denominator is the definition of a nominal anchor to "tie" agents' inflation expectations. The strategies most widely used by central banks are: 1) monetary aggregate regimes, whose nominal anchor is monetary supply growth (called monetarism); 2) exchange rate regimes, in which the nominal anchor is the exchange rate chosen within an exchange rate band; and, 3) a target inflation rate, where the anchor is the inflation forecast (Svensson and Woodford, 2003) or the medium- to long-term inflation target.

4. Orphanides (2008) presents some functional forms based on the Taylor rule.

from its target, economic cycles represented by the output gap, as well as the desire to smooth out the intervention rate and the deviation of the real exchange rate relative to a long-term target. The structure of the theoretical model developed is based on traditional foundations in the literature: a quadratic central bank loss function, a Phillips curve and an aggregate demand equation[5]. The optimal intervention rule for an open economy is found through an optimization procedure "as if" followed by the bank. A key analytical result we find is that the optimal Taylor rule for an open economy is satisfied only if the inflation rate is not a highly persistent time series or does not exhibit stochastic trend (unit root) behavior.

Once the optimal Taylor rule is obtained in its reduced form, it leads to an econometric model that is estimated as a Markov-switching model for Colombia between 1990 and 2011. This methodology seems appropriate since the Bank of the Republic (BR) of Colombia adopted an explicit inflation targeting policy halfway through this period, in October 2000. Importantly, the empirical methodology allows estimation of whether there was a monetary policy regime switch in the BR's behavior, consistent with a change in the parameters of the optimal Taylor rule with respect to the inflation rate and/or the output gap due to the explicit adoption of the inflation targeting monetary regime in October 2000, without having to specify exogenously the date of the possible structural regime switch. Empirical evidence is found that there were two monetary policy states or regimes in which the BR's behavior, in terms of the optimal Taylor rule, is the opposite of the way in which the intervention interest rate reacted to changes in the inflation rate and the output gap. In one regime, which prevailed prior to October 2000 and is labeled the *regime with no inflation targeting* (NIT), the BR maintained a policy in which the intervention interest rate reacted statistically solely to changes in the output gap; in the other regime, which prevailed after October 2000 and is labeled the *regime with inflation targeting* (IT), the intervention interest rate reacted statistically only to changes in the inflation rate.

Statistical evidence is also found that the Taylor principle was satisfied only in the IT regime. According to our analytical results, which confirm analytically Murray *et al.*'s (2008) insight, this implies that

5. Although a Phillips curve and an aggregate demand curve are not generated from an optimization exercise within our framework, it is possible to demonstrate that they come from the idea of consumer and firm maximization. See Walsh (2010) and Woodford (2003b).

in the NIT regime the inflation rate could be a time series with a stochastic trend or unit root while in the IT regime, it should be a stationary time series. Consequently, a unit root test is performed on the inflation rate series for the time period under study, also within a Markov regime-switching model, to indirectly verify whether the Taylor principle is satisfied, consistent with the inflation-targeting policy instituted after October 2000. Empirical evidence is found that the inflation rate is a time series with a unit root in the NIT regime that prevailed before October 2000, while it is stationary in the post-October 2000 IT regime. This indirectly verifies that adoption of the IT monetary policy regime produced a structural change in the way the BR conducts monetary policy after 2000. Our combined results support the idea that an inflation targeting policy is beneficial for small countries like Colombia because it allows central banks to control inflationary shocks, steadily maintaining the inflation rate at low levels in a credible way.

This article provides two original contributions to the literature regarding monetary rules. First, it develops an optimal monetary rule in the style of Taylor for an open economy that takes into account both deviations of inflation from its target and economic cycles represented by the output gap, as well as the desire to smooth out the intervention rate and the deviation of the real exchange rate relative to a long-term target. This is an important contribution because most of the empirical literature is based on an ad hoc monetary rule that is not necessarily deduced optimally in a theoretical framework consistent with the interests of central banks. In addition, the theoretical literature on monetary rules usually focuses on rules for closed economies, considering countries such as the United States. This does not appear to be suitable for understanding the monetary policy of small and open economies such as Colombia. The key result is that we show that the Taylor principle cannot be satisfied in the optimal Taylor rule for an open economy if the inflation rate is a highly persistent time series displaying a unit root or stochastic trend.

The second contribution is that the theoretical optimal monetary rule is estimated using a Markov-switching methodology for an underdeveloped economy like Colombia. This is an important contribution, especially for Latin America, because this methodology is rarely used in the literature for countries in this part of the world. Exceptions to this trend are Kuzin (2004) which studies Germany's Bundesbank and Murray *et al.* (2008) which studies the U.S. Federal Reserve also using a Markov-switching methodology. More importantly we believe this methodology models regime switches endogenously, which seems more suitable than the typical methodology of imposing a regime switch exogenously as in Orphanides (2000).

The article is organized as follows: In Section 2 we present a review of the literature on monetary rules, while Section 3 develops the theoretical model that leads to the optimal Taylor rule for an open economy. Section 4 presents a Markov-switching econometric model that is estimated for the optimal Taylor rule developed. Section 5 contains a description of the information used in the econometric model. Section 6 provides the empirical results, while Section 7 concludes.

2. Literature review

The appearance of the concept of rational expectations and its subsequent use in macroeconomic theory and policy changed how the formation of agent expectations is modeled. Kydland and Prescott (1977) showed theoretically the presence of an inflationary bias due to the existence of dynamically inconsistent economic policies. Subsequent to their work, the main theoretical concern has focused on how to eliminate that inflationary bias through monetary rules. Increasing importance has been given to ideas about the independence of the monetary authority, reputation-building in dynamic models, following transparent monetary rules, and the credibility of monetary policy announcements to prevent dynamic inconsistencies in discretionary monetary policy.[6] The work of Taylor (1993) is important to monetary policy literature because it makes operational the theoretical debate on rules versus discretion. In that seminal work, the author provides a first approximation to systematic behavior of monetary policy, using a simple, ad hoc monetary policy rule aimed at approaching the real behavior of the Federal Reserve's intervention interest rate "as if" the Fed actually followed that rule.[7] Specifically, Taylor presents the rule as follows:

$$i = \pi + gx + h(\pi - \pi^{*}) + r \tag{1}$$

6. For more details on this topic, a good reference is Walsh (2000).

7. Traditionally, assessment of monetary policy has implicitly assumed rules of behavior set in the monetary instruments, such as a fixed growth rate of monetary supply (Friedman, 1968). However, until Taylor (1993) monetary rules were not viewed as reaction functions of monetary instruments. The main reasons are that it is difficult for a central bank to follow mechanical rules of behavior and it is also quite implausible that a central bank would maintain a fixed instrument for a certain period of time.

where i is the short-term interest rate, π is the inflation rate expressed as the percentage change in the price index, π^* is the inflation target, x is the percentage deviation of real output from potential output, g is the interest rate response to changes in the output gap, $1+h$ is the response to changes in inflation relative to the target and r is the real interest rate. The values that characterize the Federal Reserve's behavior according to Taylor are the following: $g=0.5$, $h=0.5$, $\pi^*=2$ and $r=2$.[8] The Taylor rule, initially established in an ad hoc fashion and which is shown as Equation (1), is used to closely describe the trajectory of the U.S. federal funds rate for the period 1987-1992.[9]

Later studies attempt to show that the rule is also optimal in the framework of an equilibrium macroeconomic model,[10] and specifically, they have delved deeper into the design of monetary rules, in order to analyze their performance in terms of efficiency, robustness, and optimality.[11] It is worth mentioning that the Taylor rule has consistently been the way of thinking behind the strategy of inflation targeting.[12] The differences in the specification of the various rules in the style of Taylor are mainly associated with whether the inflation gap is measured using the inflation target or inflation expectations[13] and also with the variables included in the equation used to determine the value of the instrument.

Orphanides (2008) has argued, in addition to the original Taylor rule, that other types of reaction functions have been designed that are inspired by Taylor's exercise (1993). The common characteristic of these rules is the response (optimal in many cases) of the interest

8. The idea behind Equation (1) is that the bank intervenes in the monetary market by changing the interest rate if real output diverges from potential output or inflation diverges from its target. If there are no cycles ($x=0$) or inflationary pressures ($\pi=\pi^*$), the objective is for the intervention rate to approach its long-term level, that is, $i=\pi+r$. Taylor presents the derivation of his rule intuitively: Given that the velocity of money depends on the interest rate (i), it is possible to find a relationship between i and price level P and real output Y. Exogenous changes in the monetary supply or velocity displace this relationship. This relationship will be maintained over time regardless of whether the money stock is growing at a fixed rate, although the money stock will respond systematically to interest rate and output changes. The response of money will be reflected in the change in parameters of the relationship.

9. Although Taylor (1993) is not attempting to define the Federal Reserve's "exact" rule, the equation describes a reaction function "as if" the monetary authority were following that rule, even though the Fed has never explicitly adopted an inflation-targeting policy regime.

10. Taylor's work not only stimulated the assessment of monetary rules, but also the search for optimal rules in a more general framework.

11. A sample of studies in this tradition of the literature can be found in Taylor (1999).

12. See Clarida *et al.* (1998); Clarida *et al.* (1999); Clarida *et al.* (2000); Rudebusch and Svensson (1999); Svensson (1999); and Svensson (1999), among others.

13. As in the works of Clarida, Gali and Gertler.

rate to the inflationary gap (the deviation of the inflation rate from its expected value or target) and to the output gap. Recently, other factors have been included that could influence the central bank's response, such as the intervention interest rate with a one-period lag[14], the exchange rate[15] or the price of assets (Bernanke and Gertler, 2000; Bernanke and Gertler, 2001; Carlstrom and Fuerst, 2007). In light of this, it has been important in the literature to quantitatively assess how the rules perform (Svensson, 2011), as well as the reaction of the central bank "as if" it were to conduct monetary policy using the reaction function that the Taylor rule prescribes or that which results from the specification of the economic model.[16]

The definition of monetary rules based on observation and empirical evaluation has been traditional in the literature on rules in the style of Taylor. These rules are not necessarily optimal monetary rules in the sense of being derived from an optimization problem; rather, they are simple, ad hoc rules concerning the way in which monetary policy is carried out. The literature has generated a set of optimal and nonoptimal monetary rules. This paper focuses on an optimal Taylor rule because we believe that it is important that the reaction function of a central bank come from a theoretical model of optimization that involves the usual concerns that a central bank of a small, open economy has in maintaining macroeconomic stability. From a theoretical point of view, there is another characteristic that a Taylor rule must satisfy (original, ad hoc, or optimal) and which is related to the size of the coefficients, particularly the coefficient that accompanies the inflationary gap or aversion to inflation (parameter $1 + h$ of Equation (1)). As shown in Woodford (2003, Chap. 2), a necessary condition is that it has a value greater than one so that the

14. To capture the central bank's desire to smooth out and avoid large movements in interest rates. For a detailed discussion, see Clarida *et al.* (1998), Woodford (2003, Chap 4), Woodford (2003), and Walsh (2010, Chap. 8).

15. See Ball (1999), Svensson (2000), and Taylor (2001). Giraldo *et al.* (2011) assess a monetary rule in which they include the real exchange rate, with a one- or two-period lag, as well as the smoothing of the interest rate for a closed economy, and they find that the coefficients associated with the exchange rate are not significant for Colombia between 1990 and 2010. For their part, Batini *et al.* (2003) and Batini *et al.* (2009) evaluate a series of monetary rules to determine their efficiency and robustness in a macroeconomic model. The common denominator of the rules involved international variables, such as the inclusion of real exchange rate levels (lagging and contemporary) as well as real devaluation. One of the rules includes the trade balance. The articles show that improvements in wellbeing due to inclusion of the exchange rate are marginal, although they do not analyze whether the monetary authority takes that variable into account when making policy decisions.

16. This type of assessment has been common for the United States: Taylor (1993); Taylor (1999a); Clarida *et al.* (1998); Clarida *et al.* (1999); Clarida *et al.* (2000), among others. Recently, similar exercises have been performed for other countries such as Germany (Kuzin, 2004) and Colombia (López, 2004; Pérez, 2005; Giraldo, 2008; Bernal and Tautiva, 2011; Giraldo *et al.*, 2011).

central bank's intervention will significantly affect the interest rate and thus investment and consumption will react to the change made by the monetary authority. This characteristic of monetary rules is known as the *Taylor principle* and in our view, an inflation-targeting policy should satisfy this principle.

Recently, the literature has addressed the possibility that a central bank may show an asymmetrical reaction to changes in variables that affect the interest rate in the monetary rule (traditionally the inflation gap and output gap), which would reflect different monetary positions for reaching the price stabilization objective. For example, Kuzin (2004) shows that the parameters of an ad hoc monetary reaction function, particularly that of aversion to inflation, have varied over time, and provides evidence for Germany's Bundesbank that when the inflation gap is positive, the Bundesbank reacts more strongly in raising interest rates than when the gap is negative.

Kuzin finds this asymmetrical reaction by using a Kalman filter as well as estimating with a Markov-switching methodology, where the parameters vary according to the economic regime. Furthermore, Kuzin estimates a simple Taylor rule (ad hoc, modified) with interest-rate smoothing. The main finding is that aversion to inflation is not constant and shows large and sudden changes during the period in which a monetary aggregates tracking strategy is used. In particular, Kuzin finds periods of high aversion and low aversion to inflation, in which the former are those coinciding with fulfillment of the Taylor principle. Despite the strong central bank tradition of the Bundesbank, there have been periods in which aversion is low and therefore the Taylor principle is not fulfilled. From a practical point of view, the Bundesbank follows a policy that accommodates changes in inflation during these periods.

On the other hand, Murray *et al.* (2008) use a simple macroeconomic model to argue that when the inflation rate follows a stationary process, it is because the central bank follows the Taylor principle, given that the monetary authority reacts strongly to inflationary pressures, driving inflation towards its long-term target. These authors show that the way in which the Federal Reserve drives monetary policy is not always geared toward stabilizing inflation, and they suggest that it is more common to face diverse regimes in which the parameters reflect different preferences of the central bank for stabilizing output or inflation. The authors estimate a Markov-switching model to show that U.S. monetary policy can be described as one of two states: one in which the Taylor principle is fulfilled and the inflation rate is stationary, and another in which the Taylor principle is not fulfilled and the inflation rate follows a unit root process.

The findings of Murray *et al.* (2008) contrast with those of Clarida *et al.* (1998), Clarida *et al.* (2000) and Orphanides (2000), which report, for the Federal Reserve, stability of parameters in periods that have been identified as inflation-stabilizing. What is novel about Murray *et al.* (2008) is that they find the break points or regime-switch points endogenously, using the Markov-switching methodology. As they argue, it is very risky to define, *a priori*, the periods of time in which one or another monetary regime is believed to function, based only on history. In this sense, this article follows Murray *et al.* (2008) in that the regime-switching dates are allowed to be defined endogenously even for the case of Colombia, which adopted the target inflation rate monetary policy explicitly in October 2000. However, since the 1990s the central bank of Colombia had been implicitly implementing some measures that are, to a certain extent, compatible with an inflation targeting regime.

Based on all this, this article studies an optimal Taylor rule in the framework of a small, open economy such as Colombia's, which has the following characteristics: i) it responds to deviations in inflation and the output gap (as the original rule; ii) it involves the central bank's desire to smooth the interest rate; and iii) it includes the exchange rate (or variables related to it). The model that we present below is innovative in that it includes variables from other countries that could affect the monetary policy of a small, open economy, unlike what has usually been done in the literature with use of the original Taylor rule, which is designed for a (virtually) closed economy such as the United States.

3. Model

3.1. Basic assumptions

Consider the central bank of a small, open economy that takes as given an aggregate intertemporal IS curve as follows:

$$x_t = E_t x_{t+1} + \delta x_{t-1} - \sigma(r_t - r^n) + \delta^* x_t^* + \alpha_1 e_t + \varepsilon_{1t} \tag{2}$$

and an aggregate supply curve as follows:

$$\pi_t = \beta E_t \pi_{t+1} + \gamma \pi_{t-1} + kx_t + \alpha_2 e_t + \varepsilon_{2t} \tag{3}$$

where x denotes the log of the output gap for the domestic economy while x^* denotes the output gap for the country that is the trading partner of the domestic economy, the operator $\mathrm{E}_t(\cdot)$ is the expected value in period t of the variable in period $t+1$ using all observable information until period t, and r denotes the interest rate that satisfies:

$$r_t = i_t - E_t \pi_{t+1} \tag{4}$$

which is the Fisher equation, where i denotes the logarithm of the nominal (short-term) interest rate and π is the log of the inflation rate. That said, the variable r^n in Equation (2) represents the long-term level of the logarithm of the real interest rate, which is assumed to be exogenously determined in the domestic economy. The variable e in equations (2) and (3) denotes the log of the real exchange rate, which is defined as

$$e_t \equiv R_t - p_t^* - p_t \tag{5}$$

where R is the logarithm of the nominal exchange rate, p^* is the logarithm of the price level of the trading partner, and p is the logarithm of the price level of the domestic economy. We assume that a small open economy must necessarily satisfy the condition of interest rate parity given by:

$$i_t - i_t^* = E_t R_{t+1} - R_t + \varphi_t \tag{6}$$

where i^* is the logarithm of the nominal interest rate of the trading partner and φ is the risk premium of the domestic economy relative to the trading partner. The imposition of this parity condition follows Svensson (2000) who argues that a truly open economy must satisfy this condition. The risk premium φ incorporates any type of disturbance exogenous to the exchange rate, including changes in portfolio preferences, credibility effects, etc. Using Equation (5) we can describe this parity condition in terms of the logarithm of the real exchange rate[17] given the definitions $E_t p_{t+1} - p_t \equiv \pi_t$ and $E_t p^*_{t+1} - p^*_t \equiv \pi^*_t$ which generates the following form, lagged one period:

$$e_t - e_{t-1} = (i_{t-1} - i^*_{t-1}) - (\pi_t - \pi^*_t) - \varphi_{t-1} \tag{7}$$

The parameters of the equations described above are assumed to satisfy the following: $\sigma > 0$, $\delta \in (0,1)$, $\delta^* > 0$, $\gamma \in [-1,1]$, $\beta \in [0,1], k > 0$, while α_1 and α_2 can be, but aren't necessarily, positive. The terms ε_1 and ε_2 in (2) and (3) denote, respectively, the demand and supply disturbances, which are assumed to be i.i.d. normally distributed disturbances with a zero mean and variances σ_1^2 and σ_2^2 , respectively.

We assume that the country that is the trading partner of the open domestic economy is an economy that follows a Taylor rule for a closed economy with the form:

$$i^*_t = r^{n*} + \eta_x x^*_t + \eta_\pi \pi^*_t + \varepsilon^*_t \tag{8}$$

where $\eta_x \neq 0$, $\eta_\pi > 1$ and $\varepsilon^*_t \sim iid(0,\sigma^{*2})$. In addition, we assume that the expectations of the open economy's central bank concerning x^*_t, π^*_t and φ_t are stable AR(2) stochastic processes where

$$\begin{aligned} E_t x^*_{t+1} &= \rho_{x,1} x^*_t + \rho_{x,2} x^*_{t-1} \\ E_t \pi^*_{t+1} &= \rho_{\pi,1} \pi^*_t + \rho_{\pi,2} \pi^*_{t-1} + \bar{\pi}^* \\ E_t \varphi_{t+1} &= \rho_{\varphi,1} \varphi_t + \rho_{\varphi,2} \varphi_{t-1} + \bar{\varphi} \end{aligned} \tag{9}$$

such that $\bar{\pi}^* \geq 0$, $\bar{\varphi} \geq 0$ and the coefficients $\rho_{h,i}$ for $h = x^*, \pi^*, \varphi$ and $i = 1,2$ satisfy $|\rho_{h,i}| < 1$, $\rho_{h,1} + \rho_{h,2} < 1$, $\rho_{h,2} < 1 + \rho_{h,1}$, $|\rho_{h,2}| < 1$.

We also assume that the central bank has a loss function represented by

$$\begin{aligned} \frac{\mathcal{L}_t}{2} = &\left(\pi_t - \bar{\pi}\right)^2 + \lambda_1 x_t^2 + \lambda_2 \left(e_t - e^n\right)^2 \\ &+ \lambda_3 \left(e_{t-1} - e_{t-1}\right)^2 + \lambda_4 \left(i_t - i_{t-1}\right)^2 \end{aligned} \tag{10}$$

17. Take Equation (5) which is forwarded one period such that $E_t e_{t+1} = E_t R_{t+1} + E_t p^*_{t+1} - E_t p_{t+1}$ and then substract $e_t = R_t + p^*_t - p_t$, which generates $E_t e_{t+1} - e_t = E_t R_{t+1} - R_t + E_t p^*_{t+1} - p^*_t - E_t p_{t+1} + p_t$. Therefore, $E_t R_{t+1} - R_t = E_t e_{t+1} - e_t - (E_t p^*_{t+1} - p^*_t) + (E_t p_{t+1} - p_t)$. Replacing this equation in (6) generates $i_t - i^*_t = E_t e_{t+1} - e_t + (E_t p_{t+1} - p_t) - (E_t p^*_{t+1} - p^*_t) + \varphi_t$.

where $\lambda_i = 1,..,4$ are nonnegative parameters. The first two terms are common to the majority of monetary rules, where λ_1 reflects the weight of the output gap relative to the weight of one in inflation's deviation from its target $\overline{\pi}$. Having said that, the term associated with λ_2 reflects the weight that the central bank of the domestic economy assigns to the difference between its real exchange rate and the long run target rate e^n (which may be explicit or implicit for the central bank). This is justified in part by the fact that small, open economies have had explicit exchange rate bands (publicly known) or even implicit ones (not publicly known), as in Colombia. Moreover, we consider that the central bank of this small open economy intervenes when there are strong fluctuations in the real exchange rate from one period to another, which is reflected in the weight λ_3 that appears in the bank's loss function. Finally, concern about nominal interest rate fluctuations is captured by the term associated with λ_4 similar to Woodford (2003a).

3.2. Optimal Taylor rule for an open economy

We assume that the central bank's behavior can be empirically understood to be the result of a monetary policy that arises from minimizing (10) with respect to (π, x, i, e) subject to (2), (3), (4), and (7). However, this problem can be solved in a simple fashion if (7) is first replaced in Equation (10) which generates the following loss function:

$$\begin{aligned}\frac{\mathcal{L}_t}{2} &= \left(\pi_t - \overline{\pi}\right)^2 + \lambda_1 x_t^2 + \lambda_2\left(e_t - e^n\right)^2 \\ &\quad + \lambda_3\left(\left(i_{t-1} - i^*_{t-1}\right) - \left(\pi_t - \pi^*_t\right) - \varphi_{t-1}\right)^2 + \lambda_4\left(i_t - i_{t-1}\right)^2\end{aligned} \tag{11}$$

Then, Equation (4) is replaced in Equation (2) and therefore the optimal monetary policy arises from minimizing Equation (11) with respect to (π, x, i, e) subject to (2) and (3). The corresponding Lagrangian function associated with this problem is then

$$\begin{aligned}\mathcal{L} = E_t\Bigg\{\sum_{t=0}^{\infty}\beta^t\Bigg\{&\frac{L_t}{2} + \phi_1\left[x_t - x_{t+1} - \delta x_{t-1} + \sigma\left(i_t - \pi_{t+1} - r^n\right) - \alpha_1 e_t - \gamma^* x^*_t - \varepsilon_{1t}\right] \\ &+ \phi_2\left[\pi_t - \beta\pi_{t+1} - \gamma\pi_{t-1} - kx_t - \alpha_2 e_t - \varepsilon_{2t}\right]\Bigg\}\Bigg\}\end{aligned}$$

where $\beta \in (0,1)$ is the discount rate, ϕ_1 and ϕ_2 are the Lagrange multipliers, and it is assumed that they are constant in the regime implemented by the central bank.[18] Because the loss function is a quadratic function the following first-order conditions, which are necessary and sufficient, solve the optimization problem.[19] Solving the first-order equations jointly with (2) and (3), the endogenous variables ϕ_1 and ϕ_2 can be eliminated and everything can be collapsed into the following *reduced* form equation:

$$\begin{aligned} i_t = {} & \beta_0 + \beta_1 E_t x_{t+1} + \beta_2 x_t + \beta_3 x_{t-1} + \beta_4 x_t^* + \gamma_1 \pi_t + \gamma_2 \pi_{t-1} + \gamma_3 \pi_t^* \\ & + \delta_1 E_t i_{t+1} + \delta_2 i_{t-1} + \delta_3 i_{t-1}^* + \delta_4 \varphi_{t-1} + \frac{\varepsilon_{1t}}{\sigma \Gamma} - \frac{\varepsilon_{2t}}{\beta \Gamma} \end{aligned} \tag{12}$$

where the parameter vector $\theta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \gamma_1, \gamma_2, \gamma_3, \delta_1, \delta_2, \delta_3, \delta_4)$ consists of reduced form parameters that are functions of the deep structural parameters of the model as derived formally in Appendix A and the parameter Γ is defined as

$$\Gamma \equiv \left[1 + \left(\frac{\alpha_1 m_1 + \alpha_2 m_2}{\sigma m_1 \lambda_2} \right) \left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta} \right) \lambda_4 \left(1 + \beta \right) \right]$$

where constants m_1 and m_2 are defined in Appendix A.

We call the optimal monetary rule of Equation (12) the *optimal Taylor rule* for an open economy. It is important to note that exchange rates (real or nominal) do not directly appear in the optimal Taylor rule. This is interesting precisely because it could be considered that the optimal monetary rule for an open economy would involve the reaction of nominal interest rate to changes in the real exchange rate, since it has been explicitly assumed that real exchange rate fluctuations are present in the central bank's loss function (Equation (10)) under $\lambda_3 > 0$. However, the exchange rate does not appear because of the assumption of interest rate parity equation (7) between the domestic economy and its trading partner, which implies that movements in the domestic economy's real exchange rate are due to exogenous factors such as the inflation rate, interest rate, and risk premium relative to the trading partner.

18. The constancy of these Lagrange multipliers is a restriction that Woodford (2003b), for example, does not impose. However, we do this to maintain an analysis with an explicit solution; otherwise we would have to resort to numerical simulations. This constancy means that the central bank would have constant reactions over time to changes in (2) and (3).

19. See Appendix A for an explicit solution to this optimization problem.

a. Analysis of the optimal Taylor rule

Closed economy and the original Taylor rule

It is instructive to briefly analyze the sufficient conditions under which the monetary rule of Equation (12) becomes a Taylor rule for a closed economy similar to that considered by Taylor (1993) for the United States economy. A closed economy satisfies $\delta^{*} = \alpha_1 = \alpha_2 = 0$ which implies $\Gamma \equiv 1$. In addition, under naive expectations $E_t x_{t+1} = x_t$ and without inflationary inertia in the Phillips curve, $\gamma = 0$, the structural equation that gives rise to Equation (12) becomes:

$$i_t = r^n + \left(\frac{\delta}{\sigma} - \frac{k}{\beta}\right) x_t + \frac{1}{\beta}\pi_t + \frac{\varepsilon_{1t}}{\sigma} - \frac{\varepsilon_{2t}}{\beta} \tag{13}$$

Taylor (1993) proposes characterizing the monetary policy of the United States in the 1970s and 1980s as based on a rule such as that given in Equation (13) with a coefficient associated with gap x of around 0.5 and a coefficient associated with the inflation rate π around 1.5. Furthermore, as outlined above, Taylor's work generated what Woodford (2003) has called the Taylor principle, which means that the coefficient associated with π is strictly greater than one. The Taylor princple implies that the nominal interest rate varies more than proportionally to the variation in inflation in period t. Note that in the optimal rule in a closed economy the coefficient associated with π is $1/\beta$ so it is necessarily greater than one because $\beta \in (0,1)$. Therefore, the case of the Taylor rule for a closed economy satisfies the Taylor principle. In fact, the values 0.5 and 1.5 that Taylor proposes can be replicated in this model under the parametric structural assumption $\beta = 2/3,\ \sigma = 1,\ k = 1/6,\ \delta = 3/4$.

In addition, it should be emphasized that if the inflation rate presents a certain inertia in the sense of the structural parameter such that $\gamma \in (-1,1)$ then the optimal reaction of monetary policy in the long run, compatible with a steady state[20], is

20. Appendix A contains a proof that the general model has a unique steady state under certain plausible restrictions.

$$\frac{di_e}{d\pi_e} \equiv LR_\pi = \frac{1-\gamma}{\beta}$$

where $\pi_t = \pi_e$ and $i_t = i_e$ for all t in the steady state. A sufficient condition to satisfy the Taylor principle in this case is then $1 > \beta + \gamma$. In addition, the coefficient associated with the gap x is not necessarily positive in Equation (13) since it is possible that $(\delta / \sigma) - (k / \beta) < 0$, although empirically it does not appear to be a relevant case. Finally, demand shocks (changes in ε_1) and supply shocks (changes in ε_2) optimally increase and decrease, respectively, the interest rate i_t.

Open economy

We now study the *structural* optimal Taylor rule for an open economy ($\alpha_1 \neq 0$ and $\alpha_2 \neq 0$) deduced in the appendix in Equation (A13) which is the equation that gives rise to the reduced optimal Taylor rule given by Equation (12). The signs of the different coefficients in Equation (A13) are generally ambiguous because they depend specifically on whether the term $(\alpha_1/\sigma) - (\alpha_2 / \beta)$ is positive or negative. To simplify the analysis, we make an assumption on some parameters:

$$\sigma k = 1 - \beta(1 - \beta\delta) > 0.$$

Under this restritcion we obtain $((\alpha_1 m_1 + \alpha_2 m_2)/\alpha m_1) = ((\alpha_1/\sigma) - (\alpha_2/\beta))$ where the definitions of m_1 and m_2 given in equations (A6) and (A9) of the appendix were used. Hence, the constant Γ can be simplified:

$$\Gamma \equiv 1 + \left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)^2 \left(\frac{\lambda_4\left(1+\beta\right)}{\lambda_2}\right)$$

We can conclude, therefore, that $\Gamma \geq 1$ given that $\lambda_4\,(1+\beta) / \lambda_2 > 0$. Equation (A13) is simplified somewhat, generating the following equation:

$$\begin{aligned}
i_t = & \frac{1}{\Gamma}\left[r^n + \left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)e^n\right] + \frac{E_t x_{t+1}}{\sigma\Gamma} + \frac{1}{\Gamma}\left\{\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)\left(\frac{\alpha_2\lambda_1}{k\lambda_2}\right) - \left(\frac{1+\beta^2\delta}{\beta\sigma}\right)\right\}x_t \\
& + \frac{\delta}{\sigma\Gamma}x_{t-1} + \left(\frac{1}{\beta\Gamma} + \left(\frac{\lambda_{3\beta}\left(1+\frac{1}{1-\beta\gamma}\right)}{\lambda_2\Gamma}\right)\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)^2\right)\pi_t - \frac{\gamma}{\beta\Gamma}\pi_{t-1} \\
& + \frac{\delta^*}{\sigma\Gamma}x_t^* + \frac{\delta^*}{\sigma\Gamma}x_t^* - \left(\frac{\lambda_{3\beta}\left(1+\frac{1}{1-\beta\gamma}\right)}{\lambda_2\Gamma}\right)\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)^2\pi_t^* \\
& + \left(\frac{\lambda_3\left(\beta+\sigma m_1\right)}{\Gamma}\right)\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)^2 i_{t-1}^* \\
& + \left(\frac{\lambda_{3\beta}\left(1+\frac{1}{1-\beta\gamma}\right)}{\lambda_2\Gamma}\right)\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)^2\varphi_{t-1} + \frac{\beta\lambda_4}{\Gamma}\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)^2 E_t i_{t+1} \\
& + \left(\frac{\lambda_4 - \lambda_3\beta\left(1+\frac{1}{1-\beta\gamma}\right)}{\lambda_2\Gamma}\right)\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)^2 i_{t-1} + \frac{\varepsilon_{1t}}{\sigma\Gamma} - \frac{\varepsilon_{2t}}{\beta\Gamma}
\end{aligned} \tag{14}$$

Then, as the optimal Taylor rule has interest rate lags and leads, it is important to be careful when determining whether the Taylor principle is satisfied in the long term. We must find in steady state the long-term effect of a 1% change in π over i according to the optimal Taylor rule. We replace $\pi_t = \pi_e$ and $i_t = i_e$ for all t, where the subscript e means *evaluated* in steady state, and then we obtain $(di_e / d\pi_e) \equiv LR_\pi$ which yields:

$$LR_\pi = \frac{\frac{1-\gamma}{\beta\Gamma} + \left(\frac{\lambda_3\beta\left(1+\frac{1}{1-\beta\gamma}\right)}{\lambda_2\Gamma}\right)\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)^2}{1 - \left[\frac{\left(1+\beta\right)\lambda_4 - \lambda_3\beta\left(1+\frac{1}{1-\beta\gamma}\right)}{\lambda_2\Gamma}\right]\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)^2}$$

The following result gives the sufficient conditions for satisfying $LR_{\pi} > 1$ which is the long-term version of the Taylor principle (compatible with a steady state) for an open economy.

Proposition 1. If $\sigma k = 1 - \beta(1 - \beta\delta)$ and $1 > \beta + \gamma$ then the optimal Taylor rule for an open economy given by (14) satisfies $LR_{\pi} > 1$.

Proof: First, note that $LR_{\pi} > 0$ is satisfied if the denominator is positive, given that the numerator is always positive. Therefore, the following must be satisfied:

$$1 > \left[\frac{\left(1+\beta\right)\lambda_4 - \lambda_3\beta\left(1+\dfrac{1}{1-\beta\gamma}\right)}{\lambda_2\Gamma}\right]\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)^2$$

which is an inequality that is simplified if

$$\Gamma \equiv 1 + \left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)^2 \left(\frac{\lambda_4\left(1+\beta\right)}{\lambda_2}\right)$$

is replaced and then the equation is reorganized to be expressed as

$$1 + \left[\frac{\lambda_3\beta\left(1+\dfrac{1}{1-\beta\gamma}\right)}{\lambda_2\Gamma}\right]\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)^2.$$

Clearly this condition is always positive given that $1 > \beta\gamma$ and that the other terms are positive. In addition, note that $LR_{\pi} > 1$ is satisfied once the constant Γ is replaced and it is noted that $1 > \beta + \gamma$ is satisfied by assumption.

■

Corollary. The condition $1 > \beta + \gamma$ implies that it is more likely that the central bank of a small open economy that implements an optimal Taylor rule does not satisfy the Taylor principle the greater the persistence $\gamma > 0$ in the inflation rate, for a given intertemporal discount rate of the domestic economy. This shows that an inflation rate that is highly persistent, say has a unit root or a stochastic trend, should not satisfy the Taylor principle. This key observation will be crucial for our empirical results.

Now, the coefficients associated with x_{t+1} and x_{t-1} in Equation (14) are positive, while the sign associated with x_t has an ambiguous sign even under the simplifying assumption $\sigma k = 1 - \beta(1 - \beta\delta)$. The long-term effect is also ambiguous. To see this, note that we replaced $x_t = x_e$ and $i_t = i_e$ for all t, and then we obtain $(di_e / dx_e) \equiv LR_x$:

$$LR_x = \frac{\frac{1}{\Gamma}\left\{\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)\left(\frac{\alpha_2\lambda_1}{k\lambda_2}\right) - \left(\frac{1+\beta^2\delta}{\beta\sigma}\right)\right\} + \frac{1+\delta}{\sigma\Gamma}}{1 - \left[\frac{(1+\beta)\lambda_4 - \lambda_3\beta\left(1 + \frac{1}{1-\beta\gamma}\right)}{\lambda_2\Gamma}\right]\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)^2}$$

The denominator is positive as shown in Proposition 1, while the numerator can have either sign. Similar to the case of a closed economy, the optimal reaction of monetary policy to changes in the output gap can have an ambiguous sign.

Finally, it is important to note that if within the bank's loss function the weighting relative to changes in the nominal interest rate λ_4 is sufficiently large compared to the weighting λ_3 that corresponds to changes in the real exchange rate around an implicit target, the optimal Taylor rule smooths the movement of the nominal interest rate. This is summarized in the following proposition.

Proposition 2. If $\sigma k = 1 - \beta(1 - \beta\delta)$, $(\alpha_1/\sigma) \neq (\alpha_2/\beta)$ and $(\lambda_4/\lambda_3) \geq \beta\left(1 + \left(1/(1 - \beta\gamma)\right)\right)$ are satisfied, then the optimal Taylor rule in Equation (14) smooths the temporal behavior of the nominal interest rate.

> **Proof**: Note that in Equation (14) the coefficient associated with $E_t i_{t+1}$ is always positive under the assumption $(\alpha_1/\sigma) \neq (\alpha_2/\beta)$. In addition, the associated coefficient i_{t-1} is positive under $(\alpha_1/\sigma) \neq (\alpha_2/\beta)$ and $(\lambda_4/\lambda_3) \geq \beta\left(1 + \left(1/(1 - \beta\gamma)\right)\right)$. Thus, we conclude that under the conditions formulated, a central bank in a small open economy finds it optimal to smooth out changes in the interest rate.
>
> ■

Woodford (2003a) has argued that optimal monetary rules can have a smoothing effect on the nominal interest rate even in closed economies where the real exchange rate does not play a role. It is important to note that in our model we obtain an optimal monetary rule with smoothing of the nominal interest rate only if the economy is open, i.e., $\alpha_1 > 0$ and $\alpha_2 > 0$ such that $\alpha_1 \neq \alpha_2$ under the assumption $\sigma k = 1 - \beta(1 - \beta\delta)$. As can be observed in Equation (14) if $\alpha_1 = \alpha_2 = 0$ we end up with a rule that does not smooth out the behavior of the nominal interest rate. Naturally, we are not arguing that a closed economy model cannot present smoothing in the behavior of the nominal interest rate. In fact, Woodford (2003a) has argued that smoothing of the behavior of the nominal interest rate can be present in an optimal monetary policy even if the central bank's loss function does not take into account changes in the nominal interest rate. What we note here, however, is that smoothing of nominal interest rate behavior occurs in our model only for an open economy.

The following proposition gives sufficient conditions for the optimal Taylor rule to be consistent with a dynamic system with a unique steady state.

Proposition 3. If $\sigma k = 1 - \beta(1 - \beta\delta)$, $1 > \beta + \gamma$, $\gamma \in (-1,1)$, $0 > \alpha_1\beta^2\gamma > \alpha_2\sigma$, $\lambda_2 \leq \beta\lambda_3^2$ and

$$\lambda_1 < min\left\{\frac{k^2\left[\lambda_3\left(2-\beta\gamma\right)\right]+1}{\beta\gamma\left(1-\beta-\gamma\right)}, \frac{k\left(1-\beta\delta\right)\left(1-\beta\right)\left(\lambda_3\left(2-\beta\gamma\right)+1\right)}{\left(1-\gamma\right)\sigma\beta\gamma}\right\},$$

then the optimal Taylor rule for Equation (14) is consistent with a dynamic system that has a unique steady state.

See Appendix A for the proof of this proposition. It is important to note that we do not study the conditions for the local or global dynamic properties of the steady state of the underlying linear dynamic system. We assume, in fact, that the dynamic around the steady state is not divergent: It can be local or globally convergent or even a saddle point. The conditions for it to be convergent or a saddle point with a trajectory that goes to steady state is outside the scope of this article, due to the complexity of carrying out this analytical exercise, given the dimensions of the dynamic system that would require analysis. What is important is that we are assuming that there is a dynamic trajectory that goes to the steady state in the long run and that the economy ends up there, through implementation of the optimal Taylor rule. Also, since our concern is more empirical than theoretical in analyzing the behavior of the Bank of the Republic in the case of Colombia we do not go further in studying the theoretical model.

4. Estimable empirical specification

The econometric estimation of the reduced form Equation (12) is carried out using a Markov-switching model that considers the existence of different states or regimes. It should be pointed out that the change from one regime to another is characterized by a dichotomous stochastic variable (S_t) that is endogenous to the system and which represents the state or regime of the economy where $S_t \in \{1,1\}$. Therefore, from Equation (12) the econometric model can be written as:

$$\begin{aligned} i_t = {} & \beta_0 + \beta_1 E_t x_{t+1} + \beta_2 x_t + \beta_3 x_{t-1} + \beta_4 x_t^* \\ & + \gamma_1 \pi_t + \gamma_2 \pi_{t-1} + \gamma_3 \pi_t^* \\ & + \delta_1 E_t i_{t+1} + \delta_2 i_{t-1} + \delta_3 i_{t-1}^* + \delta_4 \varphi_{t-1} + u_t \end{aligned} \tag{19}$$

where $(\varepsilon_{1t}/\sigma\Gamma) - (\varepsilon_{2t}/\beta\Gamma) \equiv u_t \sim \text{iid}(0,\sigma^2_{S_t})$ is an error or disturbance and where we define $\beta_{S_t} = \beta_0(1 - S_t) + \beta_1 S_t$, $\sigma^2_{S_t} = \sigma^2{}_0(1 - S_t) + \sigma_1^2 S_t$.

The nature of the system is that the variable S_t has the property of a first-order Markov process, that is, the regime at the time is determined only by the preceding regime and the evolution of regime changes is associated with the realization of that stochastic process. The probabilities of transition between the two states represent this Markovian property:

$$\Pr[S_t = 0|S_{t-1} = 0] = \text{p}, \Pr[S_t = 1|S_{t-1} = 1] = \text{q},$$
$$\Pr[S_t = 1|S_{t-1} = 0] = 1 - \text{p}, \Pr[S_t = 0|S_{t-1} = 1] = 1 - \text{q}$$

where p is the probability of the economy being in state 0 at time t given that it was in that same state at time $t - 1$; q is the probability of the economy being in state 1 at time t given that it was in that same state in the preceding time period; and $1 - \text{p}$ and $1 - \text{q}$ are the probabilities of transition from one regime to the other.

5. Data for Colombia

The frequency of the Colombian data is quarterly and the data covers the first quarter of 1990 through the fourth quarter of 2011. The data used to estimate the econometric model are succinctly described as follows:

- Intervention interest rate: for Colombia, we take the datum of the interbank interest rate of the Bank of the Republic[21], i_t. The intervention rate of the United States, i_t^*, is taken from the Federal Reserve (FED).
- Output gap: defined as the difference in the logarithm of observed and potential output. The datum for Colombia is provided by the Bank of the Republic, x_t, and that of the United States is the result of applying the Hodrick-Prescott filter to its real observed GDP, which is taken from the Bureau of Economic Analysis (BEA) of the United States, x_t^*.
- Inflation rate: the Colombian data are from the National Statistics Department and Bank of the Republic, π_t. Data from the BEA are used for United States inflation π_t^*.
- Risk premium: This is from the Chicago Board Options Exchange Market Volatility Index, φ_t.

6. Results

6.1. Estimation of the modified Taylor rule under Markov switching[22]

Table 1 reports the results of the estimation of the reduced-form model of the optimal Taylor rule for an open economy (12) using a Markov-switching methodology with Colombian quarterly data for the period 1990-I to 2011-IV. Statistical tests were performed to determine the existence of two or more regimes, however the evidence suggests that in the period studied there are just two monetary policy regimes in terms of the estimated Taylor rule.

21. The interbank interest rate is used because it is strongly influenced by the Bank of the Republic's intervention rate. As the institution itself acknowledges, "The rate is more heavily influenced by policies of the monetary authority, as the contraction and expansion operations are concentrated in the same periods, between 1 and 14 days" (Bank of the Republic, 1999, p. 16).

22. The empirical model follows the theoretical development of Hamilton (1994) and Krolzig (1997). See details in Appendix B.

Table 1 shows the empirical results for the two regimes obtained, where the confidence interval is reported for the parameter of the variable with bootstrapped standard errors and the t statistic under standard errors from maximum likelihood. As the table shows, the conclusions are similar under both estimation methods, where the exceptions are for the forward interest rate i_{t+1} in regime 0, which is statistically significant under standard errors of the maximum likelihood procedure but not statistically significant under bootstrapped standard errors. The other exception is for the output gap variable with a one-period lag, x_{t-1} and the current output gap x_t for regime 1, which are statistically significant under maximum likelihood standard errors but are not statistically significant under bootstrapped standard errors. We prefer the statistical inference under the bootstrap procedure which is a more conservative approach.

Regime 0 coincides with the period during which the BR behaves under inflaton targeting (IT) while regime 1 coincides with the period in which the political constitution granted the central bank technical independence from the government, and the BR conducted monetary policy under a structure that was not specifically inflation targeting, as defined in the literature, but which nonetheless had similar characteristics, such as the public announcement of a quantitative inflation target. In this paper, this is referred to as a regime with no inflation targeting (NIT).

To understand why we denote regimes 0 and 1 this way, consider Figure 1, which reports the filtered probabilities or occurrence probabilities of regime 0 (IT) on the vertical axis with the quarters on the horizontal

Figure 1. IT regime 0 occurrence probabilities

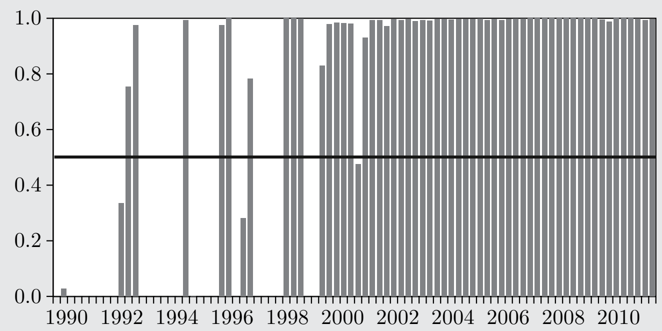


Table 1. Markov-switching model estimated by maximum likelihood, 1990-2011

Dependent variable: Intervention interest rate						
Variables	Regime 0: IT			Regime 1: NIT		
	Coefficients	10% C. interval bootstrapping	t statistics M. Likelihood	Coefficients	10% C. interval bootstrapping	t statistics M. Likelihood
i_{t-1}	0.361	[0.33, 0.56]*	13.7*	-0.35	[-0.56, -0.17]*	-2.11*
i_{t+1}	0.212	[-0.091, 0.30]	6.88*	-0.16	[-0.20, 0.16]	-0.97
x_{t+1}	-0.0165	[-0.67, 0.35]	-0.07	-9.68	[-9.79, 1.19]	-3.51*
x_t	-0.051	[-0.09, 0.08]	0.13	8.35	[-2.60, 9.12]	1.93*
x_{t-1}	0.155	[-0.22, 0.88]	0.70	4.76	[4.28, 12.4]*	1.47
π_t	-0.041	[-0.28, 0.28]	-0.46	1.18	[-1.23, 1.66]	1.01
π_{t-1}	0.39	[0.11, 0.67]*	4.41*	0.07	[-0.42, 2.00]	0.07
x_{t^*}	0.28	[0.06, 0.52]*	2.75*	4.10	[0.98, 5.54]*	2.29*
π_{t^*}	0.03	[-0.20, 0.07]	0.37	-2.87	[-3.47, 1.38]	-1.54
i_{t^*}	0.158	[0.05, 0.34]*	2.99*	-0.05	[-0.81, 2.18]	-0.05
ϕ_{t-1}	0.0138	[-0.003, 0.04]	1.31	0.45	[-0.28, 0.68]	1.18
Constant	0.003	[-0.003, 0.011]	0.80	0.16	[0.01, 0.24]	1.97*
Std error	0.004			0.058		
LR_x	0.295	[-1.11, 1.80]		3.53	[2.74, 10.59]*	
LR_π	0.61	[0.16, 1.52]*		1.23	[-1.10, 3.58]	

*Statistical Significance at 10%; Bootstrapping with 2000 replicas.
Long run Effects (LR) that have (*) are computed with coefficients that are statistically significant under bootstrapping

axis between 1990 and 2011. If the probability bar in the figure is greater than 0.5 it means that in that quarter, regime 0 is operating, while if the bar is less than 0.5, it means that in that quarter regime 1 is operating. As shown in the figure, after October 2000, when the Bank of the Republic announces its target inflation rate strategy, regime 0 predominantly obtains probabilities greater than 0.5, which explains why we call this the inflation targeting regime (IT). In contrast, prior to October 2000, regime 1 predominantly obtains probabilities greater than 0.5, and is thus referred to as the regime with no inflation targeting (NIT).

Returning to Table 1, it indicates that in regime 0 (IT) the Bank's behavior smooths movements of the intervention rate because the variable i_{t-1} is statistically significant and positive. This does not occur in regime 1 (NIT), which shows behavior not aimed at smoothing changes in the intervention rate in the period prior to October 2000. With respect to the relationship of the intervention interest rate to Colombia's output gap, what is observed is that the coefficients associated with the current level (x_t) have opposing signs in regime 0 and 1 while they have the same sign in both regimes for the forward gap (x_{t+1}) and lagged gap (x_{t-1}) variables. In addition, the values in absolute terms differ substantially in the two regimes. However, only in regime 1 is the intervention interest rate significant and statistically related to the output gap variables. This suggests that the BR followed a monetary policy prior to October 2000 that reacted strongly to changes in the output gap, especially with respect to the output gap lagged by one period according to bootstrapping standard errors. Using only the statistically significant variables according to bootstrapping standard errors, for regime 1 we can compute the long-term effect (LR) on the intervention interest rate, compatible with a steady state, of a one percentage-point change in the ouput gap in period t.

$$LR_x^1 = \frac{4.76}{1 + 0.35} = 3.53$$

This marginal effect is statistically significant because zero is not in the confidence interval at 90% with bootstrapping standard errors, which is $[2.74, 10.59]^*$. In practical terms, this long-term effect is significant because for each one-point increase in the output gap, the BR in regime 1 (NIT) increases the interest rate by 3.5 percentage points on average. In addition, using the statistically significant variables in regime 0 we find that $LR_x^0 = 0.29$, within the confidence interval at 90% with bootstrapping standard errors, which is in [-1.11, 1.80]. This means that the null hypothesis that the BR in regime 0 (IT) did not react to changes in the output gap cannot be rejected.

With respect to the inflation rate, we find that the signs of the point estimates of the parameters associated with the current inflation rate (π_t) variable in the two regimes are also different, while the signs of the lagged inflation rate (π_{t-1}) coincide. As observed with the output gap, the actual absolute values of the inflation rate also vary from one regime to another. Only in regime 0 (IT) do we find that lagged inflation is statistically significant while in regime 1 (NIT) none of the inflation variables are statistically significant. This suggests that the BR only reacted statistically to changes in the inflation rate in regime 0 under inflation targeting. Again, using only the statistically significant variables according to bootstrapping standard errors, for regime 0 we can compute the long term (LR) effect on the intervention interest rate, compatible with a steady state, of a one percentage-point change in period t of inflation:

$$LR_\pi^0 = \frac{0.39}{1 - 0.36} = 0.61$$

This estimate is in the 90% confidence interval with bootstrapping standard errors $[0.16, 1.52]^*$ indicating that at the 10% significance level, the null hypothesis under IT that the BR did not affect the intervention interest rate in the face of changes in the average interest rate in the long term can be rejected. Moreover, at the 10% significance level the null hypothesis that the Bank has a $LR_\pi^0 > 1$, as the Taylor principle suggests, cannot be rejected. In particular, note that values greater than 1, up to 1.52, are within this confidence interval, providing evidence that in regime 0 with IT the BR satisfies the Taylor principle. In addition, the long-term effect for regime 1 is $LR_\pi^1 = 1.23$ with a confidence interval of 90% with bootstrapping standard errors [-1.10, 3.58]. Although the specific value is greater than 1, the confidence interval shows that the null hypothesis that $LR_\pi^1 = 0$ cannot be rejected at 10%, indicating that the BR in the regime without inflation targeting did not react to changes in the inflation rate on average, and therefore, it probably did not satisfy the Taylor principle in that regime.

The empirical results found regarding the output gap and the inflation rate suggest that the Bank of the Republic in Colombia fundamentally changed its monetary policy in October 2000 by adopting an inflation targeting monetary strategy, as we find that in regime 1 (NIT) the BR did not react to changes in the inflation rate, while in regime 0 (IT) it reacted in a way that is compatible with Taylor's principle. In addition, while in regime 1 (NIT) the BR reacted strongly to changes in the output gap, in regime 0 (IT) we find that it did not react to changes in this variable.

Table 1 shows that in both regimes the Bank of the Republic reacted in a positive and statistically significant way, at a 10% significance level, with respect to the United States output gap (x_t^*), and this reaction was 20 times stronger in regime 1 than in regime 0 (4.1 versus 0.28). This is most likely because there is an exchange rate band in this period that the BR defended for several years. Moreover, only in regime 0 did the BR seem to have reacted in a positive and statistically significant way, at a 10% significance level, to changes in the lagged intervention rate (i_{t-1}^*) of the United States, while it did not react on average to the U.S. inflation rate (π_t^*).

As can be seen in Table 2, in probabilistic terms, both regimes are relatively persistent over time in the period studied, since the probability of remaining in regime 0 is 0.92 while for regime 1 it is 0.84, indicating incidentally that regime 0 (IT) is more persistent than regime 1 (NIT) in the period studied.

Table 2. Transition matrix

	Regime 0	Regime 1
Regime 0	0.912	0.088
Regime 1	0.162	0.838

Given the estimated probabilities of transition, the average length in each state can be calculated using the following equations:

$$\sum_{i=1}^{\infty} i q^{i-1}\left(1-q\right) \tag{25}$$

$$\sum_{i=1}^{\infty} i p^{i-1} \left(1-p\right) \tag{26}$$

Table 3 indicates that the average duration of the IT regime is 55.3 quarters, while the average duration of the NIT regime is 6.17.

Table 3. Duration of regimes

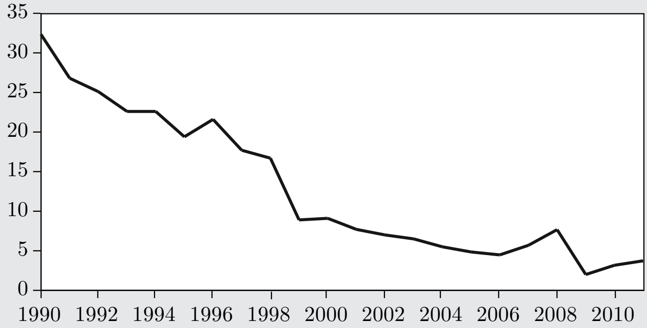
	Regime 0	Regime 1
Regime 0	55.3	12.47
Regime 1	31.7	6.17

6.2. Taylor's principle and persistent inflation rate behavior

As described in the previous section, the BR changed its monetary policy by implementing an inflation targeting strategy after October 2000 in Colombia. This change reveals that the intervention interest rate reacted prevalently to changes in the output gap in regime 1 (NIT), which coincides with the pre-inflation targeting period, while it did not react to the output gap in regime 0, which coincides with the inflation targeting period. In addition, when the BR implemented the target inflation rate strategy in October 2000, it generated a monetary policy that reacted strongly to changes in inflation and is compatible with the Taylor principle, something that did not occur in regime 1. Ideally, in order to conclude that the BR satisfied the Taylor principle in regime 0 (IT), the confidence interval for this regime would have been one in which the lower bound was strictly greater than 1. However, the evidence shows that the interval at 90% confidence for LR_{π}^{0} is $[0.16, 1.52]^{*}$ under bootstrapping standard errors, which clearly includes values that are positive but less than 1. Thus, while we cannot reject the hypothesis that in regime 0 the BR's inflation-targeting policy satisfies Taylor's principle, nor can we reject the possibility that it is not satisfied because we cannot reject null hypotheses where $LR_{\pi}^{0} < 1$.

Our analytical results show that Taylor's principle cannot be satisfied if the inflation rate is highly persistent as a time series. Hence a unit root behavior or the inflation rate is evidence that Taylor's principle is not satisfied.

Figure 2. Annual inflation rate in Colombia, 1990 to 2011



Source: Bank of the Republic of Colombia.

Moreover, as mentioned in the literature review, Murray *et al.* (2008) argue that the inflation rate should be a stationary time series under inflation targeting precisely because a central bank, by satisfying the Taylor principle, must react by intervening in the inflation rate more than proportionally to changes in the inflation rate, an argument consistent with our analytical results. This observation enables an indirect strategy for verifying whether the BR had implemented a monetary policy consistent with the Taylor principle after October 2000: we must verify that after October 2000 Colombia's inflation rate is a weakly stationary time series.

Figure 2 shows that Colombia's annual inflation rate appears to behave differently before and after the year 2000. Before 2000, double-digit inflation was rapidly decreasing while after 2000, according to the figure, it began to fall more slowly and remained stable at a 1-digit rate. This appears to suggest that 2000 is a year of regime switching consistent with what has been found above.

To indirectly verify the hypothesis that after 2000 the Bank of the Republic of Colombia began to satisfy the Taylor principle with a change to an IT policy, we must verify that the inflation rate should have a unit root in the regime that prevailed before October 2000 and then becomes a weakly stationary time series in the regime that prevailed after October 2000. To do this, a unit root test is implemented for a Markov-switching model following Camacho (2010), using a

Table 4. ADF test for inflation rate under a Markov switching process

Dependent variable: Change in inflation rate						
Variables	Regime 0			Regime 1		
	Coefficient	Std. error	t statistic	Coefficient	Std. error	t statistic
$\Delta\pi_{t-1}$	0.519	0.185	2.81	-0.02	0.112	-0.24
$\Delta\pi_{t-2}$	0.277	0.254	1.10	-0.10	0.11	-0.93
$\Delta\pi_{t-3}$	0.381	0.177	2.15	-0.18	0.088	-2.07
π_{t-1}	-0.455	0.116	-3.39	0.07	0.018	4.05
Trend	-0.0015	0.001	-3.93	0.0003	0.0001	3.32
Constant	0.132	0.0343	3.83	-0.021	0.003	-5.72
Standard error	0.0013			0.0052		

*Statistical significance at 10%; Bootstrapping with 2000 replicas

bootstrapping procedure, based on a regime-switching augmented Dickey-Fuller regression.[23]

Table 5. Bootstrapping DF unit root test for a Markov switching model

H0: Inflation rate has a Unit Root
H1: Inflation rate is a Stationary Series
Regime 0: p-value = 0.031
Regime 1: p-value = 0.883

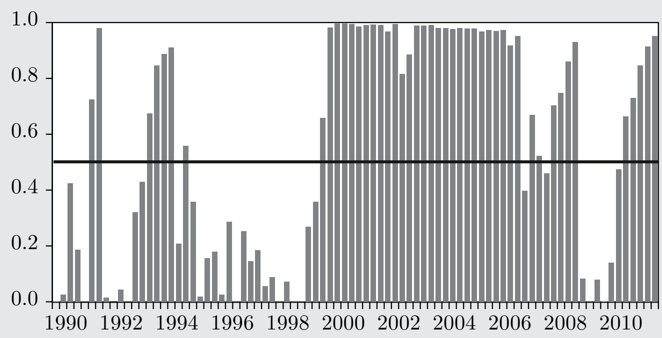
Table 4 shows the results of the estimation under maximum likelihood, while Table 5 resports the test of the hypothesis in question using standard errors under a bootstrapping procedure with 2,000 replicas, which is a more conservative procedure. According to Table 4 the coefficient associated with π_{t-1} is the DF unit root test. As seen the t stat is -3.39 for regime 0 while it is 4.05 for regime 1. Hence, according to the Dickey-Fuller table we can reject at the 10% level that in regime 0 there is a unit root while we cannot reject at the 10%

23. See Appendix B.

level that in regime 1 the inflation rate has a unit root. This same conclusion is obtained with bootstrapping standard errors as reported in Table 5, since the p value of the null for regime 0 is 0.03, which indicates that there is evidence to reject the null hypothesis of the existence of a unit root in this regime at a significance level of 10%, while the p value for regime 1 is 0.88, which supports the conclusion that there is no evidence to reject the null even at the 10% level that in this regime the inflation rate exhibits a unit root or stochastic trend.

In addition, the filtered probabilities or occurrence probabilities for the two regimes are estimated and shown in Figure 3. As can be observed, each of the two regimes is present in the entire period analyzed, although starting in 2000 the more prevalent regime is regime 0, which, as shown earlier, coincides with the adoption of inflation targeting by the BR, and in which the null hypothesis that the inflation rate has a unit root can be rejected. Likewise, regime 1 prevails before October 2000, when the BR followed a regime of no inflation targeting, and we find that in this regime the inflation rate exhibits a unit root and therefore the Taylor principle is not satisfied.

Figure 3. Regime 0 occurrence probabilities



It is the prevalence of regime 0 after October of 2000 in the unit root test that gives us the confidence to argue that this is the regime closest to the IT regime identified above given that the Markov switching unit root test is independent of the previous estimation procedure. Hence, the evidence provides greater certainty that the regime switching in the BR's monetary policy after October 2000 satisfies the Taylor principle precisely because after implementation of the IT monetary policy the prevailing regime is compatible with the inflation rate being a stationary time series.

This result allows us to conjecture, although with great prudence, that if the BR had not adopted the IT policy in October 2000 the inflation rate would have continued to be a time series with a stochastic trend or unit root. This surely would have had repercussions for the effectiveness of the BR's monetary policy in controlling inflationary shocks, because persistent series exhibit unpredictable and sometimes volatile behavior over time. Overall, our results support the idea that an IT policy is beneficial for small open countries like Colombia because it enables central banks to control inflationary shocks by maintaining the inflation rate stable at low levels in a credible way.

7. Concluding Remarks

This article has proposed and developed an optimal Taylor rule for an open economy that arises from a loss function optimization problem of a central bank that is concerned with deviations of inflation from its target, the economic cycles represented by the output gap, the desire to smooth the intervention rate and the deviation from the real exchange rate relative to a long-term goal. This optimal monetary policy is framed within a model with an intertemporal IS curve, an aggregate supply curve or Phillips curve, Fisher's equation and an interest rate parity condition, because the economy is small and open. The optimal Taylor rule has as a special case the original ad hoc Taylor rule for a closed economy as in Taylor (1993). Conditions are found under which Taylor's principle is more likely to be satisfied in the optimal Taylor rule for an open economy in which the inflation rate is not persistent over time. This shows that an inflation rate that has a unit root or a stochastic trend should not satisfy the Taylor principle.

Once the optimal Taylor's rule is obtained in its reduced form, it is estimated in a Markov-switching model for Colombia between 1990 and 2011 because the Bank of the Republic adopted an inflation targeting monetary policy midway through this period, in October 2000. We do not impose this potential structural change exogenously; rather, we allow the methodology to estimate for us the number of regimes in the period studied. We find two different regimes in the period studied, consistent with a change in the parameters of the optimal Taylor rule with respect to the inflation rate and the output gap, after the inflation targeting policy is adopted in October 2000. Evidence is found that in both regimes, the Bank of the Republic exhibits behavior that is diametrically opposed in terms of the way in which the intervention interest rate reacts to changes in the inflation rate and output gap. In one regime, which prevails before October 2000 and is labeled regime with no inflation targeting, the Bank of the Republic implements a monetary policy where the intervention interest rate reacts only statistically to changes in the output gap. In the other regime, which prevails after October 2000 and is labeled as a regime with inflation targeting, the intervention interest rate reacts only statistically to changes in the inflation rate. Moreover, we find that in the regime with inflation targeting, Taylor's principle can be, but is not necessarily, satisfied statistically.

Our analytical results show that an inflation rate that has a unit root or a stochastic trend should not satisfy the Taylor principle. Moreover, this analytical argument confirms the insight of Murray *et al.* (2008) who argue that Taylor's principle is associated with a central bank's behavior consistent with inflation that follows a stationary path over time. Hence, a unit root test is carried out for a Markov-switching model to study the relationship between Taylor's principle and the stationarity of the inflation rate. The Dickey-Fuller unit root test is carried out in a two-state Markov-switching model given the results that were obtained when the optimal Taylor rule was estimated. After October 2000, the test shows that the regime that is predominant has an inflation rate that is weakly stationary, while before October of 2000 the predominant regime has an inflation rate with unit root behavior. This additional empirical evidence provides greater certainty that the inflation targeting regime generated a policy that very likely satisfies the Taylor principle. These results support the idea that an inflation targeting policy is beneficial to small open countries like Colombia because it enables central banks to control inflationary shocks by maintaining the inflation rate at low rates in a credible way.

REFERENCES

Ball, L. (1999a), “Efficient rules for monetary policy,” *International Finance* **2**(1): 63–83.

———. (1999b), “Policy rules for open economies,” in J.B. Taylor, ed., *Monetary Policy Rules*, NBER Studies in Business Cycles, Chicago: NBER-University of Chicago Press.

Banco de la República (1999), *Informe de la Junta Directiva al Congreso de la República* (Report of the Board of Directors to the Congress of the Republic), No. 2, July.

Barro, R.J. and D. Gordon, (1983), “A positive theory of monetary policy in a natural rate model,” *Journal of Political Economy* **91**(4): 589–610.

Batini, N., R. Harrison, and S.P. Millard (2003), “Monetary policy rules for an open economy,” *Journal of Economic Dynamics and Control* **27**(11/12): 2059–94.

Batini, N., P. Levine, and J. Pearlman (2009), “Monetary rules in emerging economies with financial market imperfections,” in J. Galí and M. Gertler, eds., *International Dimensions of Monetary Policy*, National Bureau of Economic Research.

Bernal, G. and J. Tautiva (2011), “*Datos en tiempo real: una aplicación a la regla de Taylor en Colombia*” (Real-time data: An application of Taylor’s rule in Colombia), *Revista de Economía Institucional* **13**(24): 373–94.

Bernanke, B. and M. Gertler (2000), “Monetary policy and asset price volatility,” Working Paper 7559, National Bureau of Economic Research.

———. (2001), “Should central banks respond to movements in asset prices?” *American Economic Review (Papers and Proceedings of the 113th Annual Meeting of the American Economic Association)* **91**(2): 253–257.

Camacho, M. (2010), “Markov-switching models and the unit root hypothesis in real U.S. GDP,” *Economics Letters* 112(2): 161–64.

Carlstrom, C.T. and T.S. Fuerst (2007), “Asset prices, nominal rigidities, and monetary policy,” *Review of Economic Dynamics*, **10**(2): 256–75.

Clarida, R., J. Gali, and M. Gertler (1998), “Monetary policy rules in practice: Some international evidence,” *European Economic Review* **42**: 1033–67.

———. (1999), “The science of monetary policy: A new Keynesian perspective,” *Journal of Economic Literature* **37**(4): 1661–707.

———. (2000), “Monetary policy rules and macroeconomic stability: Evidence and some theory,” *Quarterly Journal of Economics* **115**(1): 147–80.

Diebold, F. X., J. Lee and G. C. Weinbach (1993), “Regime switching with time-varying transition probabilities,” Working Papers 93-12, Federal Reserve Bank of Philadelphia.

Diebold, F.X., J. Lee, and G.C. Weinbach (1994). “Regime Switching with Time-Varying Transition Probabilities,” in Hargreaves, Colin,P. ed., Nonstationary Time Series Analysis and Cointegration. New York and Oxford: Oxford University Press, pp. 283-302.

Diebold, F.X., J. Lee, and G.C. Weinbach (1999), "Regime switching with time-varying transition probabilities," in F.X. Diebold and G.D. Rudebusch, eds., *Business Cycles: Durations, Dynamics, and Forecasting*, Princeton: Princeton University Press.

Friedman, B.M. (1990), "Targets and instruments of monetary policy," in *Handbook of Monetary Economics*, Vol. 2, Chapter 22: 1185–230.

Friedman, M. (1968), "The role of monetary policy," *American Economic Review* **58**(1): 1–17.

Giraldo, A. (2008), "Aversión a la inflación y regla de Taylor en Colombia 1994-2005" (Aversion to inflation and Taylor's rule in Colombia 1994-2005), *Cuadernos de Economía* (Universidad Nacional de Colombia) **XXVII**(49): 229–62.

Giraldo, A., M. Misas, and E. Villa (2011), "Reconstructing the recent monetary policy history of Colombia from 1990 to 2010," *Ensayos sobre Política Económica, Banco de la República*. Publication pending. An initial version is available at http://www.dotec-colombia.org.

Hall, S.G., Z. Psaradakis, and M. Sola (1999), "Detecting periodically collapsing bubbles: A Markov switching unit root test," *Journal of Applied Econometrics* 14: 143–54.

Hamilton, J.D. (1994), *Time Series Analysis*, Princeton: Princeton University Press.

Krolzig, H.M. (1997), "Markov-switching vector autoregression." Lecture Notes in Economic and Mathematical Systems.

Kuzin, V. (2006), "The inflation aversion of the Bundesbank: A state space approach," *Journal of Economic Dynamics and Control* 30: 1671–86.

Kydland, F.E. and E.C. Prescott (1977), "Rules rather than discretion: The inconsistency of optimal plans," *Journal of Political Economy* **85**(June):473–92.

López, M. (2004), "Efficient policy rule for inflation targeting in Colombia," *Ensayos sobre Política Económica* (Banco de la República) **45**: 80–115.

Mishkin, F. (1999), International experiences with different monetary policy regimes," *Journal of Monetary Economics* **43**(3): 579–605.

Murray, C., A. Nikolsko-Rzhevskyy, and D. Papell (2008), "Inflation persistence and the Taylor principle," MPRA Paper 11353, University Library of Munich, Germany.

Orphanides, A. (2000), "Activist stabilization policy and inflation: the Taylor rule in the 1970s," Finance and Economics Discussion Series 2000-13, Board of Governors of the Federal Reserve System.

Orphanides, A. (2008), "Taylor rules," in E.N. Durlauf and L.E. Blume, eds., The New Palgrave Dictionary of Economics Online. Palgrave Macmillan, 2nd Edition.

Pérez, J. (2005), "Evaluación de reglas de tasa de interés en un modelo de economía pequeña y abierta" (Evaluation of interest rate rules in a small and open model economy), *Borradores de Economía* (Banco de la República), 385.

Rudebusch, G. and L.E.O. Svensson (1999b), "Policy rules for inflation targeting," in J.B. Taylor, ed., *Monetary Policy Rules*, NBER Studies in Business Cycles, Chicago: NBER-University of Chicago Press.

Svensson, L. (1997), Inflation forecast targeting: Implementing and monitoring inflation targets," *European Economic Review* **41**: 1111–46.

———. (1999a), "Inflation targeting as a monetary policy rule," *Journal of Monetary Economics* **43**(3): 607–54.

———. (1999b), "Inflation targeting: Some extensions," *Scandinavian Journal of Economics* **101**(3): 337–61.

———. (2000), "Open-economy inflation targeting," *Journal of International Economics* **50**(1): 155–83.

———. (2008), "Inflation targeting," in E.N. Durlauf and L.E. Blume, eds., The New Palgrave Dictionary of Economics Online, 2nd Edition. Palgrave Macmillan.

———. (2011), "Inflation targeting," in B.M. Friedman and M. Woodford, eds., *Handbook of Monetary Economics*, Elsevier B.V.

Svensson, L. and M. Woodford (2003), "Implementing optimal policy through inflation-forecast targeting," NBER Working Paper 9747, National Bureau of Economic Research.

Taylor, J.B. (1993), "Discretion versus policy rules in practice," Carnegie-Rochester Conference Series on Public Policy **39**(Dec.): 195–214.

———. (1999a), "A historical analysis of monetary policy rules," in Taylor, J. B., editor, *Monetary Policy Rules*, NBER Studies in Business Cycles Vol. 31, Chicago: NBER-University of Chicago Press.

———. (1999b), Monetary Policy Rules, NBER Studies in Business Cycles Vol. 31, National Bureau of Economic Research.

———. (2001), "The role of the exchange rate in monetary-policy rules," *The American Economic Review* **91**(2): 263–7.

Walsh, C. (2003), *Monetary Theory and Policy, 2nd Edition*. Cambridge: MIT Press.

———. (2010), *Monetary Theory and Policy, 3rd Edition*. Cambridge: MIT Press.

Weinbach, G.C. (1993), "Regime switching with time-varying transition probabilities: Methodological issues and application to exchange rates." doctoral dissertation, University of Pennsylvania.

Woodford, M. (2001), "The Taylor rule and optimal monetary policy," *American Economic Review* **91**(2): 232–37.

———. (2003a), *Interest and prices: Foundations of a theory of monetary policy*, Princeton: Princeton University Press.

———. (2003b), "Optimal interest-rate smoothing," *Review of Economic Studies* **70**: 861–86.

APPENDIX A

Derivation of the optimal Taylor rule for an open economy

Consider the following Lagrangian function:

$$\min_{(\pi,x,i,e)} E_t\left\{\sum_{t=0}^{\infty}\beta^t\left\{\frac{L_t}{2}+\phi_1\left[x_t-x_{t+1}-\delta x_{t-1}+\sigma\left(i_t-\pi_{t+1}-r^n\right)-\alpha_1 e_t-\gamma^* x_t^*-\varepsilon_{1t}\right]\right.\right.$$
$$\left.\left.+\phi_2\left[\pi_t-\beta\pi_{t+1}-\gamma\pi_{t-1}-kx_t-\alpha_2 e_t-\varepsilon_{2t}\right]\right\}\right\}$$

which corresponds to the problem of minimizing (11) subject to (2) and (3) where we assume $\phi_{jt}\equiv\phi_j\neq 0$ for all t and $j=1,2$. The loss function is a quadratic function and the restrictions are linear, so the following first-order conditions are necessary and sufficient to resolve the optimization problem.

$$\begin{aligned}\left(\pi_t\right):&\quad -\beta^{t+1}\left[\phi_2\gamma\right]+\beta^t\left[\pi_t-\overline{\pi}+\varphi_2\right.\\ &+\lambda_3\left(\left(i_{t-1}-i_{t-1}^*\right)-\left(\pi_t-\pi_t^*\right)-\varphi_{t-1}\right)-\beta^{t-1}\left[\sigma\phi_1+\phi_2\beta\right]\Big]=0\end{aligned}\tag{A1}$$

$$\left(x_t\right):\quad -\beta^{t+1}\left[\phi_1\delta\right]+\beta^t\left[\lambda_1 x_t+\phi_1-k\phi_2\right]-\beta^{t-1}\phi_1=0\tag{A2}$$

$$\begin{aligned}\left(i_t\right):&\quad \beta^t\left[\lambda_4\left(i_t-i_{t-1}\right)+\phi_1\sigma\right]+\beta^{t+1}\Big[\lambda_3\left(\left(i_{t-1}-i_{t-1}^*\right)\right.\\ &\left.-\left(\pi_t-\pi_t^*\right)-\varphi_{t-1}\right)-\lambda_4\left(E_t i_{t+1}-i_t\right)\Big]=0\end{aligned}\tag{A3}$$

$$\left(e_t\right):\quad \beta^t\left[\lambda_2\left(e_t-e^n\right)-\phi_1\alpha_1-\phi_2\alpha_2\right]=0\tag{A4}$$

From (A1) and (A2) we solve for ϕ_2. We then equalize both equations to solve for ϕ_1 as:

$$\phi_1\equiv m_1\left[\left(\pi_t-\overline{\pi}\right)+\lambda_3\left(\left(i_{t-1}-i_{t-1}^*\right)-\left(\pi_t-\pi_t^*\right)-\varphi_{t-1}\right)\right]+n_1 x_t\tag{A5}$$

where m_1 and n_1 are functions of the parameters given by

$$m_1 \equiv \frac{\beta k}{\sigma k - \beta\gamma\left[1-\beta\left(1-\beta\delta\right)\right]}; \quad n_1 \equiv \frac{-\beta^2\gamma\lambda_1}{\sigma k - \beta\gamma\left[1-\beta\left(1-\beta\delta\right)\right]} \tag{A6}$$

and where we assume that the parameters satisfy

$$\sigma k \neq \beta\gamma[1-\beta(1-\beta\delta)] \tag{A7}$$

Returning to find ϕ_2 we obtain

$$\phi_2 = m_2\left[\left(\pi_t - \overline{\pi}\right) + \lambda_3\left(\left(i_{t-1} - i^*_{t-1}\right) - \left(\pi_t - \pi^*_t\right) - \varphi_{t-1}\right)\right] + n_2 x_t \tag{A8}$$

where

$$m_2 \equiv -\frac{1-\beta\left(1-\beta\delta\right)}{\sigma k - \beta\gamma\left[1-\beta\left(1-\beta\delta\right)\right]}; \quad n_2 \equiv \frac{\sigma\lambda_1}{\sigma k - \beta\gamma\left[1-\beta\left(1-\beta\delta\right)\right]} \tag{A9}$$

Now, from (A3) we can solve for ϕ_1 which when equalized with the expression in (A5) and reorganized, generates:

$$\begin{aligned}\pi_t - \overline{\pi} = {} & \frac{\beta\lambda_4\left[E_t i_{t+1} - i_t - \lambda_3\left(i_{t-1} - i^*_{t-1} - \left(\pi_t - \pi^*_t\right) - \varphi_{t-1}\right)\right]}{\sigma m_1} \\ & - \frac{\lambda_4\left(i_t - i_{t-1}\right)}{\sigma m_1}\end{aligned} \tag{A10}$$

From (A4) we can solve for e_t, in which the values of ϕ_1 and ϕ_2 are replaced to obtain

$$\begin{aligned}e_t = {} & e^n + \frac{\alpha_1 m_1 + \alpha_2 m_2}{\lambda_2}\left[\pi_t - \overline{\pi} + \lambda_3\left(i_{t-1} - i^*_{t-1} - \left(\pi_t - \pi^*_t\right) - \varphi_{t-1}\right)\right] \\ & + \frac{\alpha_1 n_1 + \alpha_2 n_2}{\lambda_2} x_t\end{aligned} \tag{A11}$$

Replacing $\pi_t - \overline{\pi}$ from Equation (A10) in (A11) we obtain a simple expression that enables reduction of the four first-order conditions (A1), (A2), (A3) and (A4) in a single equation without the endogenous variables ϕ_1 and ϕ_2:

$$
\begin{aligned}
e_t = e^n + \frac{\alpha_1 m_1 + \alpha_2 m_2}{\lambda_2}\Bigg[\Bigg(&\frac{\beta\lambda_4(E_t i_{t+1} - i_t - \lambda_3(i_{t-1} - i^*_{t-1})}{\sigma m_1} \\
&+ \frac{\beta\lambda_4(-(\pi_t - \pi^*_t) - \varphi_{t-1}) - \lambda_4(i_t - i_{t-1})}{\sigma m_1}\Bigg) \\
&+ \lambda_3((i_{t-1} - i^*_{t-1}) - (\pi_t - \pi^*_t) - \varphi_{(t-1)}\Bigg] + \frac{\alpha_1 n_1 + \alpha_2 n_2}{\lambda_2} x_t
\end{aligned}
\tag{A12}
$$

From Equation (3) we can solve for $E_t\pi_{t+1}$, which is inserted in (2) and then inserted in (A12) which generates the *structural* optimal Taylor rule that has as its reduced form Equation (12):

$$
\begin{aligned}
i_t = \; & \frac{1}{\Gamma}\left[r^n + \left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)e^n\right] + \frac{E_t x_{t+1}}{\sigma\Gamma} + \frac{1}{\Gamma}\left\{\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)\right. \\
& \left.\left[\frac{\alpha_1 n_1 + \alpha_2 n_2}{\lambda_2} - \frac{\alpha_1 m_1 + \alpha_2 m_2}{\lambda_2}\frac{n_1}{m_1}\right] - \left(\frac{k}{\beta} + \frac{1}{\sigma}\right)\right\} x_t \\
& + \frac{\delta}{\sigma\Gamma} x_{t-1} + \left(\frac{1}{\beta\Gamma} + \frac{\lambda_3(\beta + \sigma m_1)}{\Gamma}\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)\frac{\alpha_1 m_1 + \alpha_2 m_2}{\sigma m_1 \lambda_2}\right)\pi_t \\
& - \frac{\gamma}{\beta\Gamma}\pi_{t-1} + \frac{\delta^*}{\sigma\Gamma}x^*_t - \frac{\lambda_3(\beta + \sigma m_1)}{\Gamma}\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)\frac{\alpha_1 m_1 + \alpha_2 m_2}{\sigma m_1 \lambda_2}\pi^*_t \\
& + \frac{\lambda_3(\beta + \sigma m_1)}{\Gamma}\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)\frac{\alpha_1 m_1 + \alpha_2 m_2}{\sigma m_1 \lambda_2} i^*_{t-1} \\
& + \frac{\lambda_3(\beta + \sigma m_1)}{\Gamma}\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)\frac{\alpha_1 m_1 + \alpha_2 m_2}{\sigma m_1 \lambda_2}\varphi_{t-1} \\
& + \frac{\beta\lambda_4}{\Gamma}\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)\frac{\alpha_1 m_1 + \alpha_2 m_2}{\sigma m_1 \lambda_2} E_t i_{t+1} \\
& + \frac{\lambda_4 - \lambda_3(\beta + \sigma m_1)}{\Gamma}\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)\frac{\alpha_1 m_1 + \alpha_2 m_2}{\sigma m_1 \lambda_2} i_{t-1} + \frac{\varepsilon_{1t}}{\sigma\Gamma} - \frac{\varepsilon_{2t}}{\beta\Gamma}
\end{aligned}
\tag{A13}
$$

where structural parameters are shown explicitly associated with the variables and the parameter Γ is defined as follows:

$$
\Gamma \equiv \left[1 + \left(\frac{\alpha_1 m_1 + \alpha_2 m_2}{\sigma m_1 \lambda_2}\right)\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)\lambda_4\left(1 + \beta\right)\right]
$$

Linear dynamic system

We can rewrite the model recursively as a discrete linear dynamic system to show the existence and uniqueness of a steady state. Under $\varepsilon_{1t} = \varepsilon_{2t} = 0$ the model can be written as a second-order, non-homogeneous linear system as follows:

$$A_0 Z_{t+1} = AZ_t + BZ_{t-1} + C \tag{A14}$$

where A_0, A, B are 8×8 matrices while C is an 8×1 matrix, and we define

$$Z_{t+1} = \begin{bmatrix} E_t x_{t+1} \\ E_t \pi_{t+1} \\ E_t i_{t+1} \\ E_t e_{t+1} \\ E_t x^*_{t+1} \\ E_t \pi^*_{t+1} \\ E_t i^*_{t+1} \\ E_t \varphi_{t+1} \end{bmatrix}.$$

A convenient way to analyze the system is to transform it into a non homogenous first order linear system. For this, define $Z_{t+1} \equiv Y_t$ such that $Z_{t+2} \equiv Y_{t+1}$ and we can rewrite the second-order system (A14) as

$$HW_{t+1} = MW_t + N \tag{A15}$$

where $W_t = \begin{bmatrix} Y_t \\ Z_t \end{bmatrix}, H = \begin{bmatrix} A_0 & 0 \\ 0 & I_8 \end{bmatrix}, M = \begin{bmatrix} A & B \\ I_8 & 0 \end{bmatrix}, N = \begin{bmatrix} C \\ 0 \end{bmatrix}.$

Note that I_8 is an 8×8 identity matrix, W is a 16×1 vector, H, and M are 16×16 matrices and N is a 16×1 vector.

Existence and uniqueness of the steady state

Proposition 3. The dynamic system (A15) has a unique steady state if $\sigma k = 1 - \beta(1 - \beta\delta)$, $1 > \beta + \gamma$, $\gamma \in (-1,1)$, $0 > \alpha_1 \beta^2 \gamma > \alpha_2 \sigma$, $\lambda_2 \leq \beta \lambda_3^2$ and

$$\lambda_1 < min\left\{\frac{k^2\left[\lambda_3\left(2-\beta\gamma\right)\right]+1}{\beta\gamma\left(1-\beta-\gamma\right)}, \frac{k\left(1-\beta\delta\right)\left(1-\beta\right)\left(\lambda_3\left(2-\beta\gamma\right)+1\right)}{\left(1-\gamma\right)\sigma\beta\gamma}\right\}$$

Proof: To prove this result, note that in steady state we have $W_t = \bar{W}$ for all t in (A15). Therefore, to prove the existence and uniqueness of the steady state of the dynamic system, we must find the conditions that guarantee that $[H - M]^{-1}$ exists such that a steady state exists in the sense $\bar{W} = [H - M]^{-1}N$ exists. A sufficient condition such that $\bar{W}$ exists is that $det[H - M] \equiv |H - M| \neq 0$. Note that $H - M = \begin{bmatrix} A_0 & 0 \\ 0 & I_8 \end{bmatrix} - \begin{bmatrix} A & B \\ I_8 & 0 \end{bmatrix} = \begin{bmatrix} A_0 - A & -B \\ -I_8 & I_8 \end{bmatrix}$.

Through the formula of the determinant for a partitioned[24] matrix it follows that

$$|H - M| = |I_8| \cdot |A_0 - A - B|$$

Therefore, we consider developing the determinant of $|A_0 - A - B|$.

This determinant can be obtained using the cofactors method on the rows or columns with the greatest quantity of zeros which leads to obtaining $|A_0 - A - B| = G \cdot |S|$, where G is the scalar given by

$$G \equiv (\rho_{\varphi,1} + \rho_{\varphi,2} - 1)(\rho_{x,1} + \rho_{x,2} - 1)(\rho_{\pi,1} + \rho_{\pi,2} - 1).$$

and $S \equiv [\ s_1 \ \ s_2 \ \ s_3 \ \ s_4\ ]$, is a 4×4 matrix where

$$s_1 = \begin{bmatrix} \delta - k\dfrac{\sigma}{\beta} \\ \dfrac{k}{\beta} \\ -\dfrac{\sigma n_1}{\beta\lambda_4} \\ -\dfrac{1}{\lambda_2}\left(n_1\alpha_1 + n_2\alpha_2\right) \end{bmatrix}, \quad s_2 = \begin{bmatrix} \dfrac{\sigma}{\beta}\left(1-\gamma\right) \\ 1 - \dfrac{1}{\beta}\left(1-\gamma\right) \\ -\dfrac{\sigma m_1}{\beta\lambda_4}\left(\lambda_3\left(\dfrac{1}{\sigma}\dfrac{\beta}{m_1}\right)+1\right) \\ -\left(\begin{matrix} m_1\alpha_1 \\ +m_2\alpha_2 \end{matrix}\right)\left(-\dfrac{\lambda_3}{\lambda_2}\right) \end{bmatrix},$$

24. Let A be a matrix that can be partitioned into four submatrices $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$. Hence, the determinant of A is given by $|A| = |A_{22}| \cdot \left|A_{11} - A_{12}[A_{22}]^{-1}A_{21}\right|$.

$$
s_3 = \begin{bmatrix} -\sigma \\ 0 \\ \dfrac{\sigma m_1}{\beta\lambda_4}\left(-\lambda_3 + \dfrac{\lambda_4 - \beta\lambda_3}{\sigma m_1}\right) - \dfrac{\beta+1}{\beta} + 1 \\ 1 \end{bmatrix}, s_4 = \begin{bmatrix} -\alpha_1 - \dfrac{\sigma\alpha_2}{\beta} \\ \dfrac{\alpha_2}{\beta} \\ 0 \\ 1 \end{bmatrix}
$$

$$
\begin{aligned}
|S| = {} & \left(\frac{\sigma\alpha_2}{\beta(1-\beta\gamma)\lambda_4}\right)\Bigg[\left(\frac{\beta^2\gamma\lambda_1}{k(1-\beta\gamma)}\right)\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)\left(\frac{1}{\beta\lambda_3} - \frac{\lambda_3}{\lambda_2}\right) \\
& + \left(\frac{\lambda_1}{k\lambda_2}\right)\left(\frac{\alpha_2\sigma - \alpha_1\beta^2\gamma}{\sigma(1-\beta\gamma)}\right)\left(\lambda_3(2-\beta\gamma)+1\right)\Bigg] \\
& + \Bigg[\frac{\sigma}{\beta(1-\beta\gamma)\lambda_4} + \left(\alpha_1 + \frac{\sigma\alpha_2}{\beta}\right)\left(\frac{\lambda_3}{\lambda_2(1-\beta\gamma)^2\lambda_4}\right)\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)\Bigg] \\
& \Bigg[k\left(\lambda_3(2-\beta\gamma)+1\right) - (1-\beta-\gamma)\left(\frac{\beta\gamma\lambda_1}{k}\right)\Bigg] \\
& + \left(\frac{\lambda_3}{\lambda_2}\right)\left(1+\frac{1}{1-\beta\gamma}\right)\left(\frac{\alpha_2}{\beta(1-\beta\gamma)}\right)\Bigg[(1-\beta\gamma)(1-\beta)\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)\left(\frac{1}{\beta\lambda_3} - \frac{\lambda_3}{\lambda_2}\right) \\
& + \left(\frac{\lambda_1}{\lambda_2}\right)\left(\frac{\alpha_2\sigma - \alpha_1\beta^2\gamma}{\beta}\right)(1-\gamma)\Bigg] + \left(\frac{\alpha_2\lambda_3}{\lambda_4(1-\beta\gamma)^2\beta\lambda_2}\right)\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right) \\
& \Bigg[\left(\frac{\sigma\beta\gamma\lambda_1}{k}\right)(1-\gamma) - (1-\beta\delta)(1-\beta)\left(\lambda_3(2-\beta\gamma)+1\right)\Bigg] \\
& + \left(\frac{\lambda_3}{\lambda_4(\beta)^2}\right)\left(1+\frac{1}{1-\beta\gamma}\right)\big[(1-\beta\delta)(1-\beta)(1-\beta-\gamma) - \\
& (1-\beta(1-\beta\delta))(1-\gamma)\big] - \left(\frac{\sigma\lambda_3}{(1-\beta\gamma)\lambda_4}\right)\left(1+\frac{1}{1-\beta\gamma}\right)\left(\frac{\alpha_1}{\sigma} + \frac{\alpha_2}{\beta}\right) \\
& \Bigg[k\left(\frac{\alpha_1}{\sigma} - \frac{\alpha_2}{\beta}\right)\left(\frac{1}{\beta\lambda_3} - \frac{\lambda_3}{\lambda_2}\right) + \left(\frac{\lambda_1}{\beta k\lambda_2}\right)\left(\frac{\alpha_2\sigma - \alpha_1\beta^2\gamma}{\sigma}\right)(1-\beta-\gamma)\Bigg]
\end{aligned}
\tag{A16}
$$

Under the assumptions given by Equations (8) and (9) in the set-up of the model we obtain $G \neq 0$ and therefore we only need to find the conditions that yield $|S| \neq 0$. Thus, we use the cofactors method to develop the determinant of S by its third column s_3 which yields the determinant of S in Equation (A16) after replacing the constants m_i and n_i for $i = 1,2$ from equations (A6) and (A9) as well as heavy manipulation. All terms are strictly negative under the proposition's assumptions; therefore we conclude that the determinant of S is negative and hence different from zero, implying that there is a steady state of the linear dynamic system (A15). The same linearity of the system implies that the steady state is unique.

APPENDIX B

Unit root test for inflation

Equation (B1) bellow presents, for the inflation rate, the auxiliary regression of the augmented Dickey-Fuller test, under a Markov switching process. In this case, the coefficients and variances are state dependent; in other words, they are governed by a stochastic and unobserved state variable $S_t \in \{0,1\}$:

$$\Delta\pi_t = \rho_{S_t}\pi_{t-1} + C_{S_t} + \theta_{S_t} t + \sum_{j=1}^{K}\eta_{jS_t}\Delta\pi_{t-j} + v_t \tag{B1}$$

This state variable follows a Markov order process, one whose probabilities of transition are defined by: $Pr[S_t = 0|S_{t-1} = 0] = p$, $Pr[\mathrm{S}_t = 1|S_{t-1} = 1] = q$, $Pr[\mathrm{S}_t = 1|S_{t-1} = 0] = 1 - p$ and $Pr[\mathrm{S}_t = 0|S_{t-1} = 1] = 1 - q$.

The unit root test is based on the t statistic corresponding to the coefficient ρ_{S_t} associated with the variable π_{t-1}. The null hypothesis is the existence of the unit root, that is, H_0: $\rho_{S_t} = 0$ versus the alternative H_1: $\rho_{S_t} < 0$, that is, a stationary process. It should be pointed out that this statistic is calculated as the ratio of the estimated coefficient and its standard deviation, derived from the negative inverse of the Hessian matrix of the likelihood function evaluated at the maximum (see Camacho, 2010 and Hall *et al.*, 1999). The construction of the unit root test for the existence of a unit root in the inflation rate, where there are two economic regimes during the period of study, requires taking the following steps:

1) Equation (B1) must be estimated under the null hypothesis and its disturbances must be grouped into two exclusive subsets. The subset assignment scheme is carried out through the filtered probabilities of transition.

a) Subset $\mathbf{A}_1$ corresponds to the residuals associated with regime 0.

b) Subset $\mathbf{A}_2$ corresponds to the residuals associated with regime 1.

where $\mathbf{A}_1 \cap \mathbf{A}_2 = \varnothing$.

2) A number B of vectors $\mathbf{A}^b$ are generated, where $b = 1,\ldots,\mathrm{B}$ are disturbances of the same sample size. With that objective,

we construct subvectors $\mathbf{A}_1{}^b$, and $\mathbf{A}_2{}^b$, where $b = 1,\ldots,$B under sampling with repetition of subsets $\mathbf{A}_1$ and $\mathbf{A}_2$, respectively. Respecting the position of the state associated with the filtered probabilities, the union of the subvectors results in the vector sought out, $\mathbf{A}^b$ for $b = 1,\ldots,$B.

3) A dichotomous state variable $S^*(t)$ is generated that is associated with the states suggested by the series of filtered transition probabilities.

4) B realizations are generated of variable π^b, $b = 1,\ldots,$B using the disturbances in step 2.

For all $b = 1,\ldots,$B

$$\Delta\pi_t^b = C_0\left(1 - S_t^*\right) + C_1 S_t^* + \theta_0\left(1 - S_t^*\right) + \theta_1 S_t^* t$$
$$+\sum_{j=1}^{K}\left[\eta_{j0}\left(1 - S_t^*\right)\Delta\pi_{t-j} + \eta_{j1} S_t^* \Delta\pi_{t-j}\right] + e_t^b$$

5) Equation (B1) is estimated for each of the $\Delta\pi^b{}_t$ and the t-statistics associated with the coefficient that accompanies $\pi^b{}_{t-1}$ are stored in a vector of size B×1. It should be pointed out that the t-statistic is constructed with the standard deviations from the negative of the inverse of the Hessian matrix associated with the optimization procedure that maximizes the likelihood function.

6) The p-value of the unit root test for each state corresponds to the percentage of t statistics found below the original statistic t_ϱ, Equation (B1).

Vol. 51 No. 1 (May, 2014), 85–107

LAJE
Latin American Journal of Economics

HAS THE NATIONAL AGREEMENT FOR THE MODERNIZATION OF BASIC EDUCATION CONTRIBUTED TO IMPROVING LEVELS OF BASIC EDUCATION AND REDUCING DISPARITIES BETWEEN THE STATES?*

Manuel Gómez-zaldívar**

In 1992, Mexico's federal government signed the ANMEB agreement as part of a series of strategic public education reforms. The agreement decentralized the education system, making state governments directly responsible for providing basic public education, in an attempt to reduce marked regional disparities in educational levels. Now that sample sizes are large enough to allow reasonable empirical analysis, I examine several indicators used to measure the characteristics of education in each state. The aim is to assess whether there is sufficient empirical evidence to affirm that the agreement has contributed to improving education levels and reducing disparities among the states.

JEL classification: H4, H75, O1

Keywords: ANMEB, difference-in-differences analysis, Mexican education system

1. Introduction

The average illiteracy level in Mexico prior to the 1910 revolution was alarming: According to the 1910 population census, 72.3% of Mexicans aged 10 or older were unable to read and write. In order to bring down the illiteracy rate, the Constituent Congress of 1917 established that public education was to be free.[1] Nevertheless, this constitutional right failed to achieve the desired effect, and the 1921 population census showed that illiteracy remained relatively high at 66.2%. Therefore, in September of that year the federal government created the Secretariat of Public Education (*Secretaría de Educación Pública*, or SEP). The main objective of this new institution was to reduce illiteracy and increase gross enrollment ratio nationally. To do

* The author gratefully acknowledges the data analysis assistance provided by Lizeth García-Belmonte.

** Department of Economics and Finance, Universidad de Guanajuato, DCEA - Campus Marfil, Fracc. I, El Establo, Guanajuato, Guanajuato, C.P. 3625, Mexico. Telephone/fax: +52 (473) 735-2925, ext. 2831. Email: manuel.gomez@ugto.org.

1. And from 1934 onwards, also mandatory.

doi 10.7764/LAJE.51.1.85

this, the federation implemented a school construction program and made it possible for local governments (state and municipal) to build and operate their own schools, in effect creating a two-tier public education system that functioned independently on two different government levels: one federal, the other local.

However, this two-tier education system did not work as well as hoped. Although from 1921 to 1930 the illiteracy rate dropped by almost 5 percentage points, it is also the case that, in absolute terms, the number of illiterate Mexicans increased. This may have been the result of the heterogeneity of development and income levels among the Mexican states combined with a homogeneous education policy implemented by the federal government. In other words, given that the federation covered the operations and maintenance costs of the schools it built while states and municipalities paid for their own, wealthier local governments built more schools and, at the same time, increased their gross enrollment ratio and raised literacy rates within their states. In 1930, in the states along the country's southern Pacific coast (Guerrero, Oaxaca, and Chiapas) only one in five people aged 10 and older knew how to read and write, whereas in the northeastern states (Coahuila, Nuevo León, and Tamaulipas) almost half of the population was literate.

Efforts to attenuate educational differences among the states have been intense, as evidenced by the secondary education laws passed by Congress at various times. Nevertheless, although education indicators have been improving throughout the time for which data is available (i.e., since 1976), disparities among the states remain. During the 1991–1992 school year, the national illiteracy rate stood at 11.7% and the gross enrollment ratio at the elementary level was 95.4%. In terms of states, the illiteracy rate in Chiapas was 28.5%, whereas in Nuevo León it was just 4.3%. In contrast, the gross enrollment ratio in Guerrero reached 100% and in Tamaulipas 89.9%.

In order to reduce these differences, in May 1992 the federal government and the governors of the 31 Mexican states signed the National Agreement for the Modernization of Basic Education (*Acuerdo Nacional para la Modernización de la Educación*, or ANMEB). The decentralization of educational services (abolishing of the two-tier system established in 1921) was intended, at least in part, to eliminate such differences. In order to achieve this goal, the SEP would, in conjunction with the states, take all actions necessary to reduce and overcome disparities, paying particular attention to those regions with the greatest deficits in terms of enrollment rates and educational achievement. The agreement also established that the Federation would guarantee that more resources would be allocated to those states with economic limitations and more pressing educational deficiencies.

The conceptual framework for addressing the questions of how and why the agreement might help to improve the quality of education and reduce inter-state inequalities in Mexico is complex, primarily due to the extensive changes that the reform entailed.[2] Nevertheless, I summarize below the three main strategies and explain the channels through which this development was expected to improve basic education.

First, the central aim of the agreement was to raise the standards for teacher training, as the educational level of teachers has been seen as one of the primary drawbacks of the Mexican education system. The reform sought to achieve this goal through: i) a revised curriculum and revision of the courses studied in teacher-training programs; ii) new in-service programs for all teachers, principals, and supervisors; iii) the creation of a more effective system for assessing teacher-training programs; iv) the creation of a single teacher-training system; and v) the development of a radio and TV-based "distance training" program for teachers in rural and indigenous schools.

Second, a merit pay system was developed to link professional performance to salaries. The career ladder, known as the *carrera magisterial*, was expected to help raise educational levels by aligning government and teacher incentives and recognizing and stimulating teacher performance by renewing their interest in ongoing improvement. The measurement of professional performance would take into account experience, professional skills, educational attainment, and the completion of accredited courses.

Third, the reform established technical councils made up of teachers, principals, and supervisors to improve teacher participation in the education process. As stated by Tatto (1999), the goal was "...to promote teachers' analytic and critical views of their own teaching practice and as forums for discussing teaching and learning, curriculum and teacher education." Furthermore, these self-governing mechanisms were expected to develop short- and long-term projects to solve the particular issues of each school.

2. More extensive discussion of the conceptual issues underlying the agreement can be found in Bray (1999), Gershberg (1999), Ornelas (1995, 2008), and Tatto (1999), among others.

With the enactment of the ANMEB, the federal government transferred 100,000 schools (66.3% of the total), 500,000 teachers (66.1%), and almost 13.5 million students (68.5%) to the purview of state governments. The administration and organization of the transferred schools presented a heterogeneous dilemma for state governments. During the period of the two-tier public educational system (1921–1992), some states had highly developed local education systems whereas others had no experience at all with providing such a service. Table 1 shows state participation in public education in the year the ANMEB was signed. As can be seen, states such as the State of Mexico, Nuevo León, Baja California, and Jalisco accounted for over 40% of all the students enrolled in the public system, whereas the provision of education in Tamaulipas, Oaxaca, and Hidalgo depended totally on the federal government.

In addition to being responsible for providing education services, local governments also received resources to cover the teachers' payroll and to maintain and operate the schools they took over after the ANMEB went into effect.

Previous research on the ANMEB has focused mainly on this last issue, i.e., analyzing which factors have determined the allocation of resources to pay for education in the states.[3] The objective of this paper differs somewhat. I am interested in analyzing whether, as a result of the signing of the Agreement, state education indicators have improved and whether regional disparities have been reduced. For this purpose, I apply the difference-in-differences (DD) technique. The results indicate that progress has been made in terms of the indicators, primarily those related to elementary schools. Nevertheless, further analysis suggests that is difficult to consider such improvement a consequence of the ANMEB.

The remainder of the paper is organized as follows: Section 2 briefly describes the decentralization processes of a number of other countries and their relationship to that of Mexico. Section 3 succinctly illustrates the difference-in-differences methodology. In Section 4 I present a description of the variables and their basics statistics. Section 5 contains the empirical analysis of different estimated models. Finally, Section 6 outlines the main conclusions.

3. See Ontiveros (2001), Hecock (2006), and Sharma and Cárdenas (2008). For education financing and the distribution of federal resources linked to the ANMEB, see Cárdenas and Luna (2007) and Latapí and Ulloa (2000).

Table 1. State participation in public education, 1991–1992 school year

(in percent)

	Enrollment	Teachers	Schools
State of Mexico	53.2	53.7	53.8
Nuevo León	45.3	50.3	37.0
Baja California	42.3	47.4	38.1
Jalisco	40.2	41.5	26.7
Sinaloa	32.5	34.1	18.2
Veracruz	32.2	30.7	28.4
Chihuahua	30.8	32.1	25.4
Sonora	30.7	26.8	22.8
Yucatán	28.6	30.0	16.6
Durango	27.7	26.2	24.5
Puebla	26.9	26.4	21.1
Guanajuato	26.9	28.2	20.2
Tlaxcala	23.1	21.2	25.3
Coahuila	22.6	22.2	14.2
Chiapas	22.4	22.9	20.2
Zacatecas	21.4	22.0	25.3
Guerrero	17.9	15.6	18.1
Nayarit	15.5	16.8	17.0
San Luis Potosí	13.3	16.3	19.2
Tabasco	13.0	15.4	15.7
Colima	9.7	12.9	9.9
Campeche	3.5	5.9	3.1
Michoacán	3.4	4.6	8.1
Baja California Sur	3.1	2.9	4.2
Quintana Roo	2.5	3.4	5.7
Aguascalientes	1.6	2.4	5.5
Morelos	1.2	1.0	2.5
Querétaro	0.9	1.3	1.1
Tamaulipas	0.3	0.5	2.2
Oaxaca	0.1	0.1	0.1
Hidalgo	0.0	0.0	0.0

Source: García-Pérez (2008), p. 31.

2. Education decentralization in Mexico and elsewhere

Educational decentralization experiences in countries around the world have been ongoing. These reforms have so many characteristics that it would be difficult to make a clear-cut comparison of all of them; to do so would require analytical frameworks to reduce the dimensionality of such reforms and enable a reasonably fair comparison.

One attempt at such a comparison was made by Tatto (1999) for the purpose of analyzing the Mexican reform. Her approach has two dimensions: the first is characterized by two different methods of teaching, i.e., didactic/routine vs. interactive/conceptual; the second is characterized by two different authority structures, i.e., formal versus organic control. Using this methodology, Tatto compares teaching dynamics in Mexico with those at schools in Brazil, China, France, Japan, and the United States. She concludes that the characteristics of Mexico's educational reform caused teaching to shift from being didactic/routine in nature towards being interactive/conceptual. Furthermore, the reform is expected to change teaching practices to make them more organic, provided teachers are able to work together and see themselves as actors playing a key role in finding solutions to education problems through their own practices.

This paper does not present a general analytical framework to contrast the decentralization experience in different countries, nor to compare different theories of what makes for more or less "successful" education reforms. Nevertheless, I believe that it is important to point out the similarities and differences between Mexico's decentralization experience and analogous reforms in other countries.

According to the Inter-American Development Bank, or IDB (1994), almost every country in Latin America implemented some sort of education decentralization policy during the 1980s and early 1990s. The diversity of the reforms is broad, as are the strategies with which they were executed. Moreover, implementation of such reforms has been influenced by the political and economic context in which they occurred, and there has been little research carried out to evaluate their success. The following are studies that relate the Mexican experience to that of other countries, and studies connected with the work I propose.

Gershberg (1999) analyzes the costs and benefits of two alternative methods of implementing education reforms: the first is to enact legislation to define and support the reforms, and the second is to implement those changes without any legal framework. Specifically, he compares Mexico to Nicaragua,[4] arguing that given Mexico's size and the power of its national teachers' union, the strategy implemented by the government—i.e., creating a legislative basis for the reform first—is more appropriate for that particular country. The Nicaraguan strategy, on the other hand, mitigates some of the pitfalls associated with the legislative approach by fostering citizen participation, giving a great deal of power to parents and local stakeholders. He concludes that countries that want to apply education decentralization reforms should use a combination of both strategies in order to achieve better results.

Faguet and Sánchez (2006) study the impact of the decentralization of education funding by evaluating a number of education statistics in Bolivia and Colombia. In Bolivia, they find evidence that after decentralization, investment in education became more responsive to local needs, especially in rural areas. Although they are unable to make a formal comparison between the situation before and after the program due to a lack of data, they find improvements in class enrollment in the post-reform period. In the case of Colombia, the availability of data made it possible to study school enrollment at a municipal level. The authors argue that in municipalities in which educational financing and policymaking are most free from central influence, enrollment increased. They suggest that it would have been more interesting to study other variables besides enrollment, such as standardized test results, but that data limitations make this unfeasible.

Lane and Murray's (1985) description of the policy of education decentralization in Sweden highlights one characteristic that made it similar to the policy followed in Mexico. Both countries' reforms were intended to strengthen local government participation by transferring central decisions and responsibilities to regional and local state bodies. Nevertheless, the Swedish reform included universities and colleges, whereas Mexico's reform included only primary and middle schools. In general, we can say that both reforms are similar in that their goals were determined by central authorities but how these reforms were achieved was decided at the local level.

4. According to Gershberg, Mexico followed the former strategy while Nicaragua followed the latter.

3. Methodology: difference-in-differences analysis

The DD approach has been widely used to analyze the effects of policy changes. This procedure helps to examine the effect of some sort of "treatment" by comparing the performance of a treatment group to the performance of a control group. In the basic set-up, the researcher analyzes the outcomes of the two groups during two periods of time: before and after the treatment. It is assumed that one of the groups has been exposed to a treatment in the second period but not in the first. The control group is not exposed to the treatment in either period.[5]

If the researcher focuses on the treatment group alone, before and after, in order to infer the consequences of the policy change, an erroneous conclusion may be reached since there may be other factors influencing events at the same time as the treatment. Therefore, the DD methodology utilizes a control group to remove the possible effects of other factors. The implicit assumption is that if there are other factors affecting both groups at the same time they will have the same effect on the treatment as on the control group.

The baseline DD model to be estimated in this study takes the following form:

$$y_{it} = \alpha + \beta d_g + \gamma d_t + \delta(d_g \cdot d_t) + \textstyle\sum_{k=1}^{3} \varphi_k X_{kt} + \varepsilon_{it} \tag{1}$$

Where y_{it} is the value of the variable of interest in state i ($i = 1,2,\ldots, 31$) at time t; d_g is a dummy variable that indicates the group, and which takes the value of zero for the control group and the value of one for the treatment group; d_t is a time dummy that indicates the period of time and takes the value of zero for the pre-treatment period and the value of one for the post-treatment period. The coefficient of interest is δ. This reveals the behavior of the variable of interest, that is, the treatment group after the agreement is implemented; the interaction term $(d_g \cdot d_t)$ takes the value of one when the observation belongs to the treatment group in the post-treatment period. The variables denoted as X's are: real gross state product per capita, percentage of population living in urban areas, and state fiscal independence. They

5. See Meyer (1995) for a detailed explanation of this approach.

are included in the model because they are important characteristics that may influence the level of education provision in the states, especially before the implementation of the ANMEB.

The independent variables employed in model (1) are defined as follows: The time dummy variable, d_t, takes the value of zero for observations up to 1992, the year in which the agreement was signed, and the value of one for the years after 1992. The group dummy variable, d_g, which distinguishes between those states that were exposed to a treatment and those that were not, deserves a more detailed explanation.

The ANMEB is a federal agreement and therefore one that affects all Mexican states at the same time. Strictly speaking, all states receive the public policy treatment; consequently, it should not be possible to separate the states into two distinct groups: control and treatment. Nevertheless, in order to analyze the impact of the ANMEB with this method, I use the following line of reasoning: García-Pérez (2008) calculates the percentage of students who attended public schools operated by state governments in the 1991–1992 school year, the year in which the agreement was signed. This variable can serve as a proxy to measure the "amount of experience" that each state had in administering and providing public education, and therefore, I believe that it can be used to establish the two different groups needed to apply DD. On the one hand, for those states that had little or no experience in providing education services, the agreement would impose a new responsibility, one with which they were unfamiliar. Thus, the agreement represents a change of policy; I place these states in the treatment group. On the other hand, for those states that already offered this service to a high percentage of students, we can regard the agreement as having had little impact and not representing a change of policy, since these states had already assumed responsibility for providing public education. I place these states in the control group. García-Pérez calculates that the range of participation by state governments in public education was wide. Among the states with the most experience were the State of Mexico, Nuevo León, Baja California, and Jalisco, with over 40% of students enrolled in public schools under the control of the state government. Among those states with least experience were Querétaro, Tamaulipas, Oaxaca, and Hidalgo, with less than 1%. In fact, the state government of Hidalgo had no participation, i.e., 100% of the students were enrolled in the federal system.

In addition, there is another issue that needs to be clarified in the construction of d_g, namely, the threshold value of the level of participation of the state governments to be placed in either group.

Since it is difficult or impossible to argue that a specific value of this variable should determine whether to place each state in one of the groups, the analysis is carried out for different levels, i.e., 30% and 40%. In each case, d_g takes the value of zero—the state is added to the control group—if the state has "sufficient" experience in providing education; otherwise, the value of the variable is one.

The unknown coefficients, α, β, and γ, represent the constant term, the effect attributed to the specific group and the time effect, respectively. The purpose of the DD methodology is to obtain a good estimator of δ, $\hat{\delta}$. Equation (1) can be estimated using the data for both groups in both periods of time using ordinary least squares (OLS), by assuming that the error term ε_{it} has the properties ordinarily required.

Determining the expected value of the variable of interest in each of the four groups, which is denoted as $Y_{i,T0}$, $Y_{i,T1}$, $Y_{i,C0}$, and $Y_{i,C1}$, is straightforward, where the subscripts T and C refer to the treatment and control group respectively, and the subscripts 0 and 1 differentiate between the pre- and post-treatment period. The expected values for each group are defined as follows:

$$\begin{aligned} E(Y_{i,T0}) &= \bar{Y}_{T0} = \alpha + \beta \\ E(Y_{i,T1}) &= \bar{Y}_{T1} = \alpha + \beta + \gamma + \delta \\ E(Y_{i,C0}) &= \bar{Y}_{C0} = \alpha \\ E(Y_{i,C1}) &= \bar{Y}_{C1} = \alpha + \gamma \end{aligned} \tag{2}$$

From these results, we observe that an unbiased estimator of δ is defined: [6]

$$\delta = (Y_{T1} - Y_{T0}) - (Y_{C1} - Y_{C0}) \tag{3}$$

The unbiased estimator that assesses the impact of the treatment is defined as the difference in the average response of the treatment group, before and after the treatment, minus the same observed difference in the control group. The name of the method employed is derived from the formula, i.e., the difference of the differences.

6. Note that $\left[E(\bar{Y}_{T1}) - E(\bar{Y}_{T0})\right] - \left[E(\bar{Y}_{C1}) - E(\bar{Y}_{C0})\right] = \alpha + \beta + \gamma + \delta - (\alpha + \beta) - (\alpha + \gamma) + \alpha = \delta.$

Table 2. Description of variables

Variable	Period	Definition
Elementary school dropout rate	1976–2010	Percentage of students who dropped out of elementary school
Cohort survival rate in elementary schools	1981–2010	Total number of students who finished elementary education, divided by total enrollment in first grade five years earlier
Failure rate in elementary schools	1976–2007	Total number of students in elementary schools who did not pass to the next grade, as a percentage of total enrollment in elementary schools
Completion rate in elementary schools	1976–2007	Total number of students who finished elementary school, divided by total population of 12-year-olds
Transition rate in middle schools	1976–2010	Percentage of population who finished elementary school and then enrolled in middle school
Middle-school dropout rate	1976–2010	Percentage of students who dropped out of middle school
Cohort survival rate in middle schools	1978–2010	Total number of students who finished middle school education, divided by total enrollment in seventh grade two years earlier
Gross enrollment ratio in middle schools	1976–2007	Total enrollment in middle schools, divided by total population between 12 and 14 years old, multiplied by 100
Completion rate in middle schools	1976–2007	Total number of students who finished elementary school, divided by total population of 15-year-olds
Average schooling	1976–2010	Average number of years of schooling among population 15 years and older
Illiteracy rate	1980–2010	Percentage of population 15 years and older who are unable to write and to read

Source: Authors' elaboration.
Note: In Mexico, compulsory education consists of six years of primary school (grades one through six), followed by three years of middle school (grades seven through nine).

4. Description of variables and basic statistics

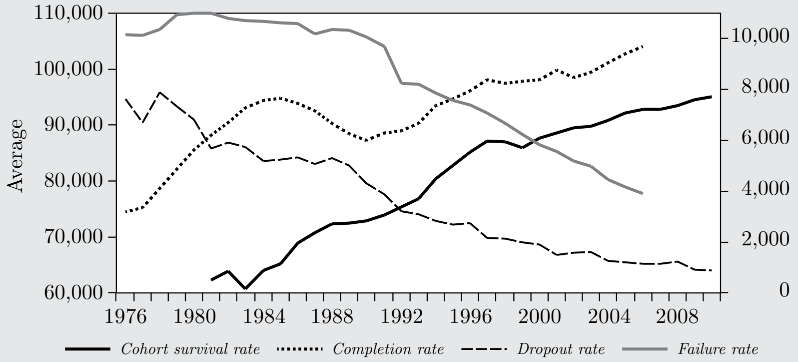
The DD methodology is used to analyze the effect of the ANMEB on the various education indicators in all states in the country. These variables are then used as the dependent variables when estimating model (1). The variables of interest are listed and described in Table 2.[7]

Figures 1 to 6 show the evolution of the mean and standard deviation of the indicators, illustrating how the indicators develop and the evolution

7. Retrieved from http://www.dgpp.sep.gob.mx/Estadi/SistesepPortal/sistesep.html (last accessed on August 10, 2011).

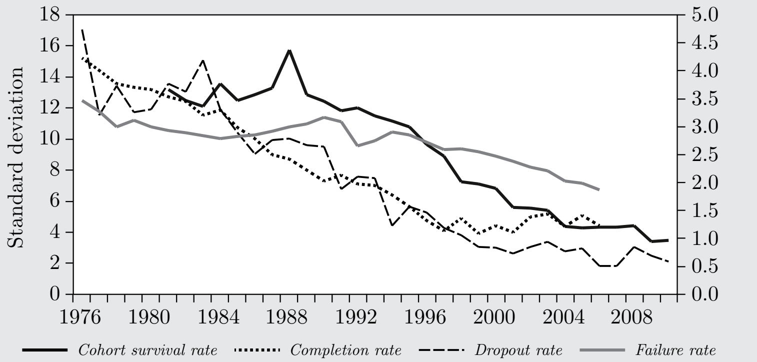
of disparities among the states. According to the figures, every indicator shows improvement as time goes on. Moreover, the standard deviations of the variables tend to decrease over time, which implies that the disparity among states is getting smaller. Since this reduction is perceptible even before the enactment of the agreement, we need more than a graph to be able to determine whether the ANMEB had a significant impact on the observed decline. In general, it seems that the improvement in elementary school indicators is greater than that in middle schools.

Figure 1. Annual mean of elementary-school variables



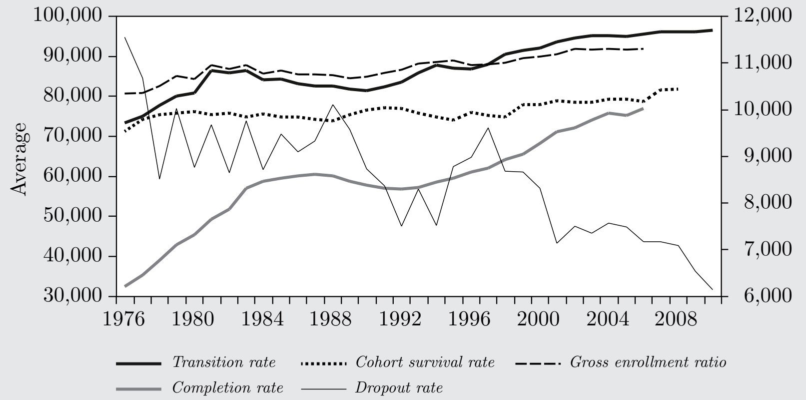
Source: Authors' calculations.
Note: Failure and dropout rates are measured on the right-hand axis.

Figure 2. Annual dispersion of elementary-school variables



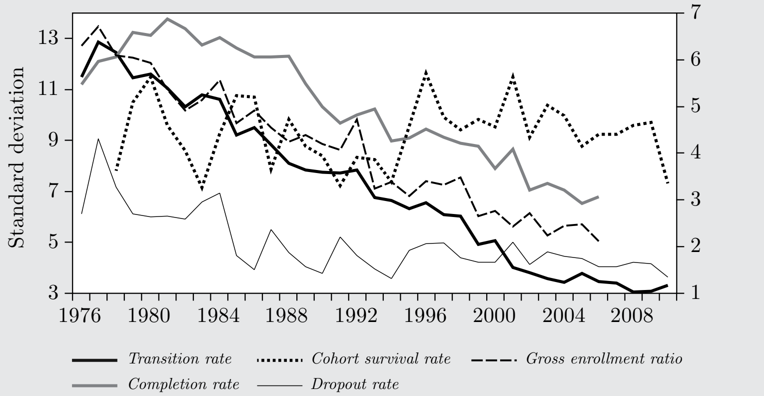
Source: Authors' calculations.
Note: Failure and dropout rates are measured on the right-hand axis.

Figure 3. Annual mean of middle-school variables



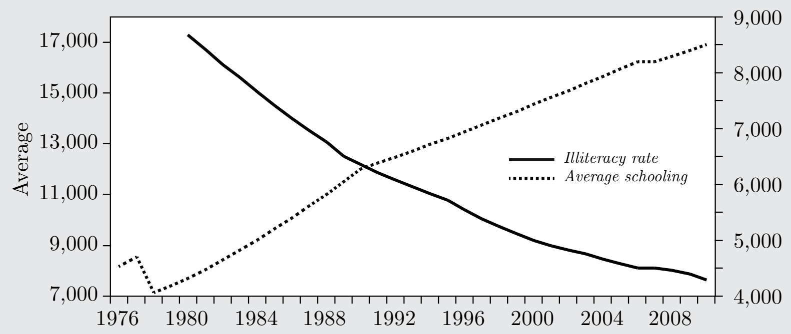
Source: Authors' calculations.
Note: Dropout rate and gross enrollment ratio are measured on the right-hand axis.

Figure 4. Annual dispersion of middle-school variables



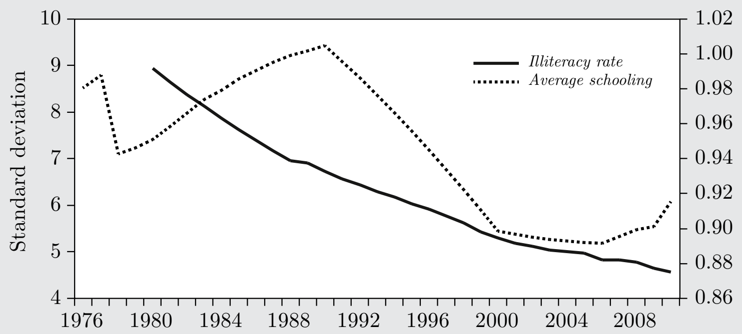
Source: Authors' calculations.
Note: Dropout rate and gross enrollment ratio are measured on the right-hand axis.

Figure 5. Annual mean of illiteracy rate and average schooling



Source: Authors' calculations.

Figure 6. Annual dispersion of illiteracy rate and average schooling



Source: Authors' calculations.

5. Empirical results

Before presenting and discussing the results, it is necessary to briefly explain how the results should be interpreted. The sign of the parameter of interest, δ, depends on the variable being analyzed. We can use equations (2) and (3) to understand this issue.

If the agreement generates the expected results,[8] the sign of δ would depend on whether the indicator measures a positive or a negative characteristic.

8. That is, if the education statistics improve in both groups after 1992 but the improvement is greater in the treatment group states.

Table 3. Expected sign of parameter

Variable	$\bar{Y}_{T1} - \bar{Y}_{T0}$	$\bar{Y}_{C1} - \bar{Y}_{C0}$	$\delta = \left[\bar{Y}_{T1} - \bar{Y}_{T0}\right] - \left[\bar{Y}_{C1} - \bar{Y}_{C0}\right]$
Measures a positive characteristic (cohort survival rate, gross enrollment ratio, completion rate, transition rate, and average schooling)	Positive	Positive	Positive
Measures a negative characteristic (dropout rate, failure rate, and illiteracy)	Negative	Negative	Negative
Source: Authors' elaboration.			

In the case that the indicator measures a positive characteristic of the education system, cohort survival rate for example, both $(\bar{Y}_{T1} - \bar{Y}_{T0})$ and $(\bar{Y}_{C1} - \bar{Y}_{C0}) > 0$ should be interpreted as an improvement after 1992. Furthermore, $(\bar{Y}_{T1} - \bar{Y}_{T0}) > (\bar{Y}_{C1} - \bar{Y}_{C0})$ implies that the improvement was greater for states in the treatment group. For those indicators that measure a negative characteristic, failure rate for example, the reasoning is analogous. Table 3 summarizes these explanations for both cases.

5.1. Results of the basic model

Table 4 shows the results of estimating model (1). The first column lists each of the dependent variables (y_{it}), while the second column shows the change in the expected value of the treatment group. The third column shows the difference from the expected value in the control group. Finally, the last column shows the estimated value of the parameter of interest.

The results of all the elementary school variables indicate that following enactment of the ANMEB there was a statistically significant decrease in dropout and failure rates and a significant increase in cohort survival rate and completion rate for states in both groups. Moreover, the last column indicates that this progress was greater in states that belong to the treatment group. This implies that after 1992, there was a reduction in disparities among states, at least as regards elementary schools.

The variables that measure the development of middle-school education and the average schooling series showed statistically significant progress in each of the groups, though my estimations do not show that the gap between the states decreased, i.e., the parameter was estimated as statistically insignificant. In contrast, the illiteracy rate declined in

Table 4. Results of DD methodology for model (1)

Variable	$\bar{Y}_{T1} - \bar{Y}_{T0}$	$\bar{Y}_{C1} - \bar{Y}_{C0}$	$\hat{\delta}$
Elementary-school dropout rate	-2.728 ***	-2.064 ***	-0.663 **
Cohort survival rate in elementary schools	16.411 ***	11.219 ***	5.192 ***
Failure rate in elementary schools	-3.156 ***	-2.481 ***	-0.675 **
Completion rate in elementary schools	6.040 **	1.652	4.388 ***
Transition rate in middle schools	6.271 **	6.479 ***	-0.208
Middle-school dropout rate	-0.912 **	-0.938 ***	-0.026
Cohort survival rate in middle schools	1.583 *	1.766 **	-0.183
Gross enrollment ratio in middle schools	3.130 **	3.258 ***	-0.128
Completion rate in middle schools	6.826 **	4.653 ***	2.173
Average schooling	1.495 **	1.512 ***	-0.017
Illiteracy rate	-2.364 **	-1.147 **	-1.216 **

Source: Authors' calculations.
Notes: The results in this table were obtained when the variable d_g was generated using a state government participation rate of 30%. Results do not change significantly if we use 40%.
The symbols *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively.

both groups and the reduction was estimated as being substantially greater in the treatment group.[9]

In general, the three variables included to control for other key features in the states that may be important to describe the level of education before and after enactment of the ANMEB were found to be statistically significant, especially so in the case of real gross state product per capita and state fiscal independence. The last variable, percentage of population living in urban areas, was occasionally found to be not significant in explaining the education indicators.

5.2. Was the ANMEB the cause of the improvement in the indicators?

The results in the previous subsection indicate that after enactment of the ANMEB there was an improvement in all of the indicators. Moreover, the computations indicate that there was a reduction in

9. As a robustness check of the results, the estimation of model (1) is extended by defining the groups differently. In these new cases, three groups are defined, and the middle one is excluded from the sample. As in the first exercise, those states in which a higher percentage of students were enrolled in the state school system in 1992 are placed in the control group and those with a lower percentage are placed in the treatment group. These new results are not significantly different from those obtained originally; they are described and explained in the appendix.

disparities between states in the control and the treatment group, though only for the elementary-school variables and illiteracy rate.

Figures 1 to 6 provide the evolution of the mean and standard deviation of the variables, showing that the indicators' progress is noticeable even before the ANMEB. Therefore, it is evident that further analysis is necessary in order to determine whether the ANMEB was responsible for the performance of these indicators.

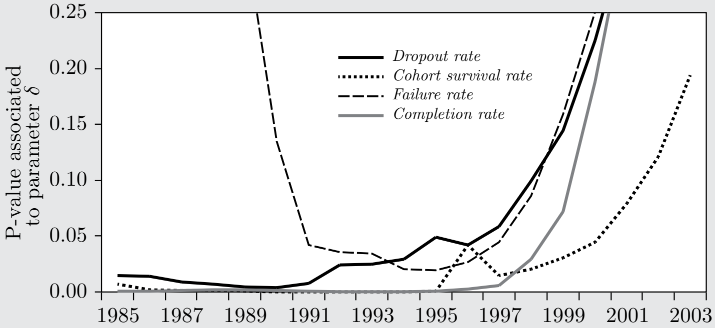
For this task, I continue estimating the same model as before (Equation (1)), although with a slight modification to the time dummy variable. I estimate this model repetitively for the period from 1985 to 2003, changing the year in which the time dummy switches on. My expectation is that if the ANMEB was responsible for improvement in the variables, the estimated value of the parameter of interest, $\hat{\delta}$, would be statistically insignificant for the years in which the time dummy switches on before 1992, and statistically significant for the years in which the time dummy switches on after 1992.

Moreover, I would not expect the results of a policy change of this nature to be immediately apparent, but rather to appear gradually over time. If this is true and the effects of the policy change took some time to materialize, the parameter would be expected to have larger absolute values—with greater statistical significance—when the years immediately following enactment of the ANMEB are excluded—i.e., when the time dummy switches on some years after 1992. Therefore, I perform the computations for all the variables, even those for which the parameter δ is not significant in Table 4.

The results of these computations are shown in Figures 7 to 9. For each variable I show the P-value—related to the null hypothesis of no significance—of the parameter δ on the Y-axis; on the X-axis I show the year in which the time dummy variable switches on for that particular estimation. If the ANMEB was the cause of greater improvement in the states in the treatment group than in the states in the control group, the P-value would be expected to decrease over time (the parameter becomes more significant over time), i.e., the P-values would be below 0.05.

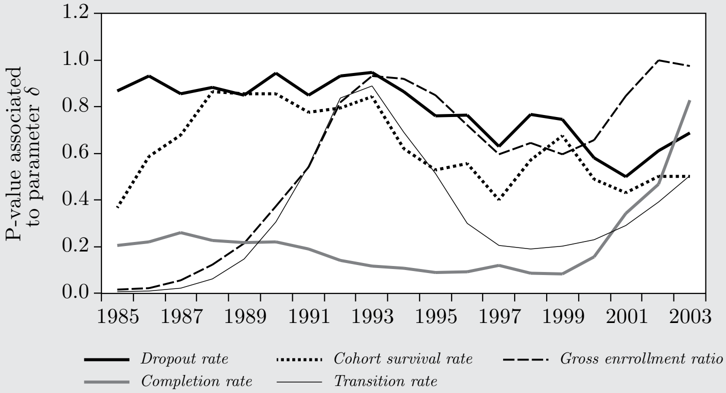
The P-values of the different elementary-school indicators in Figure 7 do not follow the expected pattern that would indicate that the ANMEB was responsible for the decrease in disparities between the states in the control and the treatment group. Parameter δ was statistically significant long before implementation of the ANMEB, except for the

Figure 7. Did the ANMEB cause improvements in elementary-school indicators?



Source: Authors' calculations.

Figure 8. Did the ANMEB cause improvements in middle-school indicators?

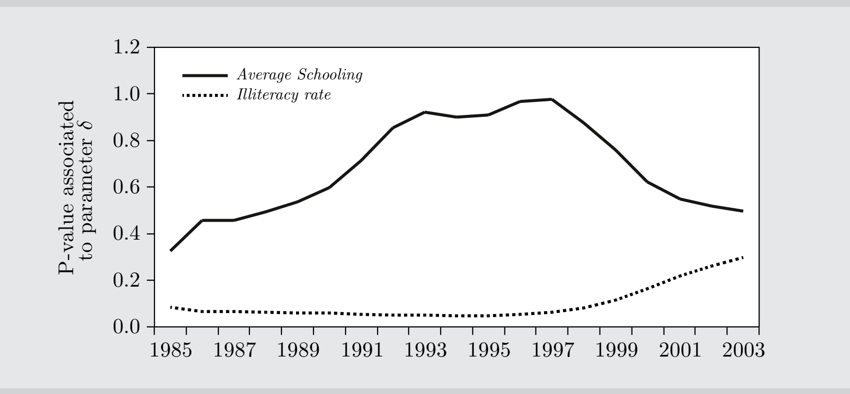


Source: Authors' calculations.

failure rate, which appears to be significant only for the 1991–1996 period. These results indicate that the reduction in disparities among states started in 1985 (or even earlier); this decline lasts for most of the 1990s and then stops, i.e., the P-values increase beyond 0.05.

The results for middle-school indicators in Figure 8 indicate that the ANMEB was not effective in reducing disparities among states.

Figure 9. Did the ANMEB cause improvements in the illiteracy rate and average schooling?



Source: Authors' calculations.

There is evidence in favor of this reduction only for the transition rate and gross enrollment ratio, and for a period prior to enactment of the ANMEB, i.e., before the 1990s. There are some indicators whose P-values do not appear in the graphs; this is because for those variables the parameter δ is not significant, i.e., it is not higher than 0.25. Figure 9 shows that the disparity in the illiteracy rate diminished only for a few years after enactment of the ANMEB.

6. Concluding remarks

I analyze various education statistics to evaluate whether there is sufficient empirical evidence to support the assertion that the ANMEB has improved education indicators and contributed to reducing disparities among the states. The results indicate that all of the variables experienced significant improvement after 1992 in both the treatment and the control group. These results are robust to changes in the specification of the treatment and control groups. Nevertheless, the question of whether this progress can be attributed to the ANMEB remains. When the model is modified to examine the period in which the disparities were decreasing, I find that for most of the elementary-school variables, the disparities began to diminish during the mid-1980s and that this trend persisted until the end of the 1990s. For the rest of the variables, evidence of a decrease in disparities is not so straightforward.

Overall, the empirical evidence shows irrefutable support for the existence of an improvement in education indicators during the period being analyzed, both in the treatment and control states. Nevertheless, this evidence is insufficient to affirm that the observed improvement was caused by the ANMEB.

REFERENCES

Arnaut, A. (1994), "*La federalización de la educación básica y normal (1978–1994)*," *Política y Gobierno* **1**(2): 237–74.

Bray, M. (1999), "Control of education: Issues and tensions in centralization and decentralization," in Arnove, R.F. and C.A Torres, eds., *Comparative education: The dialectic of the global and the local.* Oxford: Rowman & Littlefield.

Cárdenas, O.J. and F.J. Luna (2007), "*El gasto educativo: Una propuesta de financiamiento a la educación básica*," *Gestión y Política Pública* XVI(2): 261–79.

Faguet, J.P. and F. Sánchez (2008), "Decentralization's effects on educational outcomes in Bolivia and Colombia," *World Development* **36**(7): 1294–316.

García-Pérez, A.C. (2008), "*Propuesta de un mecanismo de distribución de los recursos del fondo de aportaciones para la educación básica y normal.*" Undergraduate thesis, Department of Economics and Finance, *Universidad de Guanajuato.*

Gershberg, A.I. (1999), "Education 'decentralization' process in Mexico and Nicaragua: Legislative versus ministry-led reform strategies," *Comparative Education* **35**(1): 63–80.

Hecock, D.R. (2006), "Electoral competition, globalization, and subnational education spending in Mexico, 1999–2004," *American Journal of Political Science* **50**(4): 950–61.

Inter-American Development Bank (1994), *Economic and social progress in Latin America: 1994 report*, Washington, D.C.

Lane, J.E. and M. Murray (1985), "The significance of decentralization in Swedish education," *European Journal of Education* **20**(2-3): 163–70.

Latapí, P. and M. Ulloa (2000), "*El financiamiento de la educación básica en el marco del federalism*," Mexico City: *Fondo de Cultura Económica.*

Meyer, B.D. (1995), "Natural and quasi-experiments in economics," *Journal of Business and Economic Statistics* **13**(2): 151–61.

Ontiveros, J.M. (2001), "*Gasto educativo y políticas distributivas de la educación primaria en México*," *Revista Latinoamericana de Estudios Educativos* **31**: 53–77.

Ornelas, C. (1995), *El sistema educativo mexicano: La transición de fin de siglo.* Mexico City: *Fondo de Cultura Económica.*

Ornelas, C. (2008), *Política, poder y pupitres: Crítica al nuevo federalismo educativo.* Mexico City: Siglo XXI Editores.

Sharma A. and O.J. Cárdenas (2008), "Education spending and fiscal reform in Mexico," *Journal of International and Global Economic Studies* **1**(2): 112–27.

Tatto, Maria T. (1999), "Education reform and state power in Mexico: The paradoxes of decentralization," *Comparative Education Review* **43**(3): 251–82.

APPENDIX

The table below shows the results when the control group includes those states in which more than 40% of students were enrolled in the state school system in 1992, whereas the treatment group includes those states in which fewer than 30% of students were enrolled in the state school system in 1992. In this case, the states excluded are those in which the percentage is between 30 and 40.

Table A1. Results, first alternative definition of control and treatment groups

Variable	$\bar{Y}_{T1} - \bar{Y}_{T0}$	$\bar{Y}_{C1} - \bar{Y}_{C0}$	$\hat{\delta}$
Elementary-school dropout rate	-2.797 ***	-1.464 ***	-1.333 ***
Cohort survival rate in elementary schools	16.841 ***	9.878 ***	6.963 ***
Failure rate in elementary schools	-3.238 ***	-3.304 ***	-0.066 *
Completion rate in elementary schools	6.532 **	4.715 ***	1.817 ***
Transition rate in middle schools	5.592 **	6.426 ***	-0.834
Middle-school dropout rate	-0.848 **	-1.274 ***	0.426 **
Cohort survival rate in middle schools	1.377 ***	2.742 ***	-1.365
Gross enrollment ratio in middle schools	3.008 ***	2.900 ***	0.108
Completion rate in middle schools	7.267 ***	6.818 ***	0.449
Average schooling	1.533 ***	1.636 ***	-0.103
Illiteracy rate	-2.244 **	-1.759 **	-0.485 *

Source: Authors' calculations.

The results are almost similar to those in the original model except for the middle-school dropout rate, which is now estimated as positive and significant. The estimation implies that the decrease in the dropout rate was greater in the control group than in the treatment group. One possible explanation for this is that the states that were excluded (those originally in the treatment group) performed very well in terms of the reduction in the middle-school dropout rate, and once they are excluded, the estimated improvement in the dropout rate of the group as a whole declines.

The table below shows the results when the control group includes those states in which more than 40% of students were enrolled in the state school system in 1992, and the treatment group includes those states in which fewer than 20% of students were enrolled in the state school system in 1992. In this case, the states excluded are those in which the percentage is between 20 and 40.

Table A2. Results, second alternative definition of control and treatment groups

Variable	$\bar{Y}_{T1} - \bar{Y}_{T0}$	$\bar{Y}_{C1} - \bar{Y}_{C0}$	$\hat{\delta}$
Elementary-school dropout rate	-2.751 ***	-1.107 ***	-1.644 ***
Cohort survival rate in elementary schools	17.167 ***	8.272 ***	8.895 ***
Failure rate in elementary schools	-3.774 ***	-3.128 ***	-0.064 *
Completion rate in elementary schools	6.309 ***	6.013 ***	0.296 **
Transition rate in middle schools	5.619 ***	4.874 ***	0.745
Middle-school dropout rate	-0.802 ***	-1.666 ***	0.864 **
Cohort survival rate in middle schools	-1.349 ***	3.718 ***	-2.369 **
Gross enrollment ratio in middle schools	2.977 ***	2.441 ***	0.536
Completion rate in middle schools	7.482 ***	7.255 ***	0.227
Average schooling	1.630 ***	1.633 ***	-0.003
Illiteracy rate	-2.196 **	-1.624 **	-0.572 *

Source: Authors' calculations.

The results only differ for two variables: dropout rate and cohort survival rate in middle schools. The change in the result for the first variable has already been discussed. Regarding the other variable, cohort survival rate in middle schools, this is now estimated to be negative and statistically significant. If the states in the treatment group had shown greater improvement after enactment of the agreement than those in the control group, then this parameter should be positive. The estimation reflects the fact that after 1992 the cohort survival rate increased more in control group states than in treatment group states.

Overall, the new calculations indicate that our estimations are robust and do not depend on how the treatment group is defined. It would seem that the two groups—i) the group including states in which more students were enrolled in the state school system in 1992 (control group), and ii) the group including states in which fewer students were enrolled in the state school system in 1992 (treatment group)—are homogeneous. Therefore, the results do not vary when the groups are modified to include different states.

VOL. 51 NO. 1 (MAY, 2014), 109–146

LATIN AMERICAN JOURNAL OF ECONOMICS

TEENAGE PREGNANCY IN MEXICO: EVOLUTION AND CONSEQUENCES*

EVA O. ARCEO-GOMEZ**
RAYMUNDO M. CAMPOS-VAZQUEZ***

We analyze the consequences of a teenage pregnancy event in the short and long run in Mexico. Using longitudinal and cross-section data, we match females who became pregnant and those who did not based on a propensity score. In the short run, we find that a teenage pregnancy causes a decrease of 0.6-0.8 years of schooling, lower school attendance, fewer hours of work and a higher marriage rate. In the long run, we find that a teenage pregnancy results in a 1-1.2-year loss in years of education, which implies a permanent effect on education, and lower household income per capita.

JEL classification: I00, J10, J11, O54

Keywords: Teenage pregnancy, schooling, labor outcomes, propensity score, matching

1. INTRODUCTION

According to Geronimus and Korenman (1992), "Teenage childbearing has been described as a cause of persistent poverty, and poverty that is transmitted intergenerationally" (p. 1187). As the event of teenage pregnancy may lead to an intergenerational cycle of poverty, the causes and consequences of teenage childbearing have been widely studied among social scientists (see, for example, Hoffman and Maynard, (2008), for an analysis in the United States and Stern (2012), for a sociological analysis in the Mexican case). However, most of the literature on the topic estimates associations or correlations of teenage pregnancy and socioeconomic outcomes and most of the international literature focuses on developed countries.

In this paper, we attempt to fill this void in the literature by analyzing the Mexican case. This is important because teenage mothers are far

* We gratefully acknowledge funding from the World Bank and the comments of Luis F. López-Calva. We are especially grateful to the editor and two anonymous reviewers for their constructive comments. All remaining errors are our own.

** Email: eva.arceo@cide.edu. Address: Centro de Investigación y Docencia Económicas, Carretera Mexico-Toluca 3655, Col. Lomas de Santa Fe, 01210, Mexico DF. Phone: +52-55-57279800, ext. 2759. Fax: +52-55-57279878

*** E-mail: rmcampos@colmex.mx, http://cee.colmex.mx/raymundo-campos. Address: El Colegio de México, Centro de Estudios Económicos, Camino al Ajusco 20, Pedregal de Santa Teresa, 10740, Mexico DF. Phone: +52-55-54493000, ext. 4153. Fax: +52-55-56450464.

doi 10.7764/LAJE.51.1.109

more common in Mexico than in the United States or other developed countries. According to World Bank data, in Mexico 69 of every 1,000 adolescents between 15 and 19 years old have children, whereas in the United States only 36 per 1,000 do. Compared to other countries in Latin America with similar development levels, Mexico's teenage childbearing rates are just above average: Brazil has a rate of 76 per 1,000 women, but Argentina and Chile have rates of 56 and 57 per 1,000, respectively. Pantelides (2004) reviews the evolution of the phenomenon in Latin America, pointing out that these rates have not decreased significantly in the last decades.

The problem has also been recognized by the Mexican government, which in 2007 implemented PROMAJOVEN, a program targeting teenage mothers who had not yet finished their primary education. In 2010, 41 percent of mothers between 12 and 19 years old had not completed their basic education. The problem is more serious for older cohorts than for younger cohorts. Perhaps due to this heterogeneity in school attainment of adolescent mothers, the program now also helps women to complete up to middle school (9th grade in Mexico). These figures do not take into account that teenage mothers may be systematically different from adolescents who do not have children, and hence the observed educational underachievement may not be entirely due to early motherhood. Our paper will provide additional evidence to justify these kinds of programs in Mexico.

In order to disentangle the effect of teenage childbearing on several socioeconomic outcomes, we match females who became pregnant during adolescence with those who did not, based on a propensity score. In other words, using several observable characteristics we are able to compare very similar individuals whose only difference is the pregnancy event. We find substantial evidence that there is balance and common support between the treatment and control groups after matching. Our analysis focuses on both short- and long-run outcomes. We find that the single most important effect of teenage childbearing is to lower the educational attainment of females by 0.6 to 0.8 years in the short run. Most importantly, we present evidence that this effect is permanent: Our long-run estimates suggest a loss of between 1 and 1.2 years of schooling. There does not seem to be any short-run effect on the household labor supply or household income per capita. However, and most likely due to their lower educational attainment, we find that in the long run teenage mothers live in households with lower income per capita as compared to females who did not become mothers in adolescence.

Determining the causal effects of teenage childbearing has proven to be very elusive. The main empirical challenge in the estimation of the causal effects is that teen mothers are systematically different than adolescents who do not have children. This selection bias suggests that even in the absence of a child, those females who ultimately raise a child during their teenage years would have had a lower socioeconomic status than those females who did not. The literature presents several approaches to identifying the effect of teenage childbearing in the case of the United States. For instance, Bronars and Grogger (1994) analyze the effect of out-of-wedlock motherhood by comparing twin first births to single first births using a couple of censuses. Although teenage mothers tend to be unwed, this identification strategy seems to answer a different empirical question: It estimates the effect of having an additional child in the first birth of single women rather than the effect of the first birth of single women (independently of whether it was a multiple birth or not).

Other more successful approaches have been used. Geronimus and Korenman (1992) compare teen mothers to their childless sisters using several longitudinal surveys, thus removing the unobserved heterogeneity coming from family background. Hotz, McElroy, and Sanders (2005) and Ashcraft and Lang (2006) use miscarriages as an instrumental variable of birth delays. In this way, they estimate the causal effect of age at first birth on several socioeconomic outcomes. Hotz, McElroy, and Sanders (2005) find statistically significant positive effects on the probability of earning a General Educational Development (GED) degree, on the number of hours of work per week, and on wages. In contrast, Ashcraft and Lang (2006) find adverse but modest effects. Finally, Levine and Painter (2003) implement propensity score matching within schools attended by treatment and control teenagers in the United States, finding that teenage mothers are 20 percent less likely to graduate from high school. Similarly, Chevalier and Viitanen (2003) estimate a propensity score matching model using data from Great Britain. They also find adverse effects of teenage childbearing on schooling attainment, labor market experience, and wages in adulthood.

In our view, the evidence on the consequences of teenage pregnancy is more limited for developing countries than for developed countries.[1]

1. Another strand of the literature focuses on the determinants of teenage pregnancy and other risky behaviors. For literature on developing countries, see for instance Blunch (2011) on Ghana; Cardoso and Verner (2007) on Brazil; and Marteleto, Lam, and Ranchhod (2008) on South Africa.

Ferre, Gerstenblüth, Rossi, and Triunfo (2009) estimate the impact of childbearing only on educational outcomes using matching methods in Uruguay. These authors work with a cross-section with no retrospective data. As a consequence, they are only able to match females on a very limited number of observable characteristics. Kruger, Berthelon, and Navia (2009) study the effect of teenage pregnancies on high-school completion in Chile using an instrumental variable strategy. The instruments they use reflect the society's and household tolerance for teenage births. In order to measure social acceptance, they estimate the proportion of teenagers in the county who gave birth and the average county rate of unwed births. To measure household tolerance, they use a dummy of whether the mother also had a teenage pregnancy. As for the first set of instruments, we doubt that they meet the exclusion restriction because social acceptance of teenage births may reflect preferences for gender roles, which in turn affect educational attainment. The same is also true for the measure of household tolerance: if having a teen birth reduces the probability of high school completion, the same is true for the teen mother's mother; hence, high school completion of the teen today is affected through the intergenerational transmission of educational attainment.[2] A paper that is more similar to ours is Ranchhod, Lam, Leibbrandt, and Marteleto (2011) who use the Cape Area Panel Study to estimate propensity-score weighted regression in South Africa. They find a negative effect of a teenage birth on educational attainment, but the effect tends to diminish over time, suggesting that teenage moms catch up with childless teenagers. Unlike our study, Ranchhod, Lam, Leibbrandt and Marteleto (2011) do not exploit the longitudinal nature of their data by estimating a difference-in-difference estimator. None of the studies cited above contrast the short- and long-run effects of teenage births as we do in this paper.

In the case of Mexico, most of the studies analyze the association of pregnancy with outcomes, but lack a clear control group to measure the impact of teenage pregnancy in later outcomes. For example, Stern (2012) conducts an excellent sociological review of the evolution of teenage pregnancy in Mexico. Using qualitative work, Stern (2007) finds that teenage pregnancy occurs in stable couples, and is not due to random encounters. Echarri Cánovas and Pérez Amador (2007)

2. For instance, Navarro Paniagua and Walker (2010) find that children of teenage mothers in Europe have lower educational attainment and are more likely to be teenage mothers themselves.

construct event histories of teenagers, finding that events such as dropping out of school, first consensual union, and leaving the parental home occur before the childbearing event. Menkes and Suárez (2003) find that a low schooling level is associated with lower contraceptive knowledge and a lower age at the first sexual encounter. These two factors, in turn, lead to a higher propensity of less educated women to become pregnant during adolescence. Furthermore, Menkes and Serrano (2010) find that women in poor families have higher rates of teenage pregnancy. Although these studies are relevant and important to increasing our understanding of the teenage pregnancy phenomenon, they only estimate associations of the pregnancy event with different outcomes. These studies also indicate that female teenagers with a pregnancy event are very different from females without the event. Hence, in order to estimate the impact of teenage pregnancy on outcomes like education, income, and work, we apply a novel strategy to the Mexican case in order to compare similar women in terms of observable characteristics.

Our identification strategy follows Levine and Painter (2003) and Chevalier and Viitanen (2003) in the sense that we match females who became mothers during adolescence to females who did not based on a propensity score. Due to data limitations, we are not able to match females within schools or families. However, we exploit two different databases to estimate short- and long-run effects. For the short-run effects we use the Mexican Family Life Survey (MxFLS), which is a longitudinal survey for which there are currently two waves publicly available (2002 and 2005). For the long-run effects, we use the 2011 Social Mobility Survey (EMOVI for its acronym in Spanish), which is a cross-section with socioeconomic information for the individuals when they were 14 years old.

Our results show that the most important effect of teenage childbearing is the permanent, lower educational attainment of the teenage mother. As a result, we find that in the long run, the households of those females who had their first child as teenagers tend to have lower income per capita. We also find that in the short run, teenage mothers reduce their school attendance (hence the lower educational attainment), and their labor supply. Finally, and in contrast with the literature in the United States, we find that having a child during adolescence has a positive effect on the probability of being married. This is most likely a result of cultural differences between Mexico and the United States.

The remainder of the paper is organized as follows. Section 2 shows the aggregate trends in teenage childbearing in Mexico. Section 3 describes the sources of data used in this paper and presents some descriptive statistics. Section 4 explains the empirical strategy that we implement. Section 5 presents the estimations of short- and long-run effects, and finally Section 6 provides concluding remarks and discussion of some policy implications.

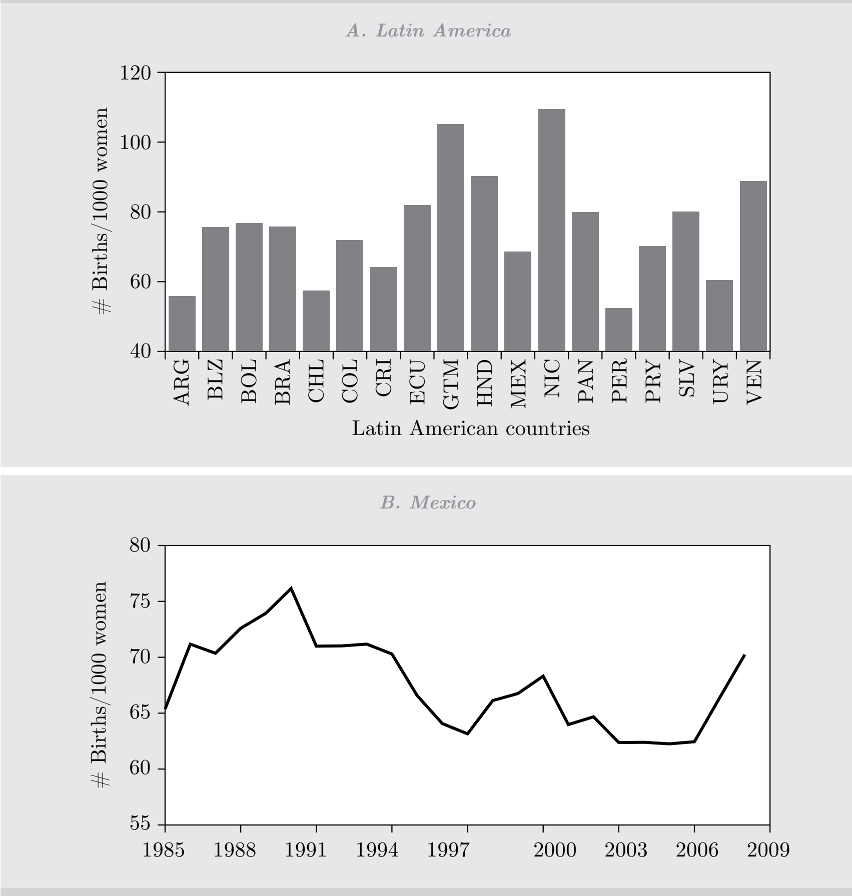
2. Aggregate Trends

In this section we discuss the aggregate trends for teenage births. The data of this section comes from the World Bank, the Mexican Population Census (1990, 2000, and 2010), and administrative birth records.[3] Figure 1, Panel A shows the number of births per 1,000 women among teenagers aged 15-19 in 2009 for a sample of Latin American countries. The unweighted average number of births per 1,000 women for this sample of countries is 75.8, whereas Mexico has a rate equal to 68.6. Among those 18 countries, Mexico has the 6th-lowest rate in the number of births per 1,000 women after Argentina, Chile, Costa Rica, Peru and Uruguay. However, using the same data source for all available countries results in an unweighted world average of 50 births per 1,000 women. Hence, although Mexico shows a slightly lower teenage pregnancy rate as compared to other Latin American countries, its rate is still higher than that of the rest of the world. Panel B shows the evolution of the number of births per 1,000 women among teenagers, based on administrative records.[4] The number of births per 1,000 women shows a decline from 1990 to 1997, then a relatively stable path from 1998-2006 at around 65 births per 1,000 women, and finally an increase in the 2007-2008 period to almost 70 births per 1,000 women.

3. Census data provides information on the numbers of childbearing women. Our results are very similar to those presented in Menkes and Serrano (2010), even though they use a different survey.

4. Administrative birth records are published by the National Statistical Institute (INEGI) in Mexico and the Ministry of Health. The data include all births registered in order to obtain a birth certificate. These administrative records include age of mother at birth, education, marital status and location of birth (county and state). We use these records in order to provide a broad picture of the evolution of teenage pregnancies. Data can be downloaded from the web sites of INEGI, http://www.inegi.org.mx/ and the Ministry of Health, http://www.sinais.salud.gob.mx/basesdedatos/index.html. We use information from the year of birth rather than year of birth registry. To calculate a series without the problem of right-censoring (births that occurred in the past may be registered at any time in the future), we restrict the data to births registered only in the same year and the year following the birth year, which represents approximately 93% of births.

Figure 1. Number of births per 1,000 women aged 15-19, Latin America and Mexico



Source: Authors' calculations.
Notes: Panel A uses World Bank data for 2009; data available at http://data.worldbank.org. ARG=Argentina, BLZ=Belize, BOL=Bolivia, BRA=Brazil, CHL=Chile, COL=Colombia, CRI=Costa Rica, ECU=Ecuador, GTM=Guatemala, HND=Honduras, MEX=Mexico, NIC=Nicaragua, PAN=Panama, PER=Peru, SLV= El Salvador, URY=Uruguay, VEN=Venezuela. Panel B uses information from the Statistical Institute (INEGI). To construct teenage births per 1,000 people, we interpolate population rates using Census data from 1990, 2000, and 2010. We use year of pregnancy rather than year of registry of birth. Due to right-censoring of the data, we limit the calculation to births registered in the same year or year following occurrence (93% of the cases on average).

Panel A in Figure 2 exhibits the fraction of births to teenage mothers, of total births. The percentage of births among teenage mothers is stable at around 16%. In contrast, the percentage of births to single mothers among all births to teenage mothers has increased in the period. As a result, the proportion of births to married women or

Table 1. Aggregate statistics, females aged 15-19, 1990-2010

	Proportions			% Childbearing		
	1990	2000	2010	1990	2000	2010
National	100.0	100.0	100.0	12.3	12.5	13.0
Rural	25.8	25.7	26.0	17.4	16.0	14.9
Urban	74.2	74.3	74.0	10.5	11.3	12.3
Education						
Primary or less	50.1	38.9	28.7	18.3	19.5	17.7
Secondary	45.0	49.1	55.4	6.4	8.3	12.0
More than secondary	5.0	12.0	15.9	4.1	4.9	7.4
Civil Status						
Single	82.5	82.3	82.1	1.3	1.7	2.5
Married	10.8	8.5	4.7	65.3	64.6	63.2
Cohabitating	5.8	8.2	11.7	60.4	60.1	60.0
Other	0.9	1.1	1.5	70.2	71.5	65.7
School Attendance						
Not attending	59.4	54.6	42.9	19.9	22.1	28.0
Attending	40.6	45.4	57.1	1.1	1.1	1.8

Source: Authors' calculations using census data.
Notes: Sample is restricted to females aged 15-19 years old with a valid answer for the number of own children. The last three columns indicate the percentage of women with at least one child born alive given the condition in the first column.

women cohabitating has decreased. These findings could be a result of a lower marriage rate triggered by teen pregnancies or a higher age at first marriage that results in fewer married teen mothers. Also, Panel B shows that while in 1985 a teenage mother was more likely to have a primary degree or less (equal to or less than 6 years of schooling), by 2002 that had changed, and a teenage mother was more likely to have a secondary degree (9 to 11 years of schooling). This last finding could be a result of higher educational achievement, and not necessarily due to a decrease in the teen childbearing rate for those with primary schooling or less.

Table 1 provides statistics for females aged 15-19 years old in Mexico for the period 1990-2010 using Census data.[5] The first three columns show the proportion of each group in the population and the last three

5. Census data are available at the web site of the National Statistical Institute (INEGI) of Mexico, http://www.inegi.org.mx.

columns show the percentage of women in that age group with at least one child born alive. The table shows that the percentage living in rural areas (less than 2,500 inhabitants) has remained relatively constant at 25%. On the other hand, education and school attendance has improved during the period of study. An interesting fact is that the proportion of single females is stable at 82% and the proportion of either married or cohabitating is stable at 16-17%. However, the percent of females who are married has decreased substantially over time, from 10.8% in 1990 to 4.7% in 2010. At the same time, the percentage of females who are cohabitating has increased from 5.8% in 1990 to 11.7% in 2010.

When examining data on childbearing teenagers only (columns 4 to 6 in the table), we find that the percentage of females with at least one child born alive has increased from 12.3% in 1990 to 13% in 2010. The increase in childbearing rates is mostly within the urban sector, as females in the rural sector have become less likely to be teenage mothers. Within education groups, the highest childbearing rate is among women with primary schooling or less (less than 8 years). Hence, the trends shown in Panel B of Figure 2 are a result of higher school attainment over time. However, the rate is decreasing slightly for the group of women with primary education and increasing for women with more education such as secondary (9-11 years of schooling) or more than secondary (more than 12 years of schooling). In terms of school attendance, if a woman is attending school the probability that the woman has children is small. When we disaggregate by marital status we find that the childbearing rate is very small (1.3-2.5%) among single women, although this rate doubled in the 1990-2010 period. In Mexico, childbearing is associated with marriage or cohabitation.[6] Moreover, the childbearing rate among married women has remained stable over time, which indicates that the increase in childbearing has been borne by single women.

3. Data and Descriptive Statistics

We are interested in the effects of teenage pregnancy on individual outcomes of the teenage mother and also on family outcomes. Most of the previous literature has focused on short-run outcomes, given

6. Census data cannot provide information about the timing of the events, so we cannot know whether pregnancy occurred before marriage or cohabitation took place.

the data availability. In this paper, we attempt to measure the consequences of teenage pregnancy both in the short and long run. For the short-run analysis, we use the Mexican Family Life Survey (MxFLS), a nationally representative longitudinal study, for the period 2002-2005.[7] In the baseline year, the MxFLS was applied to 8,440 households and approximately 35,000 individuals. The follow-up survey was applied in several months between 2005 and 2006 with an attrition rate of approximately 10% at the household level. The survey includes information on demographics, work, and health.

In the short-run analysis, we restrict the MxFLS data to females aged 14 to 18 in 2002 who are childless and not pregnant. Moreover, we further restrict the sample to females who are not married or cohabitating in 2002. Then, we follow those females into the 2005 survey. Hence, we are interested in females who became pregnant between 2002 and 2005 while still a teenager, which represents the treatment variable. Under these restrictions, the final dataset includes 1,003 females with 131 observations in the treatment group.[8, 9] The teenage pregnancy rate is around 13% in our sample, which is similar to our findings in the previous section. Due to the small sample size, we do not focus on teenage out-of-wedlock childbearing specifically, but we do present some results in the extensions section.[10] The variables in the analysis include age, years of schooling, school attendance indicators, work status, indigenous language, dropout age,[11] knowledge of contraceptives, previous sexual activity, Raven test score (percent of correct answers in the test), having been born in rural areas (i.e., localities with less than 2,500 inhabitants), and father absent from the household. We also use information about the head of household: age, years of schooling, and dummies for gender and work status. Finally, we use variables at the household level: household size, number of members ages 0 to 5, 6 to 18, and older than 65, average hours of work for members older

7. Data available at http://www.ennvih-mxfls.org.

8. This definition has an important caveat: In the post-treatment year (2005) some of the women in the control group are still teens and could potentially become adolescent mothers. If we had all the completed histories of pregnancies in our sample, then the estimates would likely be higher using the right controls.

9. The attrition rate in this sample is 9.7%. We run t-tests on those women with missing information in 2005 (either missing or lost due to attrition) and women with complete information and there are no significant differences in age, working status, schooling, Raven's test score, previous sexual experience, and knowledge of contraceptives. Hence, there is no evidence that there is sample selection bias.

10. We include cohabitation in the definition of marriage.

11. Some women have not dropped out of school, and in those cases we replace the missing value for dropout age with the age of the individual. Since we are also controlling for age, this has no effect on the estimates.

than 18, average age, income per capita, number of rooms in dwelling, and dwelling characteristics (asset ownership).

In order to measure long-run impacts, we use data from the 2011 Social Mobility Survey (EMOVI).[12] This survey is representative at the national level for both males and females between 25 and 64 years old. The main goal of the survey is to estimate intergenerational mobility. The survey not only records current characteristics, but also collects information about characteristics of the household of origin when the individual was 14 years old. For example, the survey asks about the educational level of both parents and characteristics of the dwelling. The survey includes a question on the age of the individual when he or she had his or her first child. Hence, we define the treatment variable as females who had their first child when they were 15-19 years old. We do not include teenagers who became pregnant when they were 14 years old, in order to include pre-treatment characteristics of the household of origin. This allows us to capture long-run effects because, for example, we can analyze outcomes of females from 6 to 45 years after the teenage pregnancy. However, an important drawback of this survey is that it does not include extensive information about the women when they were teenagers as the MxFLS does, which precludes us from estimating difference-in-differences effects.[13]

Table 2 presents some descriptive statistics for both samples. The MxFLS sample is restricted to the baseline year. Age is relatively similar across samples. In the MxFLS, females who became pregnant between 2002 and 2005 had less education than other females, but the difference is not statistically significant at the 5% level. On the other hand, women in the treatment group had lower school attendance levels and were more likely to work before the pregnancy event. In the case of EMOVI, schooling and proportion working refer to current outcomes. They show that after a teenage pregnancy, women have lower schooling levels and a lower probability of being employed than women without a teenage pregnancy. The subsequent rows show that women who became pregnant come from more disadvantaged backgrounds, as measured by years of schooling of the head of the household (MxFLS) or parents (EMOVI). Also, in the case of the

12. For more information, visit http://www.ceey.org.mx.

13. As we discuss below, difference-in-differences estimates are more reliable than a simple difference between treatment and control because they control for individual fixed effects and common trends between treatment and control groups.

Table 2. Descriptive statistics, MxFLS and EMOVI

	MxFLS in baseline (2002)			EMOVI		
	Control	**Treatment**	**Diff**	**Control**	**Treatment**	**Diff**
N	872	131		3378	1030	
Age	15.69	15.92	0.23	39.11	39.61	0.49
	[0.047]	[0.112]	[0.122]	[0.210]	[0.359]	[0.416]
Yrs school	8.29	8.00	-0.29	8.63	6.83	-1.81
	[0.074]	[0.210]	[0.223]	[0.072]	[0.113]	[0.134]*
Working	0.12	0.20	0.08	0.48	0.44	-0.04
	[0.011]	[0.035]	[0.037]*	[0.009]	[0.015]	[0.018]*
Attendance	0.72	0.49	-0.23			
	[0.015]	[0.044]	[0.046]*			
HH size	5.79	5.42	-0.37	5.81	6.29	0.49
	[0.067]	[0.179]	[0.192]	[0.042]	[0.079]	[0.089]*
Yrs school (head household)	5.85	5.12	-0.73			
	[0.146]	[0.359]	[0.388]			
Father: Yrs school				3.99	2.83	-1.15
				[0.075]	[0.106]	[0.130]*
Mother: Yrs school				3.70	2.65	-1.05
				[0.070]	[0.103]	[0.125]*
Knowledge contraceptives	0.90	0.91	0.00			
	[0.009]	[0.025]	[0.027]			
Previous sexual exp.	0.02	0.08	0.06			
	[0.005]	[0.023]	[0.024]*			

Source: Authors' calculations using MxFLS and EMOVI data.
Notes: . Sample is restricted to females aged 15-19 years old with a valid answer for the number of own children in the case of MxFLS. In MxFLS: Treatment is defined as women with a pregnancy event (only 3 women report a pregnancy but no child alive). In EMOVI, treatment is defined as child was born when woman was a teenager. HH Size in EMOVI refers to household size when female was 14 years old. Standard errors in brackets. * denotes significance at 5%.

MxFLS, women who became pregnant were already more sexually active than women in the control group. On the other hand, the Raven test score does not show significant differences between the treatment and control groups. In general, these results show the importance of controlling for selection bias.

4. Empirical Strategy

Our goal in this paper is to estimate the effect of teenage pregnancy on outcome variables such as years of schooling, school attendance, working status, and marriage status. The ideal experiment would be to randomly assign pregnancies to teenagers (treatment) and then compare the outcomes. Obviously, such an experiment would be unethical and unfeasible. We define Y_{1i} as the potential outcome in the treatment state and Y_{0i} as the potential outcome in the control state for individual i and define treatment as $D_i = 1$. The parameter of interest is the average treatment on the treated (ATT) defined as the mean difference in outcome variables given treatment, $ATT = E[Y_{1i} - Y_{0i} \mid D_i = 1]$. However, the term cannot be estimated given that it is not possible to observe the same individual in the treatment and the control group at the same time. This is the "fundamental problem of causal inference" (Holland, 1986). The problem is that the term $E[Y_{0i} \mid D_i = 1]$ is not observed and has to be estimated (from this point forward we will omit the subscript i for notational simplicity).

We rely on the assumption of selection on observables in order to construct a valid counterfactual. In particular, we assume that conditioning on observable characteristics before the treatment occurs removes differences in the untreated state between teenagers who became pregnant and those who did not. In other words, we assume that $(Y_0 \perp D)|X$, which is commonly referred to in the literature as the conditional independence assumption (CIA) or the unconfoundedness assumption. This assumption means that the outcome for teenagers who did not become pregnant (untreated state), for example years of schooling, is independent of treatment conditional on observable characteristics.

In order to identify the ATT, the common support also needs to hold, $\Pr(D = 1 \mid X) < 1$. This assumption means that for every X there are individuals who do not get the treatment. Ideally, we would like to match individuals in the treatment and control groups within cells of observable characteristics. However, this is not possible due to the multidimensionality problem. In order to overcome this issue, Rosenbaum and Rubin (1983) propose to estimate propensity scores. These can be easily estimated using a logit or probit of the probability of treatment on observable characteristics, $\Pr(D=1 \mid X) = P(X)$. Rosenbaum and Rubin (1983) show that under the CIA:

$$(Y_0 \perp D) \mid X \Rightarrow (Y_0 \perp D) \mid P(X) \tag{1}$$

Instead of comparing treatment and control groups within the same set of X, we compare individuals based on an index that summarizes the observable characteristics information. If the assumptions of the model are satisfied, the ATT using a propensity score is estimated as:

$$\theta_{ATT}^{PSM} = E_{P(X)|D=1}\left\{E\left[Y_1 \mid D=1, P(X)\right] - E\left[Y_0 \mid D=0, P(X)\right]\right\} \tag{2}$$

The ATT is merely the difference in mean outcomes for treated individuals and mean outcomes of individuals in the control group but reweighted or readjusted by the propensity score, $P(X)$, such that they are as similar as possible to the treatment group in the common support region.[14, 15]

We estimate the impact of teenage pregnancy in the short and long run. For the long-run estimates, we apply Equation (2). For the short-run estimates, we can improve our estimates by taking advantage of the panel structure of the data. If there is unobserved heterogeneity

14. This is the procedure we follow to estimate the effects: 1. Estimate the propensity score, 2. Match individuals based on the propensity score. In other words, we compare individuals with similar propensity scores and take the difference in outcomes for those individuals.

15. For comparison purposes we also include the results of a simple ordinary least squares (OLS) regression. However, we emphasize that PSM is preferred over regression for several reasons. First, PSM only takes into account observations with very similar values in the propensity score. Observations with dissimilar values are not taken into account to calculate the ATT. Second, we show balance in the covariates and common support tests in order to be transparent about the estimation. Third, OLS estimators also suffer from the curse of multidimensionality. And finally, the PSM estimator does not impose as many restrictions on the functional form as OLS.

that is fixed over time for individuals in the sample or common trends between the treatment and control groups, then we can eliminate this bias by estimating difference-in-differences effects:[16]

$$\theta_{ATT}^{PSM} = E_{P(X)|D=1}\begin{pmatrix} \left\{ E\left[Y_{t1} \mid D=1\right] - E\left[Y_{t0} \mid D=1\right]\right\} - \\ \left\{ E\left[Y_{t1} \mid D=0, P(X)\right] - E\left[Y_{t0} \mid D=0, P(X)\right]\right\} \end{pmatrix} \quad (3)$$

Hence, the matching difference-in-differences estimator also relies on the assumption of parallel counterfactual trends of the outcomes between the treatment and control groups.

Before estimating the ATT, three key aspects need to be considered. First, it is important to question the conditional independence assumption. Of course, the assumption is untestable, but we do have possible checks to investigate whether the assumption is likely to hold. Second, there are no strict rules as to what variables should be included in the propensity-score estimation (Caliendo and Kopeinig, 2008). Third, it is possible that the ATT is sensitive to the matching method (Smith and Todd, 2005).

The main assumption of matching on the propensity score is that observable characteristics are balanced between the treatment and control groups. In other words, within some specified values of the propensity score there should be no differences in observable characteristics between the treatment and control group. If there are differences in observable characteristics, then it is likely that there are differences in unobservable characteristics, making the estimation of the ATT unfeasible. Below, we present different tests in order to provide evidence of balance in the propensity score. However, it is important to point out that the estimate is correct only if there is no selection on unobservables bias present. If there are unobservable characteristics that differ between treatment and control, then the ATT estimate will be biased. We attempt to control for this bias by including a rich set of control variables as well as non-linear effects.

16. The procedure for estimating this equation is similar to Equation (2), however instead of taking differences in the post-treatment period, we also take differences in the pre-treatment period. For example, for the effect on schooling we take the difference after the teenage pregnancy event between treatment and control minus the difference before the teenage pregnancy event.

One of the main advantages of the propensity score is that the information on all observable characteristics is summarized in a single index. There is a trade-off of bias versus efficiency in the number of explanatory variables. On the one hand, Caliendo and Kopeinig (2008), Dehejia and Wahba (1999, 2002), and Heckman, Ichimura, and Todd (1997) mention that omitting important variables that determine treatment could bias the ATT estimate. On the other hand, Bryson, Dorsett, and Purdon (2002) point out that including irrelevant variables increases the variance of the ATT estimate. Moreover, the assumption of balance needs to hold not only for linear terms but also for non-linear terms. This implies that the propensity score may include interactions and higher-order terms (Dehejia and Wahba, 1999, 2002). This could potentially increase the variance in the ATT estimate. Instead of relying on the statistical significance of observable characteristics on the propensity score, we include variables in order to achieve balance. Nonetheless, in the robustness checks section we compare models with variations in the set of observable characteristics included in the propensity score estimation in order to compare the ATT and its standard errors.

Smith and Todd (2005) show that the ATT estimate may be sensitive to the matching method. Also, Heckman, Ichimura, and Todd (1997) suggest that the matching may be done on the log odds ratio $\left(\log\left(P(X)/(1-P(X))\right)\right)$ instead of on the propensity score $P(X)$. This is especially recommended when there is choice-based sampling in the survey. We include both recommendations in our analysis.

5. RESULTS[17]

As previous literature has pointed out that the ATT may vary according to the matching method, we present our results for three different matching methods: (1) matching with a kernel Epanechnikov and a bandwidth of 0.1; (2) matching to the three nearest neighbors within a radius of 0.01; and (3) in order to restrict even further the comparison group, we match treatment and control individuals within urban/rural, age, and school attendance status (for the long-run estimates, we only restrict to urban/rural and age). We also present the results using other matching methods as a robustness check.

17. All our matching results use the ado-file psmatch2 in Stata provided by Leuven and Sianesi (2003). We employ a logistic regression to estimate the propensity score.

The main results are presented using a propensity score that includes linear, squares and interaction terms. The model using the MxFLS data uses 108 variables and the model using the EMOVI data employs 57 variables.[18] The robustness section includes results for different specifications of the propensity score. Also, we present robustness checks with the log odds ratio as the matching score instead of the propensity score. In general, our results are stable across specifications and matching methods.

5.1. Balance of the propensity score

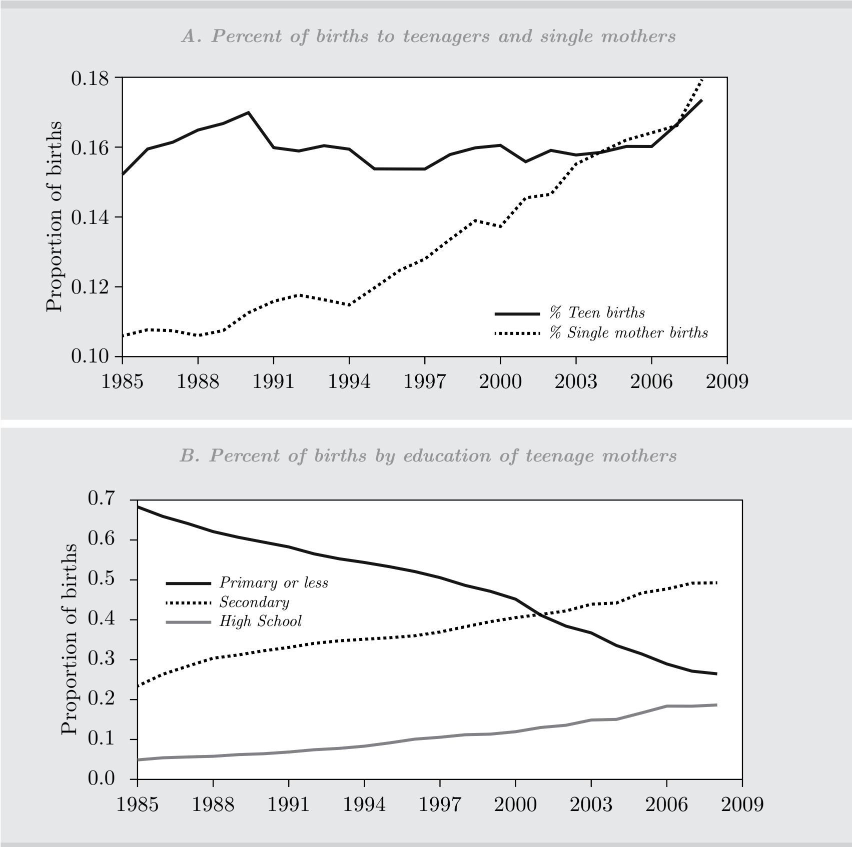
We estimate different tests to corroborate balance in the propensity score. First, we provide graphical evidence based on results by Dehejia and Wahba (1999, 2002) before and after matching to corroborate the balancing and the commom support assumptions. We also include the stratification test before and after matching proposed by Dehejia and Wahba (1999, 2002).[19] Second, we include the standardized bias measure proposed by Rosenbaum and Rubin (1985) before and after matching.[20] We report only the median standardized bias. According to Caliendo and Kopeinig (2008), a median standardized bias less than 5% is "sufficient." Third, as proposed by Sianesi (2004), we report the p-value of the joint significance test of the propensity score model

18. MxFLS 2002: age, years of schooling, school attendance, work status, indigenous language, dropout age, Raven's test score, knowledge of contraceptives, previous sexual activity, rural status, and father absent from the household. The variables included related to the head of the household are: years of education, age, female, and work status. We also include household size; number of members 0-5, 6-18, and older than 65; average hours worked in the household; mean age and income per capita of the household; number of rooms in the household; and several dummies for household assets such as indicator variables for no vehicle, no stove, no public water service and no sewage service. We also include 72 interaction terms between individual variables (age, schooling, work, indigenous, dropout age, Raven's test score, knowledge of contraceptives and previous sexual activity) and household variables and squares of age and years of schooling. We include 57 variables in the estimation of the propensity score for EMOVI: age and age squared, born in rural area, and information about both parents when individual was 14 years old, namely: education, work status, formal sector job, indigenous language, and what parent the individual was living with. The variables included about the household are: number of siblings, household size, number of rooms and cars, household assets such as no stove, no washing machine, no refrigerator, no television, no public water service, no sewage service, and no electricity. Finally, we include interactions of individual variables with household characteristics as well as squares and interactions of years of education of both parents, and work status of both parents.

19. However, they only present the stratification test before matching, while we believe the result of the test after matching is also informative. The stratification test relies on dividing observations in the treatment and control groups into quintiles or deciles. Then, within each quintile or decile, t-tests are employed for difference in means between treatment and control groups. If we have 10 variables and 5 quintiles, we have 50 tests. We report the percentage of the total tests that fail to reject the null of equal means. Dehejia and Wahba (1999, 2002) point out that this test can be used to select the variables included in the propensity score.

20. The Standardized Bias (SB) is defined as $100 \times \left(\bar{X}_1 - \bar{X}_0\right) \Big/ \sqrt{0.5\left(V(X_1) + V(X_0)\right)}$, where the subscript refers to treatment (1) and control (0).

Figure 2. Teenage pregnancy (ages 15-19) in Mexico, 1985-2008



Source: Authors' calculations.
Notes: Panels use information from the Statistical Institute (INEGI). To construct teenage births per 1,000 people, we interpolate population rates using Census data from 1990, 2000, and 2010. We use year of birth rather than year of registry of birth. Due to right-censoring of the data, we limit the calculation to births registered in the same year or year following occurrence (93% of the cases on average). In panel A, the percentage of births reported by single women excludes the percentage of women with invalid information on civil status. % Teen births refers to the percentage of teen births of total births. % Single mother births refers to the percentage of teen births with a single mother (excludes cohabitation). In panel B, around 3-5% of females have invalid education information. Primary or less refers to 8 years of schooling or less, Secondary refers to 9-11 years of schooling, High School refers to 12-15 years of schooling.

before and after matching.[21] Fourth, we report the percentage of variables that fail to reject the null hypothesis of equal means before and after matching. Finally, we report the number of observations in

21. In other words, we estimate $P(X) = \beta X$ and test the joint hypothesis that $H_0 : \beta = 0$ before and after matching. The procedure after matching includes the weights for each control.

the treatment and control for each matching method. With all these tests, our aim is to provide evidence in favor of the balancing and common support assumptions.

Figure 3 shows box plots and histograms before and after matching. To present the results, we use 3 nearest neighbors within a radius of 0.01. The figure, which includes the results for both the MxFLS and EMOVI, shows that even before matching, the treatment and control groups are not substantially different. Before matching in the MxFLS (Panel A), the mean value of the propensity score for the control group is approximately 0.10 and for the treated group it is approximately 0.25. For the EMOVI (Panel B), the mean values are even closer. Panels E and F show the box plots after matching. The box plots do not show differences in the range of the propensity score between treatment and control. Panels C and D show the number of observations in the treatment and control by deciles of the propensity score. The histograms illustrate that there is a sufficient number of observations

Figure 3. Balance in the propensity score, MxFLS and EMOVI

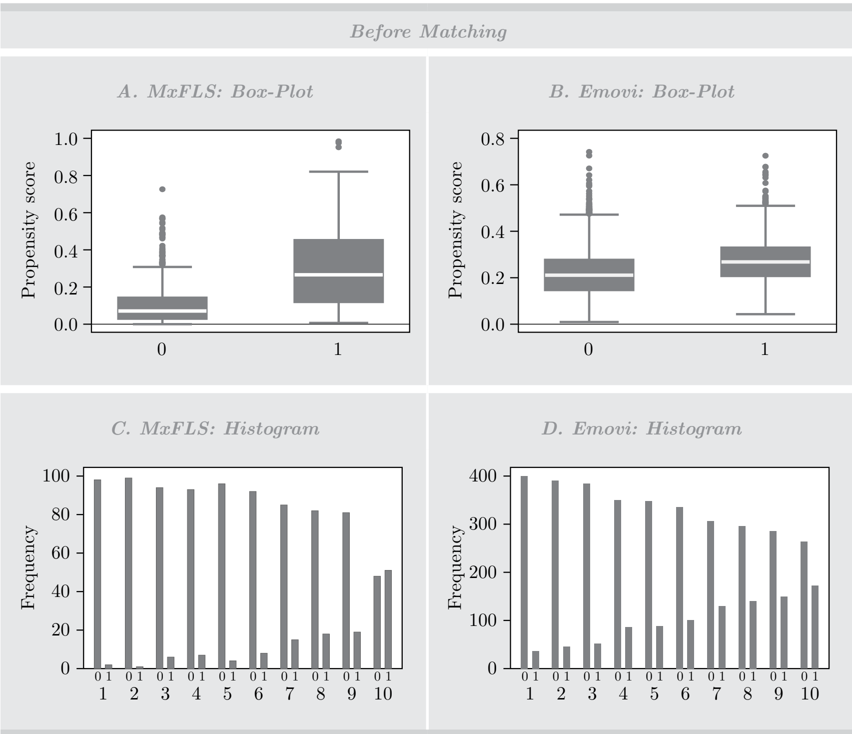
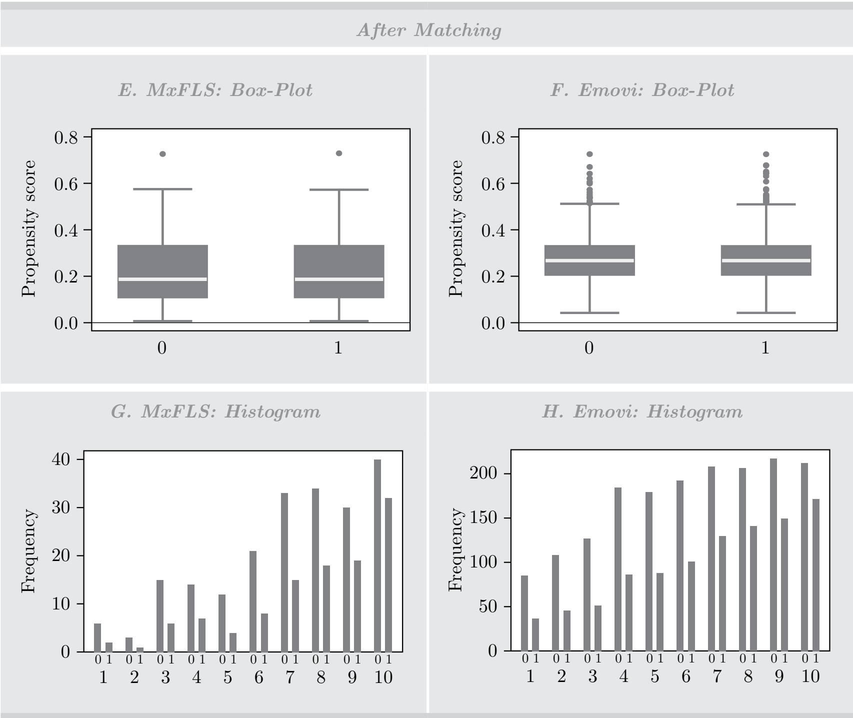


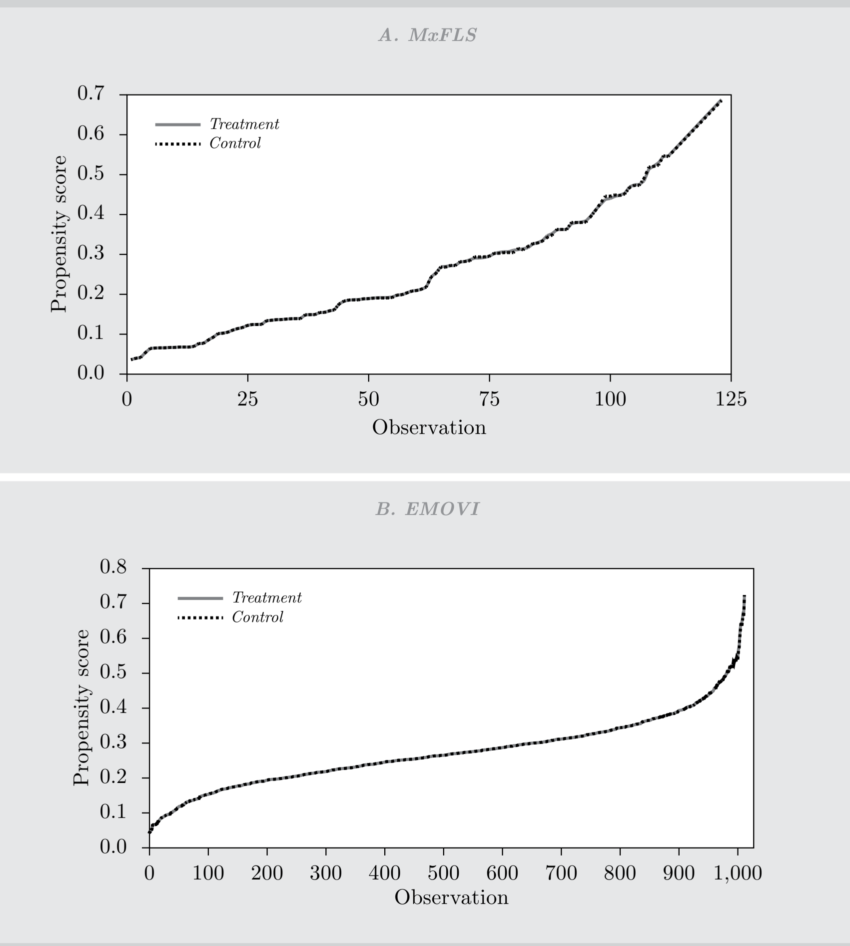
Figure 3. (continued)



Source: Authors' calculations.
Notes: After matching, figures use the method of 3 nearest neighbors within a radius of 0.01. In the box plots, 0 refers to control observations and 1 to treatment observations. In the histograms, the x-axis has two rows: the first row refers to control and treatment, and the second row to deciles of the estimated propensity score. We include 86 variables in 2002 for the estimation of the propensity score for MxFLS: age, years of schooling, school attendance, work status, indigenous language, knowledge of contraceptives, previous sexual activity, rural status, and father absent from the household. The variables included about the head of the household are: years of education, age, female, and work status. We also include household size, and household members 0-5, 6-18, older than 65, average hours worked in the household, mean age, and income per capita of the household, number of rooms in the household, and dummy variables for household assets, such as: no vehicle, no stove, no public water service and no sewage service. We also include 72 interaction terms between individual variables (age, schooling, work, indigenous, knowledge of contraceptives, and previous sexual activity) and household variables and squares of age and years of schooling. We include 57 variables in the estimation of the propensity score for EMOVI: age and age squared, born in rural areas, and information about both parents when individual was 14 years old, such as: education, work status, formal sector job, indigenous language, what parent the individual was living with. We also include information about the household: number of siblings, household size, number of rooms and cars, and dummies of household assets such as: no stove, no washing machine, no refrigerator, no television, no public water service, no sewage service, and no electricity. Finally, we include the interactions of individual variables with household characteristics as well as squares and interactions of years of education of both parents, and work status of both parents.

in the control group to match the treated group. The after matching histograms show that for each decile we have more observations in the control than in the treated group, with the exception of the top decile in the MxFLS.

Figure 4. Average propensity score in treatment and control, MxFLS and EMOVI



Source: Authors' calculations.
Notes: Matching uses the 3-nearest neighbors method within a radius of 0.01. We sort the treated observations with the propensity score (solid line) and then take the average of the propensity score for the matched controls of each treated observation (dotted line).

Figure 4 shows the estimated propensity score for each treated observation and the average propensity score for the matched controls. What it indicates is that the matching method succeeds in finding very similar observations between the treatment and control groups. In general, Figures 3 and 4 show that the common support condition for the estimation of ATT holds.

Table 3. Balance in the propensity score

	DW test		Median bias		LR test		Diff means		# Treat	# Control
	Before	After	Before	After	Before	After	Before	After		
A. MxFLS										
Epanechnikov kernel, bw=0.01	0.03	0.00	13.97	4.69	0.00	0.99	0.32	0.01	118	865
NN-3, radius 0.01	0.03	0.00	13.97	5.65	0.00	0.99	0.32	0.02	118	224
Exact match + NN-3, radius 0.01	0.03	0.00	13.97	6.85	0.00	0.98	0.32	0.02	99	195
B. EMOVI										
Epanechnikov kernel, bw=0.01	0.040	0.00	12.16	0.87	0.00	0.99	0.68	0.00	1024	3376
NN-3, radius 0.01	0.042	0.00	12.16	0.91	0.00	0.99	0.68	0.00	1024	1691
Exact match + NN-3, radius 0.01	0.042	0.00	12.16	1.38	0.00	0.99	0.68	0.00	956	1637

Source: Authors' calculations.
Notes: The first column indicates the matching method. NN refers to nearest neighbor matching. The exact matching method restricts the sample to individuals within rural or urban areas and exact age for EMOVI, and for ENNVIH it also restricts it to individuals with the same school attendance status. "DW test" refers to the Dehejia and Wahba (1999) stratification test using quintiles of the estimated propensity score. The "Median Bias" column shows the median standardized bias. The "LR test" column shows the p-value of the likelihood ratio test that all coefficients in the regression are equal to zero. The column "Diff Means" shows the percentage of tests out of total possible tests in which the null hypothesis of equal means between treatment and control is rejected. The last two columns indicate the number of observations in treatment and control after matching. For details on the variables included in the propensity score estimation, please refer to the text.

Table 3 provides the balance tests for stratification (Dehejia and Wahba 1999, 2002), standardized bias (Rosenbaum and Rubin 1985), likelihood ratio (Sianesi, 2004), the difference in means, and the number of observations after matching. We include only three matching methods for each survey (a full set of results can be found in Table A1 of the appendix). The matching method is successful in balancing treatment and control groups. After matching, there are no signficant differences in observable characteristics between treatment and control. However, balance is relatively more difficult to achieve with MxFLS than with EMOVI as measured by the standardized median bias and the difference in means. Nonetheless, the values are very small and fall within the region of "sufficient" balance mentioned by Caliendo and Kopeinig (2008). In the appendix, we show that balance is more successfully achieved in a model in which the propensity score excludes interaction terms and only includes linear terms. But since excluding important variables may bias the ATT estimates, we present the main results using the estimated propensity score with interactions and squares, and as a robustness exercise we show the ATT results using the model with linear terms.

5.2. Short-run impacts

Table 4 shows the main results using MxFLS with a difference-in-difference ATT. For a simple comparison, we also include the estimate using regression analysis, although as previously mentioned, the PSM estimates are more reliable due to a similar comparison between treatment and control. The table includes the individual outcomes of years of schooling, school attendance, marriage, working, hours of work, and whether the individual left the household by 2005. The table also includes outcomes at the household level, restricting the sample to females who did not leave the household during the period of study.

The results provide evidence that a teenage pregnancy reduces school attainment. Females who had a child between 2002 and 2005 or 2006 have 0.6-0.8 years less of schooling than a female who did not have a child. The estimate is statistically significant, although with relatively large standard errors. If they drop out of school permanently, we should expect the gap to grow; if they drop out temporarily, we should observe a reduction in the gap in the long run, or that the gap remains constant if women select the age to drop out of school. We also find that school attendance decreases. However, it is important to point out that not all teenagers who became pregnant dropped

Table 4. Short-run results, MxFLS

Individual outcomes	**Yrs school**	**Attendance**	**Married**	**Working**	**Hrs work**	**Left HH**
Epanechnikov kernel, bw=0.01	-0.652 [0.220]	-0.283 [0.074]	0.573 [0.059]	-0.176 [0.071]	-10.587 [3.503]	0.417 [0.061]
NN3, Radius 0.01	-0.685 [0.260]	-0.318 [0.082]	0.574 [0.062]	-0.139 [0.078]	-9.127 [3.841]	0.423 [0.063]
Exact match + NN-3, radius 0.01	-0.866 [0.333]	-0.267 [0.076]	0.548 [0.084]	-0.088 [0.105]	-6.475 [4.734]	0.394 [0.082]
Regression	-0.819 [0.153]	-0.281 [0.039]	0.614 [0.027]	-0.168 [0.041]	-10.443 [2.013]	0.429 [0.025]
Household outcomes	**Total hours of work**	**Parents hours of work**	**Income per capita**	**HH size**		
Epanechnikov kernel, bw=0.01	-1.63 [14.19]	2.31 [8.35]	69.71 [283.37]	1.10 [0.28]		
NN3, Radius 0.01	2.29 [15.89]	1.69 [9.21]	9.31 [324.64]	1.23 [0.29]		
Exact match + NN-3, radius 0.01	3.99 [18.45]	-0.33 [11.25]	117.85 [415.82]	1.32 [0.32]		
Regression	3.99 [6.57]	-0.33 [4.13]	117.85 [146.78]	1.32 [0.11]		

Source: Authors' calculations.
Notes: The model includes linear and interaction terms; in total, the estimated propensity score includes 108 variables. Exact matching restricts the sample to individuals within the same rural/urban, age, and school attendance cells. The first panel includes outcomes at the individual level. The second panel restricts the sample to females who did not leave the household between 2002-2005 and analyzes outcomes at the household level. Standard errors are estimated using 500 bootstrap replications. The "Regression" row includes the estimates of teenage pregnancy controlling for 26 variables: individual characteristics, rural status, parental and household characteristics, and assets at the household level.

out of school by 2005-2006. The estimate implies that between 27 and 32% of teenagers who became pregnant are not attending school after pregnancy, compared to similar teenagers in the control group.

A key difference from results in the literature on the United States is that teenage pregnancy does not reduce the likelihood of marriage. In fact, a larger share of childbearing teenagers are married as compared to similar childless teenagers. These results are very possibly due to cultural differences between Mexico and the United States. In general, Mexican females tend to marry more frequently and teenage out-of-wedlock pregnancies are severely stigmatized by Mexican society. In the extension section, we analyze outcomes for teenage out-of-wedlock pregnancy.

Additionally, there is some evidence that teenage pregnancy reduces the probability of working by 9-18 percentage points. However, the standard errors are large and in the case of exact matching the results are not statistically significant. But there is statistical evidence that teenagers who became pregnant reduce their hours of work by 6-10 hours on average. Also, teenagers who became pregnant are 39-41% more likely to leave their household than teenagers who did not become pregnant. This latter finding is a result of marriage.

It is important to analyze not only the consequences of childbearing by teenagers themselves, but the consequences for the family of origin. This is interesting but hard to measure. As we analyze longitudinal data, we observe households in two periods. But if the teenager leaves the household, we are only able to observe information for the newly formed family. We could link the information to the family of origin, but in this case the interpretation of the treatment effect would not be clear, given that the treatment on the family of origin is somewhat lost. For these reasons, we focus on teenagers who did not leave the household of origin during the period of study. Thus, we are comparing how the family is affected in the short run when a teenager becomes pregnant.

The bottom panel in Table 4 includes the results at the household level. For females who did not leave the household of origin, we observe little changes at the household level. There is no evidence that the family reacts with more hours of work (this variable excludes the labor supply of the childbearing adolescent). The results are close to zero and not statistically significant. This effect may be due to more hours of work of parents and fewer hours of work of siblings. In order to test for this possibility, we estimate the effect on parents' labor supply (as shown in the next column). However, the estimates are not statistically signifcant for hours of work of parents. There seems to be no adjustment in the labor supply of other household members. This could be due to the timing of data collection. We observe teenagers after the birth of their child, and it is possible that the household has already adjusted to previous levels of hours of work. We also do not find any significant effect on income per capita, but there is a clear increase in household size. The reason that the effect on household size is greater than one is that some teenagers became pregnant and their husband or partner moved in with her and her family. In sum, we find little evidence that a pregnancy for a teenager who stays in the household of origin has significant consequences for the family of origin itself. It is important to stress that we do not measure the immediate effects of pregnancy but rather an average of 1 to 2 years after pregnancy.

5.3. Long-run impacts

Table 5 presents the estimates using EMOVI, for both PSM and regression. Women who became pregnant when they were teenagers attain less schooling than females who did not become pregnant. We find that the difference is close to 1 year of education. Although the estimate is larger than the short-run results, it is not possible to reject the null hypothesis of equal effects. However, the results do not support the hypothesis that the gap in education is reduced in the long run. On the contrary, once a teenage pregnancy occurs, the difference in years of education will be maintained. The estimate using regression analysis is much lower, but this is likely due to a lack of adequate controls (in the PSM framework we are controlling for many more variables including interaction and non-linear terms).

Females who became pregnant while adolescents are more likely to be married, and in turn less likely to be single in the long run than their counterparts. At the same time, they are more likely to go through a divorce or separation. Hence, we do not find any evidence in the short or long run that a teenage pregnancy reduces the likelihood of marriage. Also, it seems that a teenage pregnancy is considered as an "extra child", otherwise they would have had the same total number of children as the control females. Moreover, the increase in the number of children results in a larger household size. As for the impact on the labor supply, although the effect of teenage pregnancy on work is

Table 5. Long-run results, EMOVI

	Yrs school	Married	Single	Separated	# Children	Works	HH size	Income per capita
Epanechnikov kernel, bw=0.01	-1.065	0.050	-0.124	0.051	1.085	-0.043	0.555	-322.322
	[0.106]	[0.017]	[0.012]	[0.013]	[0.049]	[0.018]	[0.065]	[72.89]
NN3, radius 0.01	-1.122	0.059	-0.132	0.048	1.120	-0.054	0.560	-338.132
	[0.169]	[0.022]	[0.017]	[0.016]	[0.066]	[0.024]	[0.086]	[140.32]
Exact match + NN-3, radius 0.01	-1.159	0.045	-0.108	0.042	1.062	-0.043	0.528	-401.429
	[0.168]	[0.024]	[0.018]	[0.020]	[0.067]	[0.026]	[0.089]	[128.12]
Regression	-0.414	0.022	-0.033	0.011	-0.163	-0.011	-0.217	286.60
	[0.128]	[0.019]	[0.015]	[0.014]	[0.055]	[0.019]	[0.064]	[140.28]

Source: Authors' calculations.
Notes: The model includes linear and interaction terms; in total, the estimated propensity score includes 57 variables. Exact matching restricts the sample to individuals within the same rural/urban, age, and school attendance cells. Standard errors are estimated using 500 bootstrap replications.

negative, it is not statistically significant. Hence, there is no evidence that having children as an adolescent reduces the likelihood of working in the long run. However, there is some evidence of a lower income per capita in the household, which is most likely a consequence of a lower educational attainment.

5.4. Extensions and robustness checks

In the previous sections, we have not analyzed outcome for pregnancies out of wedlock. It is possible that out-of-wedlock pregnancies are more costly to teenagers. The MxFLS identifies the year of pregnancy and the year of marriage. We restrict the treatment sample to females who are not married in 2005 and females who had a birth before marriage, reducing it to 76 observations instead of 131.[22] Table 6 shows the estimates for this sample.

There are no large differences between the estimates using the full sample and restricting it to out-of-wedlock pregnancies. Both the loss in years of education and the reduction in the percent working are similar to the full sample. Since we dropped pregnancies after marriage from the sample, the effect on marriage decreases but it is still high and close to 34%. Hence, there is no evidence that out-of-wedlock pregnancies are different than teenage pregnancies in a marriage. Table 6 also includes results for the EMOVI, restricting the sample to females between 25 and 39 years old. There is no evidence that the loss in years of education or the probability of working is different from the full sample. However, the percentage that is married is relatively higher than in the full sample, although we cannot reject the hypothesis of equal coefficients.

In Table 7 we provide robustness results using more matching methods and results employing a different estimated propensity score. Panel A shows the main propensity score, which includes the interactions and squares of many variables. Results are robust to changes in the matching method. Panel B modifies the estimated propensity score by including only linear terms. In total, we include only 27 and 26 variables for the MxFLS and EMOVI, respectively. The ATT are, on average, similar to previous estimations, but the standard error is lower, as suggested by Bryson, Dorsett, and Purdon (2002). Panel C

22. Of the 76 observations, 41 are unmarried and 35 are married in 2005.

Table 6. Extensions: Short-run effects of teenage, out-of-wedlock pregnancy and teenage pregnancy effects at 25-39 years old

	MxFLS: Unmarried teenagers			EMOVI: 25-39 years old		
	Yrs school	**Works**	**Marriage**	**Yrs school**	**Works**	**Marriage**
Epanechnikov kernel, bw=0.01	-0.616	-0.154	0.345	-1.177	-0.065	0.111
	[0.341]	[0.115]	[0.081]	[0.152]	[0.025]	[0.026]
NN3, radius 0.01	-0.608	-0.157	0.343	-1.061	-0.061	0.115
	[0.401]	[0.128]	[0.083]	[0.204]	[0.032]	[0.032]
Exact match + NN-3, radius 0.01	-0.813	-0.122	0.306	-1.226	-0.045	0.100
	[0.473]	[0.141]	[0.115]	[0.215]	[0.034]	[0.035]

Source: Authors' calculations.
Notes: The model includes linear and interaction terms; in total, the estimated propensity score includes 108 variables for MxFLS and 57 variables for EMOVI. Exact macthing restricts the sample to individuals within the same rural/urban, age, and school attendance (for MXFLS) cells. Standard errors are estimated using 500 bootstrap replications.

Table 7. Robustness tests

	MxFLS			EMOVI			
	Yrs school	**Works**	**Marriage**	**Yrs school**	**Works**	**Marriage**	**Income per capita**
A. Propensity score with interactions and squares							
Epanechnikov kernel, bw=0.0025	-0.539	-0.153	0.624	-1.099	-0.047	0.051	-345.757
	[0.327]	[0.112]	[0.073]	[0.118]	[0.020]	[0.017]	[82.12]
Gaussian kernel, bw=0.01	-0.615	-0.111	0.598	-1.084	-0.044	0.053	-305.646
	[0.259]	[0.073]	[0.062]	[0.103]	[0.018]	[0.017]	[71.94]
NN3, radius 0.025	-0.610	-0.147	0.574	-1.129	-0.054	0.060	-334.045
	[0.313]	[0.093]	[0.069]	[0.168]	[0.024]	[0.022]	[140.25]
B. Propensity score with linear terms							
Epanechnikov kernel, bw=0.01	-0.837	-0.177	0.622	-1.158	-0.043	0.059	-318.719
	[0.177]	[0.050]	[0.046]	[0.145]	[0.017]	[0.018]	[75.10]
NN3, radius 0.01	-1.005	-0.217	0.616	-1.163	-0.039	0.069	-279.820
	[0.238]	[0.063]	[0.050]	[0.303]	[0.051]	[0.044]	[165.54]
Epanechnikov kernel, bw=0.0025	-0.879	-0.177	0.609	-1.137	-0.043	0.065	-313.687
	[0.210]	[0.053]	[0.046]	[0.156]	[0.019]	[0.019]	[165.57]
NN3, radius 0.025	-1.021	-0.217	0.619	-1.157	-0.040	0.070	-279.689
	[0.232]	[0.064]	[0.050]	[0.302]	[0.051]	[0.044]	[80.72]

Table 7. (continued)

	MxFLS			EMOVI			
	Yrs school	**Works**	**Marriage**	**Yrs school**	**Works**	**Marriage**	**Income per capita**
C. (A) + matching in the log odds ratio							
Epanechnikov kernel, bw=0.01	-0.564	-0.123	0.589	-1.090	-0.045	0.054	-306.752
	[0.213]	[0.068]	[0.059]	[0.110]	[0.019]	[0.017]	[76.69]
NN3, radius 0.01	-0.626	-0.104	0.591	-1.130	-0.055	0.060	-336.757
	[0.257]	[0.072]	[0.062]	[0.171]	[0.024]	[0.022]	[141.56]
Epanechnikov kernel, bw=0.0025	-0.536	-0.147	0.618	-1.068	-0.042	0.050	-341.993
	[0.257]	[0.089]	[0.065]	[0.128]	[0.020]	[0.019]	[104.20]
NN3, radius 0.025	-0.600	-0.169	0.565	-1.132	-0.055	0.060	-334.589
	[0.264]	[0.085]	[0.064]	[0.171]	[0.024]	[0.022]	[140.64]

Source: Authors' calculations.
Notes: Panel A includes linear and interaction terms; in total, the estimated propensity score includes 108 variables for MxFLS and 57 variables for EMOVI. Panel B includes only linear terms in the estimation of the propensity score. We include 27 and 26 variables in the MxFLS and EMOVI, respectively. Standard errors are estimated using 500 bootstrap replications.

matches on the log of odds ratio of the main estimated propensity score as suggested by Heckman, Ichimura, and Todd (1997). In general, the impact on years of schooling in the short run varies from -0.56 to -1 and in the long run from -1.09 to -1.16. Both are within the standard errors obtained for the main estimates. The impact on income per capita in the long run is consistently negative and varies from -279 to -346. In sum, the main estimates are robust to the matching method and to the estimated propensity score.

6. Conclusions

In this paper we estimate the effect of teenage childbearing on several outcomes for the teenage mother and her family of origin in the short run, and also the long-run effects on the mother. The identification of the causal effect of teenage childbearing has proven to be very elusive due to selection bias: Those adolescents who give birth to a child are sistematically different from adolescents who do not have children. For instance, we find that in the case of Mexico, treated teenagers tend to be more sexually active before pregnancy and come from more disadvantaged backgrounds.

We attempt to solve this selection problem by implementing a propensity-score matching model using two different data sources: a longitudinal survey (Mexican Family Life Survey, or MxFLS), and a cross-section survey designed to measure mobility in Mexico (Social Mobility Survey, or EMOVI). Therefore, we have information on the individual and her household when she was 14 years old. The MxFLS allows us to estimate the short-run effects on the teenage mother and her family of origin. On the other hand, the EMOVI enables us to estimate the long-run effects on the teenage mother. However, the estimates from MxFLS are more reliable as they allow us to estimate difference-in-differences models.

According to our results, the single most important effect of teenage childbearing is a lower educational attainment of the teenage mother, both in the short and long run. As a result, we find that in the long run the households of those females who had their first child as teenagers tend to have a lower income per capita. We also find that in the short run, teenage mothers reduce their college attendance (hence the lower educational attainment) and reduce their labor supply. We do not find any significant effects on labor supply of other household members in the short run, nor on the labor supply of the teenage mothers themselves in the long run. Finally, and in contrast with the literature in the United States, we find that having a child during adolescence has a positive effect on the probability of being married. This difference is most likely a result of cultural differences between Mexico and the United States.

Although still greatly debated, there is evidence that teenage childbearing is associated with higher levels of poverty and welfare dependence in the United States. To our knowledge, there is not a large literature on the effects of teenage childbearing for developing countries. This paper contributes to fill that gap in the literature. Our findings provide evidence that teenage childbearing has adverse effects in the Mexican context. The fact that teenage childbearing prevents teenage mothers from continuing their human capital investments shows that teenage childbearing may have a deleterious effect on the probability of living in a poor household. Moreover, given that there is little social mobility in Mexico (Torche, 2010), teenage childbearing may be a gateway into an intergerational poverty trap. As such, our work has two important policy implications. First, programs aimed at preventing teenage pregnancies, such as sexual education during primary and secondary education, should be expanded, as should access to contraceptives through public health systems. Second, once a teenager becomes pregnant, the state should provide support in the form of childcare and merit scholarships, to prevent the teenage mother from dropping out of school. The latter measure is partially being addressed by PROMAJOVEN. However, the program is still limited to the poorest population even though not all teenage mothers meet that criterion.

REFERENCES

Ashcraft, A. and K. Lang (2006), "The consequences of teenage childbearing," Working Paper 12485, National Bureau of Economic Research.

Blunch, N. (2011), "Skills, schooling and non-marital teenage pregnancy in Ghana," unpublished manuscript, Washington and Lee University.

Bronars, S.G. and J. Grogger (1994), "The economic consequences of unwed motherhood: Using twin births as a natural experiment," *The American Economic Review* **84**(5): 1141–56.

Bryson, A., R. Dorsett, and S. Purdon (2002), "The use of propensity score matching in the evaluation of labour market policies," Working Paper 4, Department for Work and Pensions.

Caliendo, M. and S. Kopeinig (2008), "Some practical guidance for the implementation of propensity score matching," *Journal of Economic Surveys* **22**(1): 31–72.

Cardoso, A.R. and D. Verner (2007), "Youth risk taking behavior in Brazil: Drug use and teenage pregnancies," Discussion Paper Series 3030, Institute for the Study of Labor.

Chevalier, A. and T.K. Viitanen (2003), "The long-run labour market consequences of teenage motherhood in Britain," *Journal of Population Economics* **16**(2): 323–43.

Dehejia, R.H. and S. Wahba (1999), "Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs," *Journal of the American Statistical Association* **94**(448): 1053–62.

———. (2002). "Propensity-score matching methods for nonexperimental causal studies," *The Review of Economics and Statistics* **84**(1): 151–61.

Echarri Cánovas, C. and J. Pérez Amador (2007), "En tránsito hacia la adultez: eventos en el curso de vida de los jóvenes en México," *Estudios demográficos y urbanos* **22**(1): 43–77.

Ferre, Z., M. Gerstenblüth, M. Rossi and P. Triunfo (2009), "The impact of teenage childbearing on educational outcomes," unpublished manuscript, Society for the Study of Economic Inequality.

Geronimus, A.T. and S. Korenman (1992), "The socioeconomic consequences of teen childbearing reconsidered," *The Quarterly Journal of Economics* **107**(4): 1187–214.

Heckman, J.J., H. Ichimura and P.E. Todd (1997), "Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme," *The Review of Economic Studies* **64**(4): 605–54.

Hoffman, S.D. and R.A. Maynard (2008), *Kids having kids: Economic costs and social consequences of teen pregnancy, 2nd edition*. Washington, D.C.: Urban Institute Press.

Holland, P.W. (1986), "Statistics and causal inference," *Journal of the American Statistical Association* **81**(396): 945–60.

Hotz, V.J., S.W. McElroy, and S.G. Sanders (2005), "Teenage childbearing and its life cycle consequences: Exploiting a natural experiment," *The Journal of Human Resources* **40**(3): 683–715.

Kruger, D., M. Berthelon and R. Navia (2009), "Adolescent motherhood and secondary schooling in Chile," Discussion Paper Series 4552, Institute for the Study of Labor.

Leuven, E. and B. Sianesi (2003), "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing." Statistical Software Components, Boston College, Department of Economics.

Levine, D.I. and G. Painter (2003), "The schooling costs of teenage out-of-wedlock childbearing: Analysis with a within-school propensity-score-matching estimator," *Review of Economics and Statistics* **85**(4): 884–900.

Marteleto, L., D. Lam, and V. Ranchhod (2008), "Sexual behavior, pregnancy, and schooling among young people in urban South Africa," *Studies in Family Planning* **39**(4): 351–368.

Menkes, C. and O. Serrano (2010), "Condicionantes sociodemográficos del embarazo adolescente en México," Working paper, *X Reunión Nacional de Investigación Demográfica en México.*

Menkes, C. and L. Suárez (2003), "Sexualidad y embarazo adolescente en México," *Papeles de población* **35**: 233–62.

Navarro Paniagua, M. and I. Walker (2010), "The impact of teenage motherhood on the education and fertility decisions of their children: Evidence for Europe," unpublished manuscript, XXIV Annual Conference of the European Society for Population Economics.

Pantelides, E.A. (2004). "Aspectos sociales del embarazo y la fecundidad adolescente en América Latina," *Revista Notas de Población* **78**: 7–33.

Ranchhod, V., D. Lam, M. Leibbrandt, and L. Marteleto (2011), "Estimating the effect of adolescent fertility on educational attainment in Cape Town using a propensity score weighted regression," in *Fifth Annual PopPov Conference on Population, Reproductive Health, & Economic Development.*

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

———. (1985) "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score," *The American Statistician* **39**(1): 33–8.

Sianesi, B. (2004), "An evaluation of the Swedish system of active labor market programs in the 1990s," *The Review of Economics and Statistics* **86**(1): 133–55.

Smith, J. and P.E. Todd (2005), "Does matching overcome Lalonde's critique of nonexperimental estimators?" *Journal of Econometrics* **125**(2): 305–53.

Stern, C. (2007), "Estereotipos de género, relaciones sexuales y embarazo adolescente en las vidas de jóvenes de diferentes contextos socioculturales en México," *Estudios Sociológicos* **25**(73): 105–29.

Stern, C. (2012), *El "problema" del embarazo en la adolescencia. Contribuciones a un debate*. México: El Colegio de México.

Torche, F. (2010), "Cambio y persistencia de la movilidad intergeneracional en Mexico," in Serrano, J. and F. Torche, eds., *Movilidad social en Mexico. Poblacion, desarrollo y crecimiento*: 71–134. Mexico City: Centro de Estudios Espinosa Yglesias.

APPENDIX

Table A1. Balance in the propensity score: robustness

	DW Test		Median Bias		LR Test		Diff Means		# Treat	# Control
	Before	After	Before	After	Before	After	Before	After		
A. MxFLS										
Epanechnikov kernel, bw=0.0025	0.03	0.00	13.97	4.34	0.00	0.99	0.32	0.05	97	597
NN-1, radius 0.01	0.03	0.00	13.97	8.95	0.00	0.06	0.32	0.03	118	99
Gaussian kernel, bw=0.01	0.03	0.00	13.97	10.06	0.00	0.99	0.32	0.11	131	868
NN-3, radius 0.025	0.03	0.00	13.97	4.64	0.00	0.99	0.32	0.02	122	226
Propensity score with linear terms										
Epanechnikov kernel, bw=0.01	0.06	0.00	10.43	1.63	0.00	0.99	0.24	0.00	125	862
NN-3, radius 0.01	0.06	0.00	10.43	3.65	0.00	0.99	0.24	0.00	125	269
Exact match + NN-3, radius 0.01	0.06	0.00	10.43	6.71	0.00	0.98	0.24	0.00	114	236
Kernel Epanechnikov, bw=0.0025	0.06	0.00	10.43	3.48	0.00	0.99	0.24	0.00	120	674

Table A1. (continued)

	DW Test		Median Bias		LR Test		Diff Means		# Treat	# Control
	Before	After	Before	After	Before	After	Before	After		
B. EMOVI										
Epanechnikov kernel, bw=0.0025	0.042	0.00	12.16	0.45	0.00	0.99	0.68	0.00	1012	3302
NN-1, radius 0.01	0.042	0.00	12.16	1.43	0.00	0.99	0.68	0.00	1024	754
Gaussian kernel, bw=0.01	0.042	0.00	12.16	0.59	0.00	0.99	0.68	0.00	1030	3378
NN-3, radius 0.025	0.042	0.00	12.16	0.98	0.00	0.99	0.68	0.00	1030	1691
Propensity score with linear terms										
Epanechnikov kernel, bw=0.01	0.08	0.00	10.71	0.33	0.00	0.99	0.62	0.00	1025	3377
NN-3, radius 0.01	0.08	0.00	10.71	1.45	0.00	0.99	0.62	0.00	1025	1781
Exact match + NN-3, radius 0.01	0.08	0.00	10.71	0.93	0.00	0.99	0.62	0.00	972	1687
Epanechnikov kernel, bw=0.0025	0.08	0.00	10.71	0.75	0.00	0.99	0.62	0.00	1019	3327

Source: Authors' calculations.
Notes: The first column indicates the matching method. NN refers to nearest-neighbor matching. The exact matching method restricts the sample to individuals within rural or urban areas and exact age for EMOVI, and for ENNVIH it also restricts the sample to individuals with the same school attendance status. "DW test" refers to the Dehejia and Wahba (1999) stratification test using quintiles of the estimated propensity score, the "Median Bias" column shows the median standardized bias, while the "LR test" column shows the p-value of the likelihood ratio test that all coefficients in the regression are equal to zero. The "Diff Means" column shows the percentage of tests, of the total possible tests, in which the null hypothesis of equal means between treatment and control is rejected. The last two columns indicate the number of observations in treatment and control after matching. We include 86 variables in 2002 in the estimation of the propensity score for MxFLS including interactions.We include 57 variables in the estimation of the propensity score for EMOVI. The models with the "Propensity score with linear terms" do not include interaction terms. In this case, MxFLS includes 27 variables and EMOVI includes 26 variables.

VOL. 51 NO. 1 (MAY, 2014), 147–178

LATIN AMERICAN JOURNAL OF ECONOMICS

A COHORT ANALYSIS OF THE COLLEGE PREMIUM IN MEXICO*

FRANCISCO BENITA**

This paper provides the first empirical evidence for Mexico about relative wage differences between college-educated and high-school-educated workers across five-year age groups. Rotating panel surveys are used to implement an imperfect substitution model for similar male workers between different age groups and between the two education groups. For the period 2005-2012, the results suggest a partial elasticity of substitution of 1.7 for college- and high-school-educated workers and a partial elasticity of substitution of about 3 across age groups. Remarkably, the wage gap between younger and older workers with the same education level increased after the economic crisis of 2008.

JEL classification: J21, J31, C33

Keywords: Wage differentials, cohort effects, rotating panel, elasticity of substitution, Mexico

1. INTRODUCTION

Human capital theory proposes that education improves labor force productivity by imparting useful knowledge and skills. Over the past 40 years, many studies have examined rising rates of returns to education. Most of these studies have shown that schooling is a key factor in explaining wage gaps among individuals. On the other hand, in developed and some developing countries, average schooling has been increasing steadily over time, resulting in an expanded rise in salaries and a substantial narrowing of the wage gap. A second phenomenon of interest is the aging of the population, mainly occurring in developed countries, and more recently in some developing countries. In these countries there will be a higher absolute number of elderly people, who will make up a larger share of the population and have longer healthy life expectancies, and there will be a relatively lower number of working-age people.

The Mexican case is particularly interesting in this context for several reasons. First, Mexico's labor structure has undergone various political,

* I would like to thank the editor and two anonymous reviewers for their valuable comments and suggestions.

** Department of Systems & Industrial Engineering, ITESM (Tec de Monterrey), Campus Monterrey, Av. Eugenio Garza Sada Sur 2401, Monterrey, NL, 64849, Mexico. Email: francisco_benita@hotmail.com.

doi 10.7764/LAJE.51.1.147

economic, and demographic changes, affecting both increases in schooling and growth in the elderly population over the past three decades. Second, wage inequality trends have been substantially different than those observed in other developing countries (Tello and Ramos, 2012). According to the Mexican Population and Housing Census, the average number of years of education in the labor force has strongly increased, from 3.4 years in 1970 to 7.5 years in 2000, and this growth trend has continued, reaching 8.8 years of education in 2010. In addition, the country has seen an increase in the average age, from 22 years in 2000 to 26.7 years in 2010. Like many parts of the world, Mexico has witnessed a growth in the proportion of its elderly population compared to all other age groups.

Both demographic (aging) and economic (schooling) effects imply relative changes in the labor supply, with possible strong variations in returns to education between age groups. Growth in the pool of young, highly educated workers has resulted in an increase in the country's labor supply. Likewise, it is possible that the gap in average earnings between less-educated and highly educated workers narrows if labor demand is unchanged. This long-term phenomenon raises several questions about the future development of wage and employment structures by age groups. The issue is relevant for different economic topics, such as social security actuarial calculations, increasing wage inequality, age-related government transfers, pensions, health care, and labor market reforms.

There is a large literature on education's effect on earnings in Mexico, focused on factors that largely explain the wage gap between highly educated and less-educated workers resulting from changes in labor demand (e.g., Cragg and Epelbaum, 1996, Hanson and Harrison, 1999, Hernández, 2000, Cañonero and Werner 2002, and Chiquiar, 2008, among others). However, there is no empirical evidence on this effect according to different age groups of workers. Hence, the objective of this paper is to test whether the increase over the last decade in the number of male college graduates has led to downward pressure on the college wage premium. Furthermore, it seeks to determine whether the increase in relative supplies could possibly lead to a narrowing of the wage gap between college-educated male workers and male workers who possess only a high school diploma.

Because female labor-market participation in developing countries is the result of an underlying choice process based on utility maximization (Becker, 1965) and the modeling requires appropriate statistical methods to control self-selection bias, this study is limited to male workers. Following the econometric methodology proposed by Card and Lemieux (2001) and using rotating panel data from 2005 to 2012, I present the partial elasticity of substitution between college and high school workers and between different age groups. Since empirical evidence suggests that younger workers cannot easily replace older workers (Rones, 1983, Hairault *et al.*, 2007, Bia *et al.*, 2009, or Kalwij *et al.*, 2010), I use an imperfect substitution model for similar male workers between different age groups. In addition, I estimate not only the elasticity of substitution between younger and older male workers but also the elasticity of substitution of workers with a college degree and those with only a high school diploma. In developing countries such as Mexico, estimation of the education wage premium by age groups and its elasticity is important for determining the allocation of public spending on educational policies, which should be oriented toward educational improvement and raising worker productivity.

The model represents a competitive equilibrium from the firm side, where entrepreneurs select the amount of labor required to maximize profits. The estimation results indicate that there is a large elasticity substitution (about 3) between male workers with different levels of education, which could be interpreted to mean that college- and high-school-educated workers are considered easily interchangeable by employers. Furthermore, the small value of the estimated parameter—around 1.7—between the two education groups suggests that young and old workers are considered different by employers and therefore they are far from perfect substitutes. This result is very important because it suggests that in the current Mexican labor market it is easier to substitute skilled workers with unskilled workers than to replace older workers with a younger labor force.

The remainder of this paper is organized as follows. Section 2 provides a review of the related literature, while Section 3 outlines the theoretical framework. Section 4 presents a characterization of the estimation strategy and the dataset description. The results and some discussion are presented in Section 5, while Section 6 contains concluding remarks.

2. Literature review

Changes in labor supply and demand have no specific effect on relative wages, particularly between skilled and unskilled workers. Katz and Murphy's (1992) seminal work on movements in the college wage premium from 1963 to 1987 in the United States concludes that the premium appears to be strongly related to fluctuations in the growth rate of the supply of college graduates. Understanding these relationships is very important for economic development since labor supply affects the environment in which the labor market functions.

In general, we note that the wage gap in Mexico has been widely discussed, especially during the 1980s and the first half of the 1990s. It is generally agreed that the wage gap between college-educated male workers and those with only a high school diploma is due to changes in labor demand, which are explained using two different approaches: trade liberalization and technological change.

In relation to trade liberalization, changes have occurred as a result of employers' strategies to reward the qualifications of their employees in order to increase their international competitiveness. In regard to technological change, reduction of trade barriers and the rise in foreign direct investment have led to increased use of technology, which in turn has driven demand for skilled workers to complement the technology. As a result, the relative productivity of skilled workers, mainly those with college and advanced degrees, has been increased by such technological changes.

Following the focus on trade liberalization, in the view of Zepeda and Ghiara (1999) the trade reforms that took place in Mexico between 1986 and 1990 were accompanied by important wage gaps between skilled and unskilled workers. Applying a quantile regression to estimate the conditional distribution of income, Zamudio (2001) finds that between 1984 and 1996 the distribution of income depends to a large extent on education. Cañonero and Werner (2002) study the widening of wage differentials for skilled and unskilled workers after Mexico's entry into the General Agreement on Tariffs and Trade (GATT) and find that the average relative wage of an unskilled worker had decreased by more than 20% by the end of 1990. In a similar context, Hanson (2003) focuses on policy reforms initiated during the 1990s, finding a significant rise in demand for skilled workers in the country and concluding that the policies reduced profits in industries that paid their workers high wages prior to the reform and raised the premium paid to workers in states along the U.S. border. Airola and Juhn (2008) argue that trade liberalization per se did not widen wage differentials in the region. Finally, in the view of Chiquiar (2008), in the period following implementation of the North

American Free Trade Agreement (NAFTA), the wage premium for high-skilled workers increased in what he calls the "second stage of globalization in Mexico."

Following the second approach, Hanson and Harrison (1995, 1999) suggest that the rising wage gap during the 1980s was associated with changes internal to industries and even internal to plants. Cragg and Epelbaum (1996) report significant changes in returns to specific occupations, in particular professionals and managers, but they believe that rising inequality in the early 1990s is largely due to trade liberalization reforms. In a similar study, Meza (1999) concludes that if the skills premium had not risen over the 1988-1993 period, the wage distribution would be improved. Finally, Hernández (2000) points out that technological changes associated with trade liberalization have promoted an incipient territorial decentralization focused on the U.S. border and the periphery of Mexico City.

Among more recent studies, Huesca (2004) finds that workers with high school and college education show higher returns to schooling because of an upward trend in demand for skilled workers. Lopez (2006) notes that the Mexican population is becoming more educated. The author shows how years of schooling have increasing over the 1960-1990 period, growing from 2.76 years in 1960 to 6.72 years in 1990, a difference of nearly four years. According to Ampudia (2007), trade liberalization and international markets increased demand for skilled labor, which in turn generated higher education wage premiums. For workers with a college degree, the education premium has been shown to be cyclical in times of recession.

Although there is extensive literature documenting the role of educational level on the wage gap in Mexico, most previous studies have focused on average returns to schooling rather than differences by age or cohort. While the rise in the average wage gap between college and high school workers has been briefly documented, the fact that the increases have varied among different age groups is not as well understood.

These studies analyze the evolution of returns to schooling under the assumption that different age groups with the same level of education are perfect substitutes in production processes. This assumption implies that the aggregate supply of each type of education can be obtained by simply adding the total number of workers in each education category, such that all education-related wage differentials in the labor market in any given year are proportional to the average college-high school wage gap that year.

An analysis of the evolution of wage differentials by age groups was presented for the first time in Card and Lemieux (2001). The authors extend the model proposed by Katz and Murphy (1992) in order to allow imperfect substitution between workers of different ages for the U.S., the United Kingdom, and Canada. In the study, the authors point out that shifts in the college-high school wage gap for younger men in the U.S. reflect changes in the relative supply of highly educated workers across age groups. In their estimation of the elasticity of substitution of age groups with the same education level, they find elasticity close to 5 between different age groups in the U.S., while the elasticity of substitution between college graduates and high school graduates ranges from 1.1 to 2.5, with similar results for both the U.K. and Canada.

Roger and Wasmer (2009) describe the importance of controlling for age and skill heterogeneity to explain labor productivity. Within this framework, some recent studies have continued the line of imperfect substitution of workers between different age groups. Ferreira (2004) tests the existence of a causal relationship between the evolution of the college-educated labor supply and the behavior of the college premium in Brazil by decomposing the relative supply of college-educated workers into two components: an age-specific component and a cohort-specific component. The author's main results suggest that the parameter for elasticity of substitution between different age groups is similar to the U.S., while the elasticity of substitution between two education groups is less than 2. Some implications of this result are outlined in Prskawetz *et al.* (2012), who introduce a model of optimal education policy at the macro level allowing for heterogeneity of the workforce with respect to its age and skills. Their estimations suggest that the relationship between the elasticity of substitution of labor across age groups plays a crucial role in the way demographic changes affect the optimal educational policy.

3. Background

This work follows the two-step estimation method introduced by Card and Lemieux (2001). The theoretical model has a production function that uses labor as the sole input. In this simple model, the input can be skilled (college graduates) or unskilled (high school graduates) combined under a CES technology. The aggregate output depends on two CES sub-aggregates of high school and college workers:

$$H_t = \left[\sum_j \left(\beta_{jt} H_{jt}^{\eta} \right) \right]^{1/\eta}, \tag{1}$$

$$C_t = \left[\sum_j \left(\alpha_{jt} C_{jt}^{\eta} \right) \right]^{1/\eta}, \tag{2}$$

where H_t and C_t are the total supply of high school and college labor in period t, and $\eta \in (-\infty,1]$ is a function of the partial elasticity of substitution, σ_A, between different age groups j with the same level of education ($\eta = 1 - 1/\sigma_A$). The α_j and β_j coefficients are relative efficiency parameters that vary over time. Perfect substitutability across age groups necessarily implies that $\eta = 1$. Therefore, total college or high school labor input is simply a weighted sum of the quantity of labor supplied by each group. In this work, the total supply of each type of work in each year is measured as the sum of the average weekly hours worked by members of the different education categories.

The model assumes an aggregate CES production function $y_t = (\theta_{ct} C_t^{\rho} + \theta_{ht} H_t^{\rho})^{1/\rho}$, where y_t is the total output in period t, and $\rho \in (-\infty,1]$ is a function of the elasticity of substitution σ_{E} between the two education groups ($\rho = 1 - 1/\sigma_E$). θ_{ct} and θ_{ht} are college and high-school productivity terms.

Under this specification, the marginal product of labor for a given age-education group depends on both the group's own supply of labor and the aggregate supply of labor in its education category. Formally, we have the following:

$$\frac{\partial \mathrm{y}_t}{\partial \mathrm{H}_{jt}} = \frac{\partial \mathrm{y}_t}{\partial \mathrm{H}_t} \frac{\partial \mathrm{H}_t}{\partial \mathrm{H}_{jt}} = \theta_{ht} H_t^{\rho-1} \Psi_t \beta_{jt} H_{jt}^{\eta-1} H_t^{1-\eta} = \theta_{ht} H_t^{\rho-\eta} \Psi_t \beta_{jt} H_{jt}^{\eta-1} \tag{3}$$

where $\Psi_t = (\theta_{ct} C_t^{\rho} + \theta_{ht} H_t^{\rho})^{1/\rho - 1}$. Similarly, the marginal product of college workers in age group j is $\partial y_t / \partial C_{jt} = \theta_{ct} C_t^{\rho-\eta} \Psi_t \alpha_{jt} C_{jt}^{\eta-1}$. Additionally, efficient utilization of different skill groups requires that relative wages are equal to relative marginal products:

$$\ln\left(w_{jt}^{H}\right) = \ln\left(\frac{\partial \mathrm{y_t}}{\partial \mathrm{H}_{jt}}\right) = \ln\left(\theta_{ht} H_t^{\rho-\eta} \Psi_t\right) + \ln\left(\beta_{jt}\right) - \frac{1}{\sigma_A}\ln\left(H_{jt}\right) \tag{4}$$

$$\ln\left(w_{jt}^{C}\right) = \ln\left(\frac{\partial \mathrm{y_t}}{\partial C_{jt}}\right) = \ln\left(\theta_{ct} C_t^{\rho-\eta} \Psi_t\right) + \ln\left(\alpha_{jt}\right) - \frac{1}{\sigma_A}\ln\left(C_{jt}\right) \tag{5}$$

From equations (4) and (5) it is possible to infer that the ratio of the wage rate of skilled workers in age group j, w_{jt}^{C}, to the wage rate of unskilled workers in the same age group, w_{jt}^{H}, satisfies:

$$\ln\left(\frac{w_{jt}^{C}}{w_{jt}^{H}}\right) = \ln\left(\frac{\theta_{ct}}{\theta_{ht}}\right) + \ln\left(\frac{\alpha_{jt}}{\beta_{jt}}\right) + \left(\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right)\ln\left(\frac{C_t}{H_t}\right) - \frac{1}{\sigma_A}\ln\left(\frac{C_{jt}}{H_{jt}}\right) \tag{6}$$

If relative employment ratios are taken as exogenous[1], Equation (6) leads to a simple model for the observed college-high school wage gap in age group j in period t:

$$\begin{aligned} r_{jt} \equiv \ln\left(\frac{w_{jt}^{C}}{w_{jt}^{H}}\right) = \ln\left(\frac{\theta_{ct}}{\theta_{ht}}\right) + \ln\left(\frac{\alpha_{jt}}{\beta_{jt}}\right) + \left(\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right)\ln\left(\frac{C_t}{H_t}\right) \\ - \frac{1}{\sigma_A}\ln\left(\frac{C_{jt}}{H_{jt}}\right) + e_{jt} \end{aligned} \tag{7}$$

where e_{jt} is the sampling variation in the measured gap and/or any other sources of variation in age-specific wage premiums.

The model represents a competitive equilibrium from the firm side, where entrepreneurs choose the amount of labor required to maximize profits. The equilibrium wage for the perfectly competitive market industry is computed by allowing imperfect substitution between workers of different ages.

As a result, the model specification represents a firm's production equilibrium. General equilibrium analysis deals explicitly with the

1. Total supply of high school and college labor in each period can be varied exogenously as a part of a scenario, e.g., simulating the effects of increasing participation of certain population groups. However, they do not respond endogenously to variations in other parameters such as the productivity or wage rate of college workers in each age group. In the case of Mexico, due to the effects of GATT and NAFTA on the economy, this can be assumed since input prices and technology can be more closely linked to tariff changes (Robertson, 2004).

interdependence of households and firms as mediated by markets. With only one production factor it is particularly easy to illustrate production equilibrium. However, the model could be extended by adding other inputs such as capital. Also, within the problem statement it must be clear that the results are valid only in a perfectly competitive equilibrium scenario. Therefore, the model is not applicable to imperfect markets with monopoly or oligopoly labor demand (e.g., monopsony or oligopsony). Finally, Equation (7) becomes the equation to estimate. Note that the assumption of employment ratios as exogenous becomes essential for the results; otherwise, the weighted least squares of elasticity of substitution will have a positive basis.

4. Estimation

The empirical analysis is based on data extracted from the National Survey of Occupation and Employment (ENOE). These datasets are collected by the National Institute of Statistics and Geography of Mexico (INEGI) and are the basis for official employment statistics. The datasets are based on household surveys conducted on a quarterly basis, the main objective of which are to obtain information about demographic and economic characteristics of employment. The surveys are designed as panel data with overlapping blocks of observations that are renewed after being followed up over five consecutive quarters. This data is representative of the whole country, each of the 32 states, and urban and rural areas.

I analyze datasets from the first quarter of 2005 through the fourth quarter of 2012. These rotating panel data are used to derive labor market conditions for male workers ages 25-59 with either a high school diploma or a college degree. In addition, an important difference from previous studies is that I consider full-time workers (those working 48 hours or more per week) as well as part-time workers (those working between 48 and 96 hours per week).

For reasons related to analysis and interpretation, the strategy consists of estimating panels by following the individuals from the first through the fourth quarter of each year. This technique allows the estimation of robust coefficients, increasing the degrees of freedom and reducing collinearity among the explanatory variables. In general, panel data sets possess major advantages over conventional cross-sectional data sets. The data set in question consists of eight panels, each one made up of individuals who are interviewed in each of the four quarters of the year. The top and bottom 1% of the log real hourly wage, using 2010 as the base year, were dropped in order to avoid atypical values.

For a general overview of the effect of a college education on wages, I first estimate the college wage premium for each year t with no distinctions between age groups, formally:

$$\ln\left(w_{iq}\right) = \zeta_0 + X_{iq}^{'}\zeta + v_{iq} \tag{8}$$

where w_{iq} contains the hourly wage of individual i in quarter q, $q \in \{1,2,3,4\}$. ζ denotes the coefficient vector, matrix X contains the regressors used to explain the dependent variable, and v is a residual assumed to be $N(0,\sigma^2)$. Because schooling cannot change in each quarter, $v_{iq} = u_i + \mu_{iq}$ and the model becomes a generalized least squares[2]. The regressor matrix X contains information about age and its quadratic, a married category dummy variable, six economic activity dummies, five regional dummies, and a dummy variable for college degree. The proposed model specification is based on previous studies of Mexico described in Section 2. Although Equation (8) is not directly related to the estimation of Equation (7), it is a useful first approach to understanding the college premium. The assumption behind Equation (8) is that different age groups with the same level of education (college or high school) are perfect substitutes in production.

However, in order to capture the college wage premium within age groups, for each year t I estimate separate regressions for each five-year age group, j:

$$\ln\left(w_{jq}\right) = \gamma_0 + X_{jq}^{'}\gamma + \varepsilon_{jq} \tag{9}$$

In Equation (9), the college degree dummy variable contained in matrix X measures the college wage premium by five-year age groups. Secondly, we compute Equation (7) using the estimated college premium from Equation (9), but there is a problem because the aggregate supplies of the two education groups depend on the elasticity of substitution across

2. The critical assumption when using random effects is that regressors are not endogenous. When the assumption of zero conditional mean of the error term given the regressors is not satisfied, the estimators will be inconsistent. This can be corrected by using instrumental variables estimation to find a variable correlated with wages but not with the error term. Further work is required to test endogeneity and instrumental variables regression should be used if the problem is detected.

age groups. Card and Lemieux (2001) suggest a two-step estimation procedure to identify both σ_A and σ_E.

The first stage of the two-step estimation consists of estimating σ_A from a regression of age-group-specific college wage gaps on age-group-specific relative supplies of college educated labor, cohort effects, and time effects:

$$r_{jt} = b_{t-j} + d_t - \frac{1}{\sigma_A} \ln\left(\frac{C_{jt}}{H_{jt}}\right) + e_{jt} \tag{10}$$

where $b_{t\text{-}j}$ are cohort dummies and d_t are time dummies. Supplies C_{jt} and H_{jt} for each age group j in year t are computed by:

$$s_{jt} = \varphi_{jt}^{s} + \sum_{k} \delta_{jtk}^{s} \hat{\tau}_{tk}^{s} \tag{11}$$

where s, $s \in \{C,H\}$, denotes the category that the calculation is carried out for, φ is the sum of the average weekly hours worked by male workers in the dataset, and δ contains the sum of the average weekly hours worked by the different education categories, k. When $s = H$, $k \in$ {less than high school, incomplete college}, for $s = C$, $k \in$ {higher than college}. Finally, $\hat{\tau}$ is the estimated wage gap between k-workers with respect to s-workers, and the coefficient is used as a weight for δ. The parameter $\hat{\tau}$ is estimated three times for each year, depending on k, and the explanatory variables considered in the model are a second-degree polynomial for age, a dummy variable for marital status, six economic activity dummies, five regional dummies, and educational level dummies according to k. Note that in Equation (11) $\hat{\tau}$ does not depend on age group j, however further studies could provide more insight into this issue. Equation (11) attempts to take into account differences in the effective supply of labor by different groups. Card and Lemieux (2001) and Ferreira (2004) adopt a similar procedure. The difference is that Ferreira (2004) is unable to identify college dropouts in order to determine the respective labor efficiency, and therefore they are considered college-educated workers.

In the second stage of the two-step estimation procedure, given an estimate of $1/\sigma_A$, the relative efficiency parameters α_{jt} and β_{jt} are computed since equations (4) and (5) are equivalent to estimating:

$$\ln\left(w_{jt}^{H}\right)+\frac{1}{\sigma_{A}}\ln H_{jt}=\ln\left(\theta_{ht}H_{t}^{\rho-\eta}\Psi_{t}\right)+\ln\left(\beta_{jt}\right) \tag{12}$$

$$\ln\left(w_{jt}^{C}\right)+\frac{1}{\sigma_{A}}\ln C_{jt}=\ln\left(\theta_{ht}C_{t}^{\rho-\eta}\Psi_{t}\right)+\ln\left(\alpha_{jt}\right) \tag{13}$$

The terms on the left are computed using a first-step estimate of $1/\sigma_A$, while the terms on the right are estimated using a set of time dummies and cohort dummies. Thus, given estimates of α_{jt}, β_{jt} and η, it is possible to construct estimates of aggregate supply of college-educated and high-school-educated labor in each year. According to Card and Lemieux (2001), since the sampling variances of the estimated r_{jt}'s are known, it is straightforward to construct goodness-of-fit tests for the null hypothesis of no specification error, depending on the included effects. Finally, the model could be estimated by weighted least squares, where the dependent variable is the college-high school wage gap for each age group. The weights are the inverse of the sampling variance of the estimated wage gaps. This procedure is used to account for sample-generated heteroscedasticity, and the estimates of the coefficient errors are corrected for the various problems inherent in the two-stage method.

5. Results

Table 1 presents the summary statistics for key variables for individuals from the data used in the analysis. In the eight periods, most of the variables listed are very similar. As expected, average schooling increased from 9.46 years in 2005 to 10.28 years in 2012; growth in schooling as an investment in human capital has been one of Mexico's most impressive achievements[3]. The most significant growth was recorded for the group of high-school-educated workers, which jumped from 35 to 41% during the period. On the other hand, the small share of college-educated workers grew by just 1 percentage point, from 22 to 23%. The standard deviation for high-school-educated male workers is around 0.5, while for college-graduates it is around 0.4. These large values for standard deviations indicate the possible existence of heteroscedasticity. In addition, we see incipient growth of the average real log hourly wage.

3. Despite the trend toward increased schooling, the possibility of entering the formal employment sector has declined (see, e.g., Rodríguez-Oreggia, 2009, Gutiérrez, 2009, and Islas and Cortez, 2011).

Table 1. Summary statistics, male workers 25-59 years old

Variable	2005		2006	
	Mean	Std. dev.	Mean	Std. dev.
Age	40.103	9.110	40.026	9.135
Years of schooling	9.463	4.875	9.518	4.864
Real log hourly wage	3.027	0.750	3.077	0.759
Weekly work hours	49.691	13.132	49.594	13.009
Primary school	0.839	0.368	0.840	0.367
Junior high school	0.607	0.489	0.618	0.486
High school	0.348	0.476	0.354	0.478
College	0.215	0.411	0.213	0.409
Postgraduate education	0.015	0.121	0.019	0.138
N	40,034		41,217	
	2007		2008	
Age	40.235	9.133	40.379	9.240
Years of schooling	9.743	4.799	9.662	4.803
Real log hourly wage	3.166	0.744	3.186	0.743
Weekly work hours	49.559	12.988	49.592	13.078
Primary school	0.858	0.349	0.854	0.353
Junior high school	0.639	0.480	0.638	0.481
High school	0.371	0.483	0.366	0.482
College	0.225	0.417	0.218	0.413
Postgraduate education	0.017	0.130	0.018	0.132
N	40,430		40,293	
	2009		2010	
Age	40.343	9.256	40.324	9.250
Years of schooling	9.846	4.702	10.063	4.634
Real log hourly wage	3.190	0.725	3.196	0.710
Weekly hours of work	49.202	13.057	49.212	12.978
Primary school	0.871	0.335	0.881	0.323
Junior high school	0.659	0.474	0.686	0.464
High school	0.375	0.484	0.394	0.489
College	0.221	0.415	0.228	0.419
Postgraduate education	0.019	0.138	0.019	0.136
N	38,138		38,416	

Table 1. (continued)

Variable	2011		2012	
	Mean	Std. dev.	Mean	Std. dev.
Age	40.397	9.259	40.579	9.307
Years of schooling	10.047	4.594	10.281	4.530
Real log hourly wage	3.205	0.703	3.250	0.697
Weekly work hours	49.390	13.080	49.556	13.009
Primary school	0.888	0.315	0.897	0.304
Junior high school	0.688	0.463	0.716	0.451
High school	0.391	0.488	0.409	0.492
College	0.224	0.417	0.234	0.424
Postgraduate education	0.018	0.132	0.020	0.139
N	38,897		37,363	

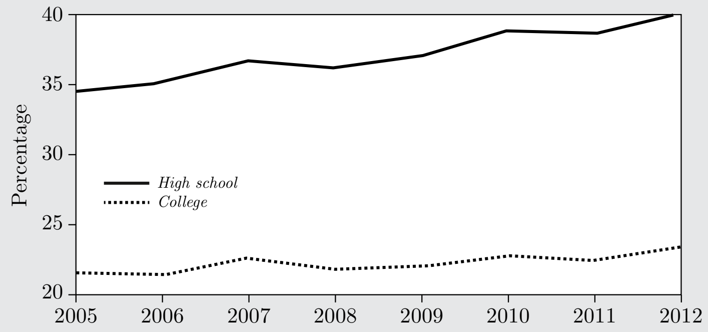
Source: ENOE, Q1 2005 to Q4 2012.
Note: In Mexico, primary school = 1^{st} through 6^{th} grades; junior high school = 7^{th} through 9^{th} grades; and high school = 10^{th} through 12^{th} grades.

Figure 1 provides the historical proportion of male workers with a college degree in the sample, capturing varying perspectives across individuals and settings. Although Mexico currently has a broad and diverse system of higher education that includes public and private institutions such as universities, technological institutes, technological universities, polytechnic universities, teachers' colleges, research centers, and specialized education centers, the percentage of workers with college degrees is less than 25.

Each column of Table 2 contains the estimated coefficients of Equation (8) for men aged 25-59. Following Card and Lemieux (2001) and Ferreira (2004), only men with a high school diploma and men with a college degree are considered. The random effect procedure is appropriate according to the Breusch-Pagan Lagrange multiplier test, suggesting its use instead of pooled regression. In addition, F- and Hausman test results favor random effects over fixed effects[4]. The models for each year include a second-degree polynomial for age, marital status, six economic activity

4. For all eight years the associated p-values from the Breusch-Pagan Lagrange multiplier test, distributed as $\chi^2_{(1)}$, are less than 0.01. In the case of the F-test, a p-value < 0.01 is obtained for all eight years. Finally, a p-value > 0.95 in all years is obtained by the Hausman test.

Figure 1. Proportion of workers with only a high school diploma and those with a college degree



Source: Table 1.

dummies, and five regional dummies[5]. The results are in agreement with previous studies (Hanson, 2003 or Varela, 2010); as is well known, the log mean wage difference between college and high school workers is around 0.5. Tables A1 and A2 in the appendix show alternative specifications that include, separately, controls for type of occupation and employment position. Table A3 includes both types of controls simultaneously. The wage-gap differences across different models are significant: about 0.4 and 0.5. In principle, all four specifications capture the strong effect of having a college degree on wages, so the explanatory variables shown in Table 1 will be those used in the subsequent models.

The estimated college premium remains almost constant over time; Table 3 presents the estimated college wage premiums for five-year age groups using Equation (9). All estimations are based on the mean log average real hourly wage between men with a college degree and those with only a high school diploma. They are estimated in separate regression models for each cohort in each year, including the age term and its quadratic, marital status, economic activity dummies, and regional dummies.

5. The economic activity dummies were selected based on the ENOE questionnaire, where a1 = Construction; a2 = Manufacturing; a3 = Trade; a4 = Services; a5 = Other activities and a6 = Agricultural. The regionalization corresponds to that used by the Presidency since 2002, where R1 = Baja California, Baja California Sur, Sonora, and Sinaloa; R2 = Chihuahua, Coahuila, Durango, Tamaulipas, and Nuevo León; R3 = Aguascalientes, Colima, Guanajuato, Jalisco, Michoacán, Nayarit, Querétaro, San Luis Potosí, and Zacatecas; R4 = Distrito Federal, Hidalgo, México, Morelos, Puebla, and Tlaxcala; and R5 = Campeche, Chiapas, Guerrero, Oaxaca, Quintana Roo, Tabasco, Veracruz, and Yucatán.

Table 2. Estimated college-high school wage differentials, males aged 25-59

	2005	2006	2007	2008	2009	2010	2011	2012
Age	0.022* (0.007)	0.028* (0.006)	0.021* (0.006)	0.027* (0.006)	0.023* (0.006)	0.024* (0.006)	0.027* (0.006)	0.033* (0.006)
Age^2	-1.7E-04* (8.7E-05)	-2.5E-04* (8.1E-05)	-1.8E-04* (7.8E-05)	-2.5E-04* (7.9E-05)	-2.1E-04* (7.8E-05)	-2.2E-04* (7.6E-05)	-2.5E-04* (7.4E-05)	-3.4E-04* (7.3E-05)
College	0.531* (0.014)	0.532* (0.014)	0.537* (0.014)	0.529 (0.014)	0.545* (0.014)	0.506* (0.014)	0.541* (0.014)	0.527* (0.014)
Married	0.116* (0.015)	0.128* (0.015)	0.112* (0.014)	0.095* (0.014)	0.086* (0.014)	0.077* (0.013)	0.080* (0.013)	0.086* (0.013)
a1	-0.180* (0.041)	-0.198* (0.039)	-0.166* (0.037)	-0.189* (0.042)	-0.173* (0.04)	-0.214* (0.042)	-0.244* (0.042)	-0.189* (0.037)
a2	-0.264* (0.040)	-0.245* (0.039)	-0.216* (0.035)	-0.291* (0.041)	-0.236* (0.039)	-0.265* (0.041)	-0.332* (0.041)	-0.268* (0.037)
a3	-0.329* (0.041)	-0.36* (0.039)	-0.305* (0.036)	-0.366* (0.041)	-0.329* (0.039)	-0.361* (0.042)	-0.432* (0.041)	-0.379* (0.037)
a4	-0.238* (0.039)	-0.228* (0.038)	-0.215* (0.034)	-0.255* (0.04)	-0.190* (0.038)	-0.220* (0.041)	-0.279* (0.040)	-0.224* (0.036)
a6	-0.756* (0.057)	-0.763* (0.053)	-0.778* (0.055)	-0.795* (0.053)	-0.765* (0.053)	-0.757* (0.053)	-0.818* (0.050)	-0.748* (0.051)
R1	0.295* (0.021)	0.282* (0.020)	0.329* (0.002)	0.320* (0.020)	0.329* (0.022)	0.288* (0.019)	0.287* (0.019)	0.259* (0.020)
R2	0.154* (0.020)	0.162* (0.019)	0.151* (0.020)	0.141* (0.020)	0.147* (0.021)	0.104* (0.019)	0.127* (0.018)	0.099* (0.019)
R3	0.108* (0.018)	0.163* (0.018)	0.154* (0.017)	0.127* (0.018)	0.149* (0.018)	0.109* (0.017)	0.134* (0.017)	0.113* (0.017)
R5	-0.037 (0.020)	-0.024 (0.019)	-0.016 (0.019)	-0.010 (0.019)	-0.017 (0.019)	0.006 (0.019)	-0.024 (0.018)	-0.025 (0.018)
N	20,468	21,081	21,082	20,083	19,173	20,019	19,785	19,575
R^2	0.232	0.240	0.239	0.231	0.244	0.224	0.250	0.227

Source: Own calculations using ENOE data from Q1 2005 to Q4 2012.
Note: Significant at 5%; robust standard errors in parentheses.

According to standard human capital theory, this differential increases steadily with age, but conventional models ignore differences in the age distribution of educational attainment. Unlike the results found in Table 2, the decomposition of age groups reveals the difference in wages between younger and older male workers. In general, the results support the standard statement of Mincer's (1974) equation: Older people have greater returns. Another important result is that by substituting age and its square for five-year age group dummies in a model such as in Table 2, it is possible to perform the Wald test. The results[6] indicate that the estimated coefficients of younger men (aged 25-29) in relation to older workers (aged 45-49, 50-54, and 55-59) are statistically different. In other words, it is necessary to dig deeper into the phenomenon by considering its relative supplies.

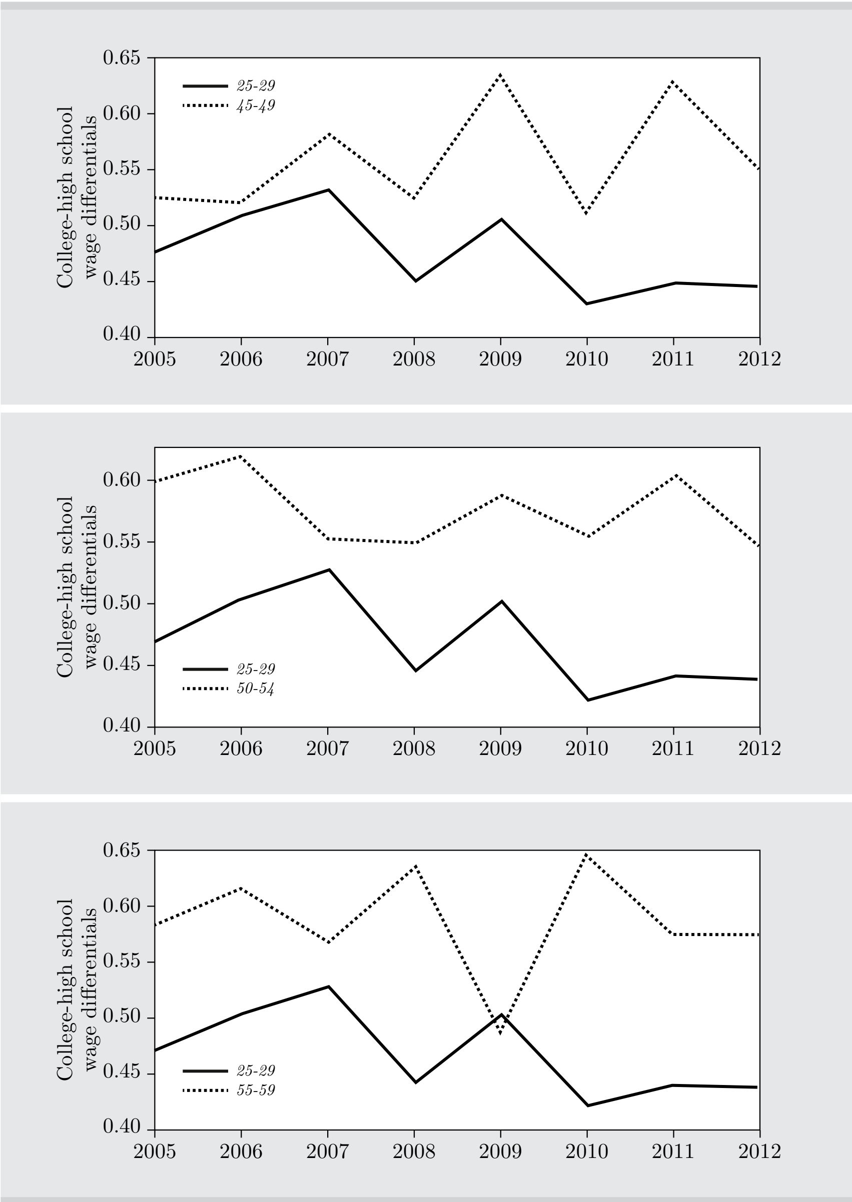
Figure 2 plots the college-high school wage gap for younger and older groups over the period analyzed. The overall patterns for the college premium for men aged 25-29 and 45-49 are very similar across years, but it appears that the evolution of the college premium for younger and older workers has tended to diverge since the financial and economic crisis of 2008. Once again, these results underscore the importance of studying differentials in the college premium among cohort groups. By examining the wage gaps for older workers aged 50-54 and 55-59, it is possible to infer that the difference has remained practically constant, with slight decreases toward the end of 2012. The employment stability earned by members of this age group may be a potential explanatory factor.

The more relevant change revealed by Figure 2 is the growing trend in the wage differential in the group of male workers aged 45-49. Conversely, the college premium among younger workers seems to have declined; it is important to highlight the fact that the financial crisis has had a particularly adverse effect on younger workers in Mexico (Villarreal, 2010). The importance of this lies in the fact that on the one hand, in the short term younger people have an incentive not to continue to higher levels of education. On the other hand, if young people decide to invest in higher levels of education, the main effect on the college premium will be observed in the long term.

To understand this change, we need to be aware not only of the gap between workers with the same education and different ages, but also of

6. The p-values in chi-square tests are $p < 0.01$ in all cases.

Figure 2. College-high school wage differentials by age group



Source: Table 3.

Table 3. Estimated college-high school wage differentials by age groups

Year/age	25-29	30-34	35-39	40-44	45-49	50-54	55-59
2005	0.470	0.530	0.534	0.528	0.521	0.599	0.584
	(0.032)	(0.032)	(0.032)	(0.034)	(0.038)	(0.047)	(0.074)
2006	0.504	0.469	0.598	0.508	0.517	0.619	0.613
	(0.030)	(0.031)	(0.030)	(0.031)	(0.039)	(0.048)	(0.070)
2007	0.528	0.518	0.531	0.515	0.579	0.552	0.568
	(0.030)	(0.032)	(0.030)	(0.032)	(0.035)	(0.046)	(0.061)
2008	0.444	0.526	0.541	0.566	0.521	0.549	0.634
	(0.031)	(0.033)	(0.033)	(0.032)	(0.036)	(0.044)	(0.062)
2009	0.501	0.574	0.491	0.528	0.633	0.588	0.488
	(0.034)	(0.029)	(0.032)	(0.035)	(0.037)	(0.044)	(0.065)
2010	0.422	0.501	0.517	0.502	0.506	0.554	0.645
	(0.030)	(0.031)	(0.031)	(0.034)	(0.034)	(0.044)	(0.067)
2011	0.441	0.500	0.545	0.516	0.627	0.602	0.575
	(0.029)	(0.030)	(0.032)	(0.033)	(0.035)	(0.044)	(0.059)
2012	0.438	0.506	0.561	0.594	0.546	0.545	0.573
	(0.030)	(0.031)	(0.031)	(0.033)	(0.039)	(0.042)	(0.056)

Source: Own calculations using ENOE data from Q1 2005 to Q4 2012.
Note: Robust standard errors in parentheses.

the composition of the relative supplies. Using Equation (11)[7], we can compute the average weekly hours worked by men aged 25-59 with any level of education; the results are contained in Table 4. According to the notation, C_{jt} corresponds to the number of hours per week worked by college graduates by age group j in year t including postgraduate workers, weighted by their wage gap. Likewise, H_{jt} is the total number of weekly hours worked by the labor force with incomplete college or less education, weighted by wage differentials.

The estimation of the relative supply of workers with a college degree shows an important growth trend in young workers, with an increase from 0.3 to 0.4 in less than a decade. This ratio may be explained because returns to education in Mexico are substantial and higher than those estimated for developed countries (see Psacharopoulos, *et al.* (1996), Psacharopoulos and Patrinos (2004), or Canton (2007) for details), whereas the natural laws of supply and demand would typically

7. The procedure is to compute Equation (11) estimating $\hat{\tau}_{tk}^{s}$ in separate regressions depending on s. A total of 24 panel data models are used to construct Table 4, taking into account differences in the effective supply of skilled and low-skilled workers. The results are not reported but are available from the author upon request.

Table 4. Relative college labor supply by age group

Year/age	25-29	30-34	35-39	40-44	45-49	50-54	55-59
				College			
2005	59,657	50,169	52,764	41,174	45,751	28,685	17,697
2006	60,193	55,440	55,785	50,308	47,343	32,394	17,109
2007	62,544	51,116	56,128	55,115	45,513	30,686	19,388
2008	62,478	48,750	49,749	52,419	49,444	33,686	21,281
2009	58,072	51,055	53,408	49,704	46,495	34,073	22,090
2010	60,201	49,328	50,085	51,993	48,848	34,541	20,516
2011	63,335	49,672	49,258	50,424	45,126	36,840	23,187
2012	64,801	50,613	49,830	49,923	44,150	38,609	22,775
				High school			
2005	186,224	169,608	167,428	155,711	126,039	103,986	77,399
2006	184,196	173,038	180,254	159,046	129,565	101,655	77,633
2007	177,967	169,256	169,789	163,969	132,831	107,536	78,591
2008	179,068	164,582	172,613	159,286	131,924	112,373	78,108
2009	173,832	152,572	166,466	153,568	131,646	109,908	78,071
2010	171,487	150,761	166,517	155,081	133,894	107,476	76,403
2011	176,408	158,454	175,127	162,847	138,741	113,022	82,016
2012	162,888	139,709	165,205	159,786	134,524	110,889	81,603
				Relative supply			
2005	0.320	0.296	0.315	0.264	0.363	0.276	0.229
2006	0.327	0.320	0.309	0.316	0.365	0.319	0.220
2007	0.351	0.302	0.331	0.336	0.343	0.285	0.247
2008	0.349	0.296	0.288	0.329	0.375	0.300	0.272
2009	0.334	0.335	0.321	0.324	0.353	0.310	0.283
2010	0.351	0.327	0.301	0.335	0.365	0.321	0.269
2011	0.359	0.313	0.281	0.310	0.325	0.326	0.283
2012	0.398	0.362	0.302	0.312	0.328	0.348	0.279

Source: Own calculations using ENOE data from Q1 2005 to Q4 2012.

yield a decrease in the college premium. Even more notably, the trend for the estimated supplies of older workers has a differentiated effect on the college premium. Individuals aged 45-49 recorded a decrease of around 4%, while conversely, individuals aged 50-54 and 55-59 increased their participation. In general, the results contained in tables 3 and 4 show a surprisingly mixed effect on wage differentials and relative supplies.

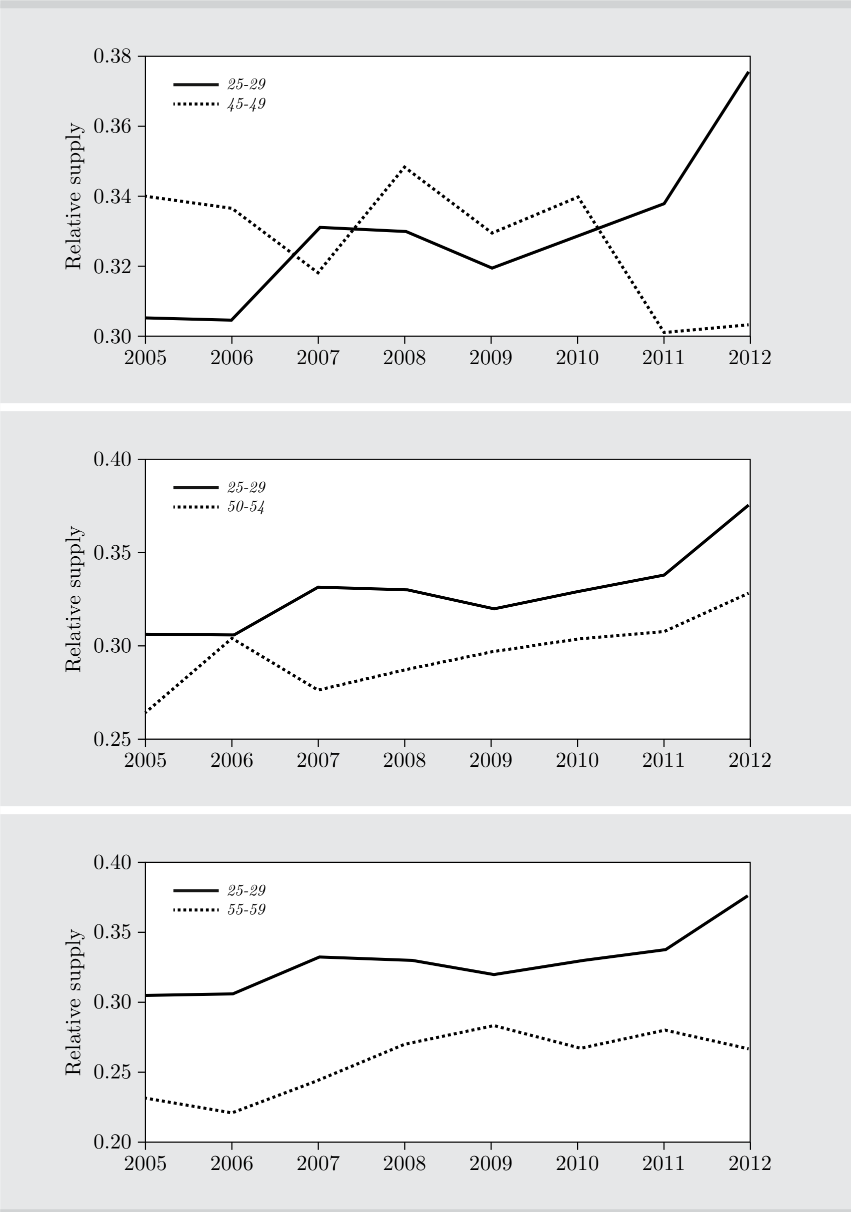
Figure 3 shows the behavior of relative supplies for each age group of interest. In all cases, with the exception of specific years for workers aged 45-49, the relative supply of the younger group is higher than the relative supply of older groups. One possible reason for this large difference is strongly suggested by the weighting parameter $\hat{\tau}$ in Equation (11). As previous work (Zepeda and Ghiara, 1999, Zamudio, 2001, Cañonero and Werner, 2002, and Hanson, 2003) has suggested, the recent increase in inequality in the income distribution is largely due to education. Thus, after trade liberalization occurred, the average relative wage of unskilled workers decreased and then, the effective supply may have been reduced for older cohorts.

Within this framework, the increase in the college-high school wage gap in workers aged 45-49 is attributable to steadily rising relative demand for college-educated labor, coupled with a dramatic slowdown in the rate of growth of the relative supply of college-educated workers. Card and Lemieux (2001) and Ferreira (2004) find similar behavior for groups of younger workers, but in the case of Mexico, younger workers are the segment of the population that is most affected, with unemployment rates twice as high as those of older adults. Furthermore, better-educated professionals have the highest unemployment rate (Villarreal, 2010).

The final step is to estimate two parameters of interest: the elasticity of substitution between the two education groups and the elasticity of substitution between different age groups. Using Equation (7) I establish the cohort born in 1956-1960 (aged 45-49) as the control group, hence the results presented correspond to differences with regard to older workers. The model presented in Table 5 suggests that a 1% increase in the relative supply of college labor causes a decrease of 0.33 percentage points in the college-high school wage differential, in the absence of non-neutral technology changes. Also, a 1% increase in the age-specific relative supply of college labor decreases the college premium by 0.6 percentage points, for that particular age group.

The empirical evidence provides an estimated elasticity of substitution between different age groups of around 1.7, while the parameter for the two education groups is about 3. We note that this specification does not seem to capture the annual wage gap (year effect) that would have, on average, increased in the absence of the age/cohort productivity factor and changes in supply according to educational level. In fact, the coefficients are not statistically different from 0 because the yearly dummies cannot capture technology shocks, which would require a long time-series sample and five-year interval dummies.

Figure 3. Relative college labor supply by age group



Source: Table 4.

Table 5 Estimated college-high school wage gap by cohort and year

Age-specific relative supply	-0.601*	(0.100)
Aggregate supply index	-0.332*	(0.121)
Age effect		
25-29	-0.121*	(0.012)
30-34	-0.090*	(0.011)
35-39	-0.063*	(0.009)
40-44	-0.050*	(0.007)
50-54	0.013	(0.008)
55-59	0.026*	(0.013)
Cohort effect		
1946-1950	-0.050*	(0.010)
1951-1955	-2E-04	(0.009)
1961-1965	0.015*	(0.006)
1966-1970	0.019*	(0.006)
1971-1975	0.024*	(0.007)
1976-1980	0.032*	(0.008)
1981-1985	0.029*	(0.009)
1986-1990	0.021	(0.012)
Year effect dummies		
2006	-0.010	(0.007)
2007	0.0010	(0.007)
2008	-0.010	(0.007)
2009	-0.003	(0.008)
2010	-0.003	(0.007)
2010	-0.024*	(0.008)
2012	-0.121*	(0.010)

Source: Own calculations using ENOE data from Q1 2005 to Q4 2012.
Note: Significant at 5%; robust standard errors in parentheses.

As a general rule, the elasticity of substitution captures the percentage change in relative demand for the two factors due to the change in relative factor prices at constant output. Also, the parameter of substitution between college and high school labor influences the impact of schooling on the education wage premium. In the traditional sense, the reduction is larger if the elasticity is low; i.e., they are far from perfect substitutes.

Similar studies for U.S, the U.K. and Canada (Card and Lemieux, 2001) and for Brazil (Ferreira, 2004) suggest that the elasticity of substitution between different age groups is 4.4. At the same time, the elasticity of substitution between college and high school labor types is about 2.5, and less than 2 in the case of Brazil. Nevertheless, these studies have analyzed long time-series data sets: 1959-1995 for the U.S., the U.K. and Canada, and 1976-1998 for Brazil. In marked contrast, this study of Mexico is focused on the short term, taking advantage of quarterly datasets. The results in Table 5 indicate that the two education groups are more easily substituted in Mexico than in other countries. In contrast, the substitution of workers between different age groups seems to be more complicated, based on the small estimated value for the elasticity of substitution.

This finding could be explained by the structural change occurring in Mexico's labor market. According to the literature review, empirical studies for Mexico have concluded that the wage gap between the two education groups has tended to widen with trade liberalization. The intuition behind previous studies is that employers reward skilled workers with an important college wage premium, making them hard to substitute. However, the combination in recent years of a high unemployment rate for skilled workers, slow economic growth, and the negative effects of the financial crisis could potentially explain this result.

6. Concluding remarks

This paper has introduced the first evidence in Mexico of the estimated evolution of the college wage gap by age groups, controlling for the relative supply of college graduate workers. Following the econometric methodology proposed by Card and Lemieux (2001) and using rotating panel data from 2005 to 2012, I present the partial elasticity of substitution between college and high school workers and across age groups. The results of the estimation indicate the existence of a large elasticity substitution (around 3) between male workers with different levels of education, which could mean that college- and high-school-educated workers are considered easily substitutable by employers. Furthermore, the small value of the estimated parameter (around 1.7) between the two education groups suggests that younger and older workers are viewed as different by employers and they are far from perfect substitutes. This result is very important because it suggests that in the current Mexican labor market it is easier to substitute skilled workers with unskilled workers rather than replace older workers with a younger labor force.

Overall, the model presents a negative and significant effect on the college premium for cohort variations in relative supply. This study highlights a decreasing trend in the college wage gap for younger workers (aged 25-29) combined with an increasing trend in the wage gap for older workers (aged 45-49). On the other hand, the behavior of the college premium for the oldest groups (aged 50-54 and 55-59) describes a constant trend in the wage gap.

The overall patterns for the college premium for men aged 25-29 and 45-49 are very similar across years but it appears that the evolution of the college premium for younger and older workers has tended to diverge after the financial and economic crisis of 2008. This underscores the need for wage gap studies in Mexico to consider the specific composition of workers by age group. According to this model, valid in a scenario of perfectly competitive equilibrium, the recent downward trend in the college premium for younger men depends mainly on the age effect. Although the results imply imperfect substitution between skilled and unskilled labor, it seems that in the case of Mexico, in contrast with the U.S., the U.K, Canada or Brazil, there is a small elasticity of substitution between the two age groups. An important implication of this finding is that, because of the aging of the population and increased levels of schooling, younger, educated workers are the segment of the population that is most affected.

The demographic and schooling transformation now underway in Mexico has the potential to both help and hinder its overall economic development agenda. Modifications of Mexico's Federal Labor Law enacted in 2012 brought important changes from an employer's perspective. One of the most interesting is the expansion of the types of employment relationships that are legally allowed. In addition to the already existing contracts for an indefinite term or a specific project, the reform introduced the seasonal employment category, which allows short-term employment to cover the need for additional workforce requirements during seasonal peaks, and the temporary employment contract, which permits short-term employment to cover immediate needs. In principle, the new recruitment scheme could have a positive impact on highly educated younger workers since seasonal or temporary employment could provide them with jobs and enable them to begin gaining experience quickly.

Finally, the results imply not only imperfect substitution between older and younger men but also between skilled and unskilled men. Future studies could include more than two categories of educated workers, including for example those with incomplete high school, incomplete college, and advanced degrees, as well as an extension for women in labor markets, controlling for self-selection.

REFERENCES

Airola, J. and C. Juhun (2008), "Wage inequality in post-reform Mexico," *Journal of Income Distribution* **17**(1): 110–34.

Ampudia, N. (2007), "Política monetaria, restricción crediticia y empleo," in Mantéy, M. and N. Levy, eds., *Políticas macroeconómicas para países en desarrollo.* Mexico: Miguel Ángel Porrúa-UNAM.

Becker, G. (1965), "A theory of the allocation of time," *The Economic Journal* **75**(299): 493–517.

Bia, M., R. Leombruni, and P. Messe (2009), "Young in-old out: A new evaluation based on generalized propensity score," *LABORatorio R. Revelli*, Working Paper No. 93.

Cañonero, G. and A. Werner (2002), "Salarios relativos y liberación del comercio en México," *El Trimestre Económico* **69**(273): 123–42.

Canton, E. (2007), "Social returns to education: Macro-evidence," *De Economist* **155**(4): 449–68.

Card, D. and T. Lemieux (2001), "Can falling supply explain the rising return to college for younger men? A cohort-based analysis," *Quarterly Journal of Economics* **116**(2): 705–46.

Chiquiar, D. (2008), "Globalization, regional wage differentials and the Stolper-Samuelson theorem: Evidence from Mexico," *Journal of International Economics* **74**(1): 70–93.

Cragg, M. and M. Epelbaum (1996), "Why has wage dispersion grown in Mexico? Is it the incidence of reforms or the growing demand for skills?" *Journal of Development Economics* **51**(1): 99–116.

Ferreira, S. (2004), "The provision of education and its impacts on college premium in Brazil," *Revista Brasileira de Economia* **58**(2): 211–33.

Gutiérrez, R. (2009). "La recesión y la medición de la tasa de desempleo de México ante una eventual reforma laboral," *Economía, Teoría y Práctica* **31**: 43–72.

Hairault, J., A. Chéron, and F. Langot (2007), "Job creation and job destruction over the life cycle: The older workers in the spotlight," IZA Discussion Paper No. 2597.

Hanson, G. (2003), "What has happened to wages in Mexico since NAFTA? Implications for hemispheric free trade," NBER Working Paper 9563.

Hanson, G. and A. Harrison (1995), "Trade, technology and wage inequality," NBER Working Paper 5110.

Hanson, G. and A. Harrison (1999), "Who gains from trade reform? Some remaining puzles," *Journal of Development Economics* **59**(1): 125–54.

Hernández, E. (2000), *Empleo y necesidades de capacitación en mercados regionales de trabajo en México.* Mexico: Secretaría del Trabajo y Previsión Social.

Huesca, L. (2004), "La rentabilidad de la escolaridad en los hogares asalariados en México durante 1984-2000," *Problemas del Desarrollo* **35**(138): 125–254.

Islas, A. and W. Cortez (2011), "How relevant is monetary policy to explain Mexican unemployment fluctuations?" MPRRA Paper No. 30027.

Kalwij, A., A. Kapteyn, and K. Vos (2010), "Retirement of older workers and employment of the young," *De Economist* **158**(4): 341–59.

Katz, L. and K. Murphy (1992), "Changes in relative wages, 1963-1987: Supply and demand factors," *Quarterly Journal of Economics* **107**(1): 35–78.

López, G. (2006), "Mexico: Two decades of the evolution of education and inequality," World Bank Policy Research Working Paper 3919.

Meza, L. (1999), "Cambios en la estructura salarial de México en el período 1988-1993 y el aumento en el rendimiento de la educación superior," *El Trimestre Económico* **66** (262): 189–226.

Mincer, J. (1974), *Schooling, experience and earnings.* New York: National Bureau of Economic Research.

Prskawetz, A., T. Tsachev, and V. Veliov (2012), "Optimal education in an age-structured model under changing labor demand and supply," *Macroeconomic Dynamics* **16**(2): 159–83.

Psacharopoulos, G. and H. Patrinos (2004), "Returns to investment in education: A further update," *Education Economics* **12**(2): 111–34.

Psacharopoulos, G., E. Velez, A. Panagides, and H. Yang (1996), "Returns to education during economic boom and recession: Mexico, 1984, 1989 and 1992," *Education Economics* **4**(3): 219–30.

Robertson, R. (2004), "Relative prices and wage inequality: Evidence from Mexico," *Journal of International Economics* **64**(2): 387–409.

Rodríguez-Oreggia, E. (2009), "La dinámica comparativa del sector informal en México: Temas actuales de política económica," in I. Soloaga, ed., *Sobre México.* Mexico: UPAEP.

Roger, M. and M. Wasmer (2009), "Heterogeneity matters: Labor productivity differentiated by age and skills," Paris School of Economics Working Paper No. 2009-51.

Rones, P. (1983), "Labor market problems of older workers," *Monthly Labor Review* May: 1–12.

Tello, C. and R. Ramos (2012), "Wage inequality and economic growth in Mexican regions," *Investigaciones Regionales* **24**: 115–52.

Varela, R., *et al.* (2010), "Determinantes de los ingresos salariales en México: Una perspectiva de capital humano," *Región y Sociedad* **22**(49): 117–42.

Villarreal, M. (2010), *The Mexican economy after the global crisis.* Mexico: UNT Libraries Government Documents Department.

Zamudio, A. (2001), "La escolaridad y la distribución condicional del ingreso: Una aplicación de regresión cuantil," *El Trimestre Económico* **68**(269): 39–70.

Zepeda, E. and R. Ghiara (1999), "Determinación del salario y capital humano en México: 1987–1993," *Economía, Sociedad y Territorio* **2**(5): 67–116.

APPENDIX

Table A1. Estimated college-high school wage differentials for men aged 25-59, specification 2

	2005	2006	2007	2008	2009	2010	2011	2012
Age	0.023* (0.006)	0.028* (0.006)	0.023* (0.006)	0.029* (0.006)	0.024* (0.006)	0.028* (0.006)	0.028* (0.006)	0.034* (0.006)
Age^2	-1.8E-04* (8.5E-05)	-2.5E-04* (7.6E-05)	-2E-04* (7.3E-05)	-2.8E-04* (7.5E-05)	-2.3E-04* (7.4E-05)	-2.6E-04* (7.1E-05)	-2.8E-04* (7E-05)	-3.6E-04* (7E-05)
College	0.376* (0.015)	0.399* (0.015)	0.409* (0.015)	0.393* (0.015)	0.400* (0.016)	0.365* (0.015)	0.402* (0.015)	0.387* (0.015)
Married	0.103* (0.014)	0.119* (0.014)	0.113* (0.013)	0.093* (0.013)	0.081* (0.013)	0.075* (0.012)	0.078* (0.012)	0.078* (0.012)
o2	0.226* (0.024)	0.185* (0.025)	0.210* (0.025)	0.187* (0.027)	0.218* (0.028)	0.237* (0.025)	0.202* (0.025)	0.282* (0.025)
o3	0.172* (0.027)	0.136* (0.027)	0.149* (0.026)	0.135* (0.029)	0.103* (0.032)	0.091* (0.027)	0.125* (0.031)	0.180* (0.031)
o4	-0.085* (0.017)	-0.073* (0.018)	-0.059* (0.016)	-0.094* (0.018)	-0.065* (0.018)	-0.093* (0.018)	-0.073* (0.016)	-0.065* (0.017)
o5	-0.191* (0.017)	-0.164* (0.017)	-0.140* (0.016)	-0.192* (0.017)	-0.188* (0.017)	-0.195* (0.017)	-0.176* (0.016)	-0.167* (0.016)
o6	-0.261* (0.021)	-0.282* (0.021)	-0.219* (0.019)	-0.253* (0.020)	-0.277* (0.021)	-0.297* (0.021)	-0.288* (0.020)	-0.271* (0.019)
o7	-0.305* (0.021)	-0.297* (0.022)	-0.277* (0.022)	-0.297* (0.022)	-0.319* (0.021)	-0.313* (0.022)	-0.318* (0.021)	-0.290* (0.021)
o8	-0.331* (0.024)	-0.301* (0.025)	-0.238* (0.023)	-0.326* (0.024)	-0.276* (0.024)	-0.281* (0.024)	-0.261* (0.023)	-0.290* (0.020)

Table A1. (continued)

	2005	2006	2007	2008	2009	2010	2011	2012
o9	-0.395*	-0.393*	-0.393*	-0.367*	-0.346*	-0.404*	-0.345*	-0.278*
	(0.028)	(0.029)	(0.027)	(0.027)	(0.027)	(0.027)	(0.026)	(0.025)
o10	-0.784*	-0.767*	-0.774*	-0.739*	-0.793*	-0.755*	-0.744*	-0.724*
	(0.044)	(0.043)	(0.047)	(0.039)	(0.042)	(0.038)	(0.035)	(0.040)
R1	0.285*	0.272*	0.316*	0.311*	0.313*	0.288*	0.273*	0.254*
	(0.02)	(0.019)	(0.019)	(0.019)	(0.021)	(0.018)	(0.019)	(0.019)
R2	0.156*	0.155*	0.150*	0.142*	0.143*	0.110	0.118*	0.103*
	(0.019)	(0.018)	(0.018)	(0.019)	(0.019)	(0.018)	(0.017)	(0.018)
R3	0.104*	0.153*	0.140*	0.122*	0.145*	0.108*	0.126*	0.117*
	(0.017)	(0.017)	(0.016)	(0.016)	(0.017)	(0.016)	(0.016)	(0.016)
R5	-0.039*	-0.026	-0.016	-0.007	-0.015	0.011	-0.018	-0.022
	(0.018)	(0.018)	(0.019)	(0.019)	(0.018)	(0.018)	(0.017)	(0.017)
N	20,468	21,081	21,082	20,083	19,173	20,019	19,785	19575
R^2	0.297	0.292	0.289	0.277	0.294	0.274	0.291	0.271

Source: Own calculations using ENOE data from Q1 2005 to Q4 2012.
Note: o1 = Professional, technical and craft workers; o2 = Education; o3 = Officers and directors; o4 = Clerks; o5 = Industrial craftsmen and assistants; o6 = Merchants; o7 = Transport operators; o8 = Personal service; o9 = Protection and surveillance; o10 = Farmworkers.
Significant at 5%; robust standard errors in parentheses.

Table A2. Estimated college-high school wage differentials for men aged 25-59, specification 3

	2005	2006	2007	2008	2009	2010	2011	2012
Age	0.021* (0.007)	0.028* (0.007)	0.023* (0.006)	0.030* (0.007)	0.028* (0.006)	0.027* (0.006)	0.025* (0.006)	0.034* (0.006)
Age^2	-1.6E-04 (8.8E-08)	-2.4E-04* (8.3E-05)	-1.9E-04* (8E-05)	-2.8E-04* (8.1E-05)	-2.5E-04* (8E-05)	-2.4E-04* (7.8E-05)	-2.3E-04* (7.7E-05)	-3.5E-04* (7.5E-05)
College	0.542* (0.014)	0.545* (0.014)	0.551* (0.014)	0.549* (0.014)	0.564* (0.014)	0.531* (0.014)	0.572* (0.014)	0.551* (0.014)
Married	0.113* (0.015)	0.121* (0.015)	0.111* (0.014)	0.091* (0.014)	0.081* (0.014)	0.074* (0.013)	0.078* (0.013)	0.074* (0.013)
p2	0.212* (0.021)	0.211* (0.021)	0.178* (0.021)	0.143* (0.021)	0.129* (0.022)	0.089* (0.022)	0.118* (0.023)	0.106* (0.022)
p3	-0.087* (0.018)	-0.111* (0.018)	-0.095* (0.017)	-0.089* (0.018)	-0.161* (0.018)	-0.124* (0.018)	-0.103* (0.018)	-0.119* (0.018)
R1	0.264* (0.021)	0.267* (0.021)	0.302* (0.020)	0.312* (0.021)	0.311* (0.022)	0.275* (0.020)	0.269* (0.020)	0.240* (0.020)
R2	0.148* (0.020)	0.161* (0.019)	0.147* (0.020)	0.148* (0.021)	0.149* (0.021)	0.108* (0.020)	0.128* (0.019)	0.107* (0.019)
R3	0.103* (0.019)	0.154* (0.018)	0.142* (0.018)	0.124* (0.018)	0.139* (0.018)	0.104* (0.018)	0.116* (0.017)	0.104* (0.017)
R5	-0.047* (0.020)	-0.029 (0.020)	-0.023 (0.020)	-3.03E-04 (0.020)	-0.016 (0.020)	0.005 (0.019)	-0.027 (0.019)	-0.028 (0.019)
N	20,468	21,081	21,082	20,083	19,173	20,019	19,785	19,575
R^2	0.215	0.22	0.217	0.196	0.213	0.189	0.206	0.189

Source: Own calculations using ENOE data from Q1 2005 to Q4 2012.
Note: p1 = Employees and subordinates; p2 = Employers; p3 = Self-employed.
Significant at 5%; robust standard errors in parentheses.

Table A3. Estimated college-high school wage differentials for men aged 25-59, specification 4

	2005	2006	2007	2008	2009	2010	2011	2012
Age	0.022*	0.026*	0.020*	0.028*	0.023*	0.027*	0.027*	0.032*
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
Age^2	-1.8E-04*	-2.3E-04*	-1.8E-04*	-2.7E-04*	-2.1E-04*	-2.6E-04*	-2.7E-04*	-3.3E-04*
	7.80E-08	7.50E-05	7.20E-05	7.30E-05	7.30E-05	7.00E-05	6.80E-05	6.80E-05
College	0.362*	0.375*	0.391*	0.379*	0.386*	0.354*	0.384*	0.372*
	(0.015)	(0.015)	(0.014)	(0.015)	(0.015)	(0.015)	(0.014)	(0.014)
Married	0.097*	0.114*	0.107*	0.089*	0.075*	0.073*	0.078*	0.075*
	(0.014)	(0.013)	(0.013)	(0.013)	(0.013)	(0.012)	(0.012)	(0.012)
Economic activity dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Position dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	20,468	21,081	21,082	20,083	19,173	20,019	19,785	19,575
R^2	0.321	0.318	0.312	0.299	0.319	0.294	0.313	0.294

Source: Own calculations using ENOE data from Q1 2005 to Q4 2012.
Note: Significant at 5%; robust standard errors in parentheses.

VOL. 51 NO. 1 (MAY, 2014), 147–178

LATIN AMERICAN JOURNAL OF ECONOMICS

A COHORT ANALYSIS OF THE COLLEGE PREMIUM IN MEXICO*

FRANCISCO BENITA**

This paper provides the first empirical evidence for Mexico about relative wage differences between college-educated and high-school-educated workers across five-year age groups. Rotating panel surveys are used to implement an imperfect substitution model for similar male workers between different age groups and between the two education groups. For the period 2005-2012, the results suggest a partial elasticity of substitution of 1.7 for college- and high-school-educated workers and a partial elasticity of substitution of about 3 across age groups. Remarkably, the wage gap between younger and older workers with the same education level increased after the economic crisis of 2008.

JEL classification: J21, J31, C33

Keywords: Wage differentials, cohort effects, rotating panel, elasticity of substitution, Mexico

1. Introduction

Human capital theory proposes that education improves labor force productivity by imparting useful knowledge and skills. Over the past 40 years, many studies have examined rising rates of returns to education. Most of these studies have shown that schooling is a key factor in explaining wage gaps among individuals. On the other hand, in developed and some developing countries, average schooling has been increasing steadily over time, resulting in an expanded rise in salaries and a substantial narrowing of the wage gap. A second phenomenon of interest is the aging of the population, mainly occurring in developed countries, and more recently in some developing countries. In these countries there will be a higher absolute number of elderly people, who will make up a larger share of the population and have longer healthy life expectancies, and there will be a relatively lower number of working-age people.

The Mexican case is particularly interesting in this context for several reasons. First, Mexico's labor structure has undergone various political,

* I would like to thank the editor and two anonymous reviewers for their valuable comments and suggestions.

** Department of Systems & Industrial Engineering, ITESM (Tec de Monterrey), Campus Monterrey, Av. Eugenio Garza Sada Sur 2401, Monterrey, NL, 64849, Mexico. Email: francisco_benita@hotmail.com.

doi 10.7764/LAJE.51.1.147

economic, and demographic changes, affecting both increases in schooling and growth in the elderly population over the past three decades. Second, wage inequality trends have been substantially different than those observed in other developing countries (Tello and Ramos, 2012). According to the Mexican Population and Housing Census, the average number of years of education in the labor force has strongly increased, from 3.4 years in 1970 to 7.5 years in 2000, and this growth trend has continued, reaching 8.8 years of education in 2010. In addition, the country has seen an increase in the average age, from 22 years in 2000 to 26.7 years in 2010. Like many parts of the world, Mexico has witnessed a growth in the proportion of its elderly population compared to all other age groups.

Both demographic (aging) and economic (schooling) effects imply relative changes in the labor supply, with possible strong variations in returns to education between age groups. Growth in the pool of young, highly educated workers has resulted in an increase in the country's labor supply. Likewise, it is possible that the gap in average earnings between less-educated and highly educated workers narrows if labor demand is unchanged. This long-term phenomenon raises several questions about the future development of wage and employment structures by age groups. The issue is relevant for different economic topics, such as social security actuarial calculations, increasing wage inequality, age-related government transfers, pensions, health care, and labor market reforms.

There is a large literature on education's effect on earnings in Mexico, focused on factors that largely explain the wage gap between highly educated and less-educated workers resulting from changes in labor demand (e.g., Cragg and Epelbaum, 1996, Hanson and Harrison, 1999, Hernández, 2000, Cañonero and Werner 2002, and Chiquiar, 2008, among others). However, there is no empirical evidence on this effect according to different age groups of workers. Hence, the objective of this paper is to test whether the increase over the last decade in the number of male college graduates has led to downward pressure on the college wage premium. Furthermore, it seeks to determine whether the increase in relative supplies could possibly lead to a narrowing of the wage gap between college-educated male workers and male workers who possess only a high school diploma.

Because female labor-market participation in developing countries is the result of an underlying choice process based on utility maximization (Becker, 1965) and the modeling requires appropriate statistical methods to control self-selection bias, this study is limited to male workers. Following the econometric methodology proposed by Card and Lemieux (2001) and using rotating panel data from 2005 to 2012, I present the partial elasticity of substitution between college and high school workers and between different age groups. Since empirical evidence suggests that younger workers cannot easily replace older workers (Rones, 1983, Hairault *et al.*, 2007, Bia *et al.*, 2009, or Kalwij *et al.*, 2010), I use an imperfect substitution model for similar male workers between different age groups. In addition, I estimate not only the elasticity of substitution between younger and older male workers but also the elasticity of substitution of workers with a college degree and those with only a high school diploma. In developing countries such as Mexico, estimation of the education wage premium by age groups and its elasticity is important for determining the allocation of public spending on educational policies, which should be oriented toward educational improvement and raising worker productivity.

The model represents a competitive equilibrium from the firm side, where entrepreneurs select the amount of labor required to maximize profits. The estimation results indicate that there is a large elasticity substitution (about 3) between male workers with different levels of education, which could be interpreted to mean that college- and high-school-educated workers are considered easily interchangeable by employers. Furthermore, the small value of the estimated parameter—around 1.7—between the two education groups suggests that young and old workers are considered different by employers and therefore they are far from perfect substitutes. This result is very important because it suggests that in the current Mexican labor market it is easier to substitute skilled workers with unskilled workers than to replace older workers with a younger labor force.

The remainder of this paper is organized as follows. Section 2 provides a review of the related literature, while Section 3 outlines the theoretical framework. Section 4 presents a characterization of the estimation strategy and the dataset description. The results and some discussion are presented in Section 5, while Section 6 contains concluding remarks.

2. Literature review

Changes in labor supply and demand have no specific effect on relative wages, particularly between skilled and unskilled workers. Katz and Murphy's (1992) seminal work on movements in the college wage premium from 1963 to 1987 in the United States concludes that the premium appears to be strongly related to fluctuations in the growth rate of the supply of college graduates. Understanding these relationships is very important for economic development since labor supply affects the environment in which the labor market functions.

In general, we note that the wage gap in Mexico has been widely discussed, especially during the 1980s and the first half of the 1990s. It is generally agreed that the wage gap between college-educated male workers and those with only a high school diploma is due to changes in labor demand, which are explained using two different approaches: trade liberalization and technological change.

In relation to trade liberalization, changes have occurred as a result of employers' strategies to reward the qualifications of their employees in order to increase their international competitiveness. In regard to technological change, reduction of trade barriers and the rise in foreign direct investment have led to increased use of technology, which in turn has driven demand for skilled workers to complement the technology. As a result, the relative productivity of skilled workers, mainly those with college and advanced degrees, has been increased by such technological changes.

Following the focus on trade liberalization, in the view of Zepeda and Ghiara (1999) the trade reforms that took place in Mexico between 1986 and 1990 were accompanied by important wage gaps between skilled and unskilled workers. Applying a quantile regression to estimate the conditional distribution of income, Zamudio (2001) finds that between 1984 and 1996 the distribution of income depends to a large extent on education. Cañonero and Werner (2002) study the widening of wage differentials for skilled and unskilled workers after Mexico's entry into the General Agreement on Tariffs and Trade (GATT) and find that the average relative wage of an unskilled worker had decreased by more than 20% by the end of 1990. In a similar context, Hanson (2003) focuses on policy reforms initiated during the 1990s, finding a significant rise in demand for skilled workers in the country and concluding that the policies reduced profits in industries that paid their workers high wages prior to the reform and raised the premium paid to workers in states along the U.S. border. Airola and Juhn (2008) argue that trade liberalization per se did not widen wage differentials in the region. Finally, in the view of Chiquiar (2008), in the period following implementation of the North

American Free Trade Agreement (NAFTA), the wage premium for high-skilled workers increased in what he calls the "second stage of globalization in Mexico."

Following the second approach, Hanson and Harrison (1995, 1999) suggest that the rising wage gap during the 1980s was associated with changes internal to industries and even internal to plants. Cragg and Epelbaum (1996) report significant changes in returns to specific occupations, in particular professionals and managers, but they believe that rising inequality in the early 1990s is largely due to trade liberalization reforms. In a similar study, Meza (1999) concludes that if the skills premium had not risen over the 1988-1993 period, the wage distribution would be improved. Finally, Hernández (2000) points out that technological changes associated with trade liberalization have promoted an incipient territorial decentralization focused on the U.S. border and the periphery of Mexico City.

Among more recent studies, Huesca (2004) finds that workers with high school and college education show higher returns to schooling because of an upward trend in demand for skilled workers. Lopez (2006) notes that the Mexican population is becoming more educated. The author shows how years of schooling have increasing over the 1960-1990 period, growing from 2.76 years in 1960 to 6.72 years in 1990, a difference of nearly four years. According to Ampudia (2007), trade liberalization and international markets increased demand for skilled labor, which in turn generated higher education wage premiums. For workers with a college degree, the education premium has been shown to be cyclical in times of recession.

Although there is extensive literature documenting the role of educational level on the wage gap in Mexico, most previous studies have focused on average returns to schooling rather than differences by age or cohort. While the rise in the average wage gap between college and high school workers has been briefly documented, the fact that the increases have varied among different age groups is not as well understood.

These studies analyze the evolution of returns to schooling under the assumption that different age groups with the same level of education are perfect substitutes in production processes. This assumption implies that the aggregate supply of each type of education can be obtained by simply adding the total number of workers in each education category, such that all education-related wage differentials in the labor market in any given year are proportional to the average college-high school wage gap that year.

An analysis of the evolution of wage differentials by age groups was presented for the first time in Card and Lemieux (2001). The authors extend the model proposed by Katz and Murphy (1992) in order to allow imperfect substitution between workers of different ages for the U.S., the United Kingdom, and Canada. In the study, the authors point out that shifts in the college-high school wage gap for younger men in the U.S. reflect changes in the relative supply of highly educated workers across age groups. In their estimation of the elasticity of substitution of age groups with the same education level, they find elasticity close to 5 between different age groups in the U.S., while the elasticity of substitution between college graduates and high school graduates ranges from 1.1 to 2.5, with similar results for both the U.K. and Canada.

Roger and Wasmer (2009) describe the importance of controlling for age and skill heterogeneity to explain labor productivity. Within this framework, some recent studies have continued the line of imperfect substitution of workers between different age groups. Ferreira (2004) tests the existence of a causal relationship between the evolution of the college-educated labor supply and the behavior of the college premium in Brazil by decomposing the relative supply of college-educated workers into two components: an age-specific component and a cohort-specific component. The author's main results suggest that the parameter for elasticity of substitution between different age groups is similar to the U.S., while the elasticity of substitution between two education groups is less than 2. Some implications of this result are outlined in Prskawetz *et al.* (2012), who introduce a model of optimal education policy at the macro level allowing for heterogeneity of the workforce with respect to its age and skills. Their estimations suggest that the relationship between the elasticity of substitution of labor across age groups plays a crucial role in the way demographic changes affect the optimal educational policy.

3. Background

This work follows the two-step estimation method introduced by Card and Lemieux (2001). The theoretical model has a production function that uses labor as the sole input. In this simple model, the input can be skilled (college graduates) or unskilled (high school graduates) combined under a CES technology. The aggregate output depends on two CES sub-aggregates of high school and college workers:

$$H_t = \left[\sum_j \left(\beta_{jt} H_{jt}^{\eta} \right) \right]^{1/\eta}, \tag{1}$$

$$C_t = \left[\sum_j \left(\alpha_{jt} C_{jt}^{\eta} \right) \right]^{1/\eta}, \tag{2}$$

where H_t and C_t are the total supply of high school and college labor in period t, and $\eta \in (-\infty,1]$ is a function of the partial elasticity of substitution, σ_A, between different age groups j with the same level of education ($\eta = 1 - 1/\sigma_A$). The α_j and β_j coefficients are relative efficiency parameters that vary over time. Perfect substitutability across age groups necessarily implies that $\eta = 1$. Therefore, total college or high school labor input is simply a weighted sum of the quantity of labor supplied by each group. In this work, the total supply of each type of work in each year is measured as the sum of the average weekly hours worked by members of the different education categories.

The model assumes an aggregate CES production function $y_t = (\theta_{ct} C_t^{\rho} + \theta_{ht} H_t^{\rho})^{1/\rho}$, where y_t is the total output in period t, and $\rho \in (-\infty,1]$ is a function of the elasticity of substitution σ_{E} between the two education groups ($\rho = 1 - 1/\sigma_E$). θ_{ct} and θ_{ht} are college and high-school productivity terms.

Under this specification, the marginal product of labor for a given age-education group depends on both the group's own supply of labor and the aggregate supply of labor in its education category. Formally, we have the following:

$$\frac{\partial \mathrm{y_t}}{\partial \mathrm{H}_{jt}} = \frac{\partial \mathrm{y_t}}{\partial \mathrm{H}_t} \frac{\partial \mathrm{H}_t}{\partial \mathrm{H}_{jt}} = \theta_{ht} H_t^{\rho-1} \Psi_t \beta_{jt} H_{jt}^{\eta-1} H_t^{1-\eta} = \theta_{ht} H_t^{\rho-\eta} \Psi_t \beta_{jt} H_{jt}^{\eta-1} \tag{3}$$

where $\Psi_t = (\theta_{ct} C_t^{\rho} + \theta_{ht} H_t^{\rho})^{1/\rho - 1}$. Similarly, the marginal product of college workers in age group j is $\partial y_t / \partial C_{jt} = \theta_{ct} C_t^{\rho-\eta} \Psi_t \alpha_{jt} C_{jt}^{\eta-1}$. Additionally, efficient utilization of different skill groups requires that relative wages are equal to relative marginal products:

$$\ln\left(w_{jt}^{H}\right)=\ln\left(\frac{\partial \mathrm{y_t}}{\partial \mathrm{H}_{jt}}\right)=\ln\left(\theta_{ht}H_t^{\rho-\eta}\Psi_t\right)+\ln\left(\beta_{jt}\right)-\frac{1}{\sigma_A}\ln\left(H_{jt}\right) \quad (4)$$

$$\ln\left(w_{jt}^{C}\right)=\ln\left(\frac{\partial \mathrm{y_t}}{\partial C_{jt}}\right)=\ln\left(\theta_{ct}C_t^{\rho-\eta}\Psi_t\right)+\ln\left(\alpha_{jt}\right)-\frac{1}{\sigma_A}\ln\left(C_{jt}\right) \quad (5)$$

From equations (4) and (5) it is possible to infer that the ratio of the wage rate of skilled workers in age group j, w_{jt}^C, to the wage rate of unskilled workers in the same age group, w_{jt}^H, satisfies:

$$\ln\left(\frac{w_{jt}^C}{w_{jt}^H}\right)=\ln\left(\frac{\theta_{ct}}{\theta_{ht}}\right)+\ln\left(\frac{\alpha_{jt}}{\beta_{jt}}\right)+\left(\frac{1}{\sigma_A}-\frac{1}{\sigma_E}\right)\ln\left(\frac{C_t}{H_t}\right)-\frac{1}{\sigma_A}\ln\left(\frac{C_{jt}}{H_{jt}}\right) \quad (6)$$

If relative employment ratios are taken as exogenous[1], Equation (6) leads to a simple model for the observed college-high school wage gap in age group j in period t:

$$\begin{aligned} r_{jt} \equiv \ln\left(\frac{w_{jt}^C}{w_{jt}^H}\right) &= \ln\left(\frac{\theta_{ct}}{\theta_{ht}}\right)+\ln\left(\frac{\alpha_{jt}}{\beta_{jt}}\right)+\left(\frac{1}{\sigma_A}-\frac{1}{\sigma_E}\right)\ln\left(\frac{C_t}{H_t}\right) \\ &\quad -\frac{1}{\sigma_A}\ln\left(\frac{C_{jt}}{H_{jt}}\right)+e_{jt} \end{aligned} \quad (7)$$

where e_{jt} is the sampling variation in the measured gap and/or any other sources of variation in age-specific wage premiums.

The model represents a competitive equilibrium from the firm side, where entrepreneurs choose the amount of labor required to maximize profits. The equilibrium wage for the perfectly competitive market industry is computed by allowing imperfect substitution between workers of different ages.

As a result, the model specification represents a firm's production equilibrium. General equilibrium analysis deals explicitly with the

1. Total supply of high school and college labor in each period can be varied exogenously as a part of a scenario, e.g., simulating the effects of increasing participation of certain population groups. However, they do not respond endogenously to variations in other parameters such as the productivity or wage rate of college workers in each age group. In the case of Mexico, due to the effects of GATT and NAFTA on the economy, this can be assumed since input prices and technology can be more closely linked to tariff changes (Robertson, 2004).

interdependence of households and firms as mediated by markets. With only one production factor it is particularly easy to illustrate production equilibrium. However, the model could be extended by adding other inputs such as capital. Also, within the problem statement it must be clear that the results are valid only in a perfectly competitive equilibrium scenario. Therefore, the model is not applicable to imperfect markets with monopoly or oligopoly labor demand (e.g., monopsony or oligopsony). Finally, Equation (7) becomes the equation to estimate. Note that the assumption of employment ratios as exogenous becomes essential for the results; otherwise, the weighted least squares of elasticity of substitution will have a positive basis.

4. Estimation

The empirical analysis is based on data extracted from the National Survey of Occupation and Employment (ENOE). These datasets are collected by the National Institute of Statistics and Geography of Mexico (INEGI) and are the basis for official employment statistics. The datasets are based on household surveys conducted on a quarterly basis, the main objective of which are to obtain information about demographic and economic characteristics of employment. The surveys are designed as panel data with overlapping blocks of observations that are renewed after being followed up over five consecutive quarters. This data is representative of the whole country, each of the 32 states, and urban and rural areas.

I analyze datasets from the first quarter of 2005 through the fourth quarter of 2012. These rotating panel data are used to derive labor market conditions for male workers ages 25-59 with either a high school diploma or a college degree. In addition, an important difference from previous studies is that I consider full-time workers (those working 48 hours or more per week) as well as part-time workers (those working between 48 and 96 hours per week).

For reasons related to analysis and interpretation, the strategy consists of estimating panels by following the individuals from the first through the fourth quarter of each year. This technique allows the estimation of robust coefficients, increasing the degrees of freedom and reducing collinearity among the explanatory variables. In general, panel data sets possess major advantages over conventional cross-sectional data sets. The data set in question consists of eight panels, each one made up of individuals who are interviewed in each of the four quarters of the year. The top and bottom 1% of the log real hourly wage, using 2010 as the base year, were dropped in order to avoid atypical values.

For a general overview of the effect of a college education on wages, I first estimate the college wage premium for each year t with no distinctions between age groups, formally:

$$\ln\left(w_{iq}\right) = \zeta_0 + X_{iq}^{'}\zeta + v_{iq} \tag{8}$$

where w_{iq} contains the hourly wage of individual i in quarter q, $q \in \{1,2,3,4\}$. ζ denotes the coefficient vector, matrix X contains the regressors used to explain the dependent variable, and v is a residual assumed to be $N(0,\sigma^2)$. Because schooling cannot change in each quarter, $v_{iq} = u_i + \mu_{iq}$ and the model becomes a generalized least squares[2]. The regressor matrix X contains information about age and its quadratic, a married category dummy variable, six economic activity dummies, five regional dummies, and a dummy variable for college degree. The proposed model specification is based on previous studies of Mexico described in Section 2. Although Equation (8) is not directly related to the estimation of Equation (7), it is a useful first approach to understanding the college premium. The assumption behind Equation (8) is that different age groups with the same level of education (college or high school) are perfect substitutes in production.

However, in order to capture the college wage premium within age groups, for each year t I estimate separate regressions for each five-year age group, j:

$$\ln\left(w_{jq}\right) = \gamma_0 + X_{jq}^{'}\gamma + \varepsilon_{jq} \tag{9}$$

In Equation (9), the college degree dummy variable contained in matrix X measures the college wage premium by five-year age groups. Secondly, we compute Equation (7) using the estimated college premium from Equation (9), but there is a problem because the aggregate supplies of the two education groups depend on the elasticity of substitution across

2. The critical assumption when using random effects is that regressors are not endogenous. When the assumption of zero conditional mean of the error term given the regressors is not satisfied, the estimators will be inconsistent. This can be corrected by using instrumental variables estimation to find a variable correlated with wages but not with the error term. Further work is required to test endogeneity and instrumental variables regression should be used if the problem is detected.

age groups. Card and Lemieux (2001) suggest a two-step estimation procedure to identify both σ_A and σ_E.

The first stage of the two-step estimation consists of estimating σ_A from a regression of age-group-specific college wage gaps on age-group-specific relative supplies of college educated labor, cohort effects, and time effects:

$$r_{jt} = b_{t-j} + d_t - \frac{1}{\sigma_A} \ln\left(\frac{C_{jt}}{H_{jt}}\right) + e_{jt} \tag{10}$$

where $b_{t\text{-}j}$ are cohort dummies and d_t are time dummies. Supplies C_{jt} and H_{jt} for each age group j in year t are computed by:

$$s_{jt} = \varphi^s_{jt} + \sum_k \delta^s_{jtk} \hat{\tau}^s_{tk} \tag{11}$$

where s, $s \in \{C,H\}$, denotes the category that the calculation is carried out for, φ is the sum of the average weekly hours worked by male workers in the dataset, and δ contains the sum of the average weekly hours worked by the different education categories, k. When $s = H$, $k \in$ {less than high school, incomplete college}, for $s = C$, $k \in$ {higher than college}. Finally, $\hat{\tau}$ is the estimated wage gap between k-workers with respect to s-workers, and the coefficient is used as a weight for δ. The parameter $\hat{\tau}$ is estimated three times for each year, depending on k, and the explanatory variables considered in the model are a second-degree polynomial for age, a dummy variable for marital status, six economic activity dummies, five regional dummies, and educational level dummies according to k. Note that in Equation (11) $\hat{\tau}$ does not depend on age group j, however further studies could provide more insight into this issue. Equation (11) attempts to take into account differences in the effective supply of labor by different groups. Card and Lemieux (2001) and Ferreira (2004) adopt a similar procedure. The difference is that Ferreira (2004) is unable to identify college dropouts in order to determine the respective labor efficiency, and therefore they are considered college-educated workers.

In the second stage of the two-step estimation procedure, given an estimate of $1/\sigma_A$, the relative efficiency parameters α_{jt} and β_{jt} are computed since equations (4) and (5) are equivalent to estimating:

$$\ln\left(w_{jt}^{H}\right)+\frac{1}{\sigma_A}\ln H_{jt}=\ln\left(\theta_{ht}H_t^{\rho-\eta}\Psi_t\right)+\ln\left(\beta_{jt}\right) \tag{12}$$

$$\ln\left(w_{jt}^{C}\right)+\frac{1}{\sigma_A}\ln C_{jt}=\ln\left(\theta_{ht}C_t^{\rho-\eta}\Psi_t\right)+\ln\left(\alpha_{jt}\right) \tag{13}$$

The terms on the left are computed using a first-step estimate of $1/\sigma_A$, while the terms on the right are estimated using a set of time dummies and cohort dummies. Thus, given estimates of α_{jt}, β_{jt} and η, it is possible to construct estimates of aggregate supply of college-educated and high-school-educated labor in each year. According to Card and Lemieux (2001), since the sampling variances of the estimated r_{jt}'s are known, it is straightforward to construct goodness-of-fit tests for the null hypothesis of no specification error, depending on the included effects. Finally, the model could be estimated by weighted least squares, where the dependent variable is the college-high school wage gap for each age group. The weights are the inverse of the sampling variance of the estimated wage gaps. This procedure is used to account for sample-generated heteroscedasticity, and the estimates of the coefficient errors are corrected for the various problems inherent in the two-stage method.

5. Results

Table 1 presents the summary statistics for key variables for individuals from the data used in the analysis. In the eight periods, most of the variables listed are very similar. As expected, average schooling increased from 9.46 years in 2005 to 10.28 years in 2012; growth in schooling as an investment in human capital has been one of Mexico's most impressive achievements[3]. The most significant growth was recorded for the group of high-school-educated workers, which jumped from 35 to 41% during the period. On the other hand, the small share of college-educated workers grew by just 1 percentage point, from 22 to 23%. The standard deviation for high-school-educated male workers is around 0.5, while for college-graduates it is around 0.4. These large values for standard deviations indicate the possible existence of heteroscedasticity. In addition, we see incipient growth of the average real log hourly wage.

3. Despite the trend toward increased schooling, the possibility of entering the formal employment sector has declined (see, e.g., Rodríguez-Oreggia, 2009, Gutiérrez, 2009, and Islas and Cortez, 2011).

Table 1. Summary statistics, male workers 25-59 years old

Variable	2005		2006	
	Mean	Std. dev.	Mean	Std. dev.
Age	40.103	9.110	40.026	9.135
Years of schooling	9.463	4.875	9.518	4.864
Real log hourly wage	3.027	0.750	3.077	0.759
Weekly work hours	49.691	13.132	49.594	13.009
Primary school	0.839	0.368	0.840	0.367
Junior high school	0.607	0.489	0.618	0.486
High school	0.348	0.476	0.354	0.478
College	0.215	0.411	0.213	0.409
Postgraduate education	0.015	0.121	0.019	0.138
N	40,034		41,217	
	2007		**2008**	
Age	40.235	9.133	40.379	9.240
Years of schooling	9.743	4.799	9.662	4.803
Real log hourly wage	3.166	0.744	3.186	0.743
Weekly work hours	49.559	12.988	49.592	13.078
Primary school	0.858	0.349	0.854	0.353
Junior high school	0.639	0.480	0.638	0.481
High school	0.371	0.483	0.366	0.482
College	0.225	0.417	0.218	0.413
Postgraduate education	0.017	0.130	0.018	0.132
N	40,430		40,293	
	2009		**2010**	
Age	40.343	9.256	40.324	9.250
Years of schooling	9.846	4.702	10.063	4.634
Real log hourly wage	3.190	0.725	3.196	0.710
Weekly hours of work	49.202	13.057	49.212	12.978
Primary school	0.871	0.335	0.881	0.323
Junior high school	0.659	0.474	0.686	0.464
High school	0.375	0.484	0.394	0.489
College	0.221	0.415	0.228	0.419
Postgraduate education	0.019	0.138	0.019	0.136
N	38,138		38,416	

Table 1. (continued)

Variable	2011		2012	
	Mean	Std. dev.	Mean	Std. dev.
Age	40.397	9.259	40.579	9.307
Years of schooling	10.047	4.594	10.281	4.530
Real log hourly wage	3.205	0.703	3.250	0.697
Weekly work hours	49.390	13.080	49.556	13.009
Primary school	0.888	0.315	0.897	0.304
Junior high school	0.688	0.463	0.716	0.451
High school	0.391	0.488	0.409	0.492
College	0.224	0.417	0.234	0.424
Postgraduate education	0.018	0.132	0.020	0.139
N	38,897		37,363	

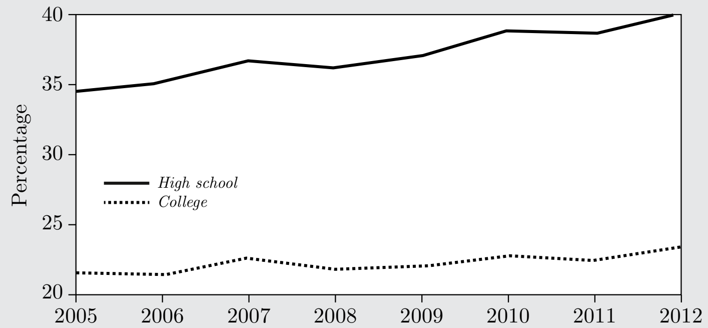
Source: ENOE, Q1 2005 to Q4 2012.
Note: In Mexico, primary school = 1st through 6th grades; junior high school = 7th through 9th grades; and high school = 10th through 12th grades.

Figure 1 provides the historical proportion of male workers with a college degree in the sample, capturing varying perspectives across individuals and settings. Although Mexico currently has a broad and diverse system of higher education that includes public and private institutions such as universities, technological institutes, technological universities, polytechnic universities, teachers' colleges, research centers, and specialized education centers, the percentage of workers with college degrees is less than 25.

Each column of Table 2 contains the estimated coefficients of Equation (8) for men aged 25-59. Following Card and Lemieux (2001) and Ferreira (2004), only men with a high school diploma and men with a college degree are considered. The random effect procedure is appropriate according to the Breusch-Pagan Lagrange multiplier test, suggesting its use instead of pooled regression. In addition, F- and Hausman test results favor random effects over fixed effects[4]. The models for each year include a second-degree polynomial for age, marital status, six economic activity

4. For all eight years the associated p-values from the Breusch-Pagan Lagrange multiplier test, distributed as $\chi^2_{(1)}$, are less than 0.01. In the case of the F-test, a p-value < 0.01 is obtained for all eight years. Finally, a p-value > 0.95 in all years is obtained by the Hausman test.

Figure 1. Proportion of workers with only a high school diploma and those with a college degree



Source: Table 1.

dummies, and five regional dummies[5]. The results are in agreement with previous studies (Hanson, 2003 or Varela, 2010); as is well known, the log mean wage difference between college and high school workers is around 0.5. Tables A1 and A2 in the appendix show alternative specifications that include, separately, controls for type of occupation and employment position. Table A3 includes both types of controls simultaneously. The wage-gap differences across different models are significant: about 0.4 and 0.5. In principle, all four specifications capture the strong effect of having a college degree on wages, so the explanatory variables shown in Table 1 will be those used in the subsequent models.

The estimated college premium remains almost constant over time; Table 3 presents the estimated college wage premiums for five-year age groups using Equation (9). All estimations are based on the mean log average real hourly wage between men with a college degree and those with only a high school diploma. They are estimated in separate regression models for each cohort in each year, including the age term and its quadratic, marital status, economic activity dummies, and regional dummies.

5. The economic activity dummies were selected based on the ENOE questionnaire, where a1 = Construction; a2 = Manufacturing; a3 = Trade; a4 = Services; a5 = Other activities and a6 = Agricultural. The regionalization corresponds to that used by the Presidency since 2002, where R1 = Baja California, Baja California Sur, Sonora, and Sinaloa; R2 = Chihuahua, Coahuila, Durango, Tamaulipas, and Nuevo León; R3 = Aguascalientes, Colima, Guanajuato, Jalisco, Michoacán, Nayarit, Querétaro, San Luis Potosí, and Zacatecas; R4 = Distrito Federal, Hidalgo, México, Morelos, Puebla, and Tlaxcala; and R5 = Campeche, Chiapas, Guerrero, Oaxaca, Quintana Roo, Tabasco, Veracruz, and Yucatán.

Table 2. Estimated college-high school wage differentials, males aged 25-59

	2005	2006	2007	2008	2009	2010	2011	2012
Age	0.022*	0.028*	0.021*	0.027*	0.023*	0.024*	0.027*	0.033*
	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Age^2	-1.7E-04*	-2.5E-04*	-1.8E-04*	-2.5E-04*	-2.1E-04*	-2.2E-04*	-2.5E-04*	-3.4E-04*
	(8.7E-05)	(8.1E-05)	(7.8E-05)	(7.9E-05)	(7.8E-05)	(7.6E-05)	(7.4E-05)	(7.3E-05)
College	0.531*	0.532*	0.537*	0.529	0.545*	0.506*	0.541*	0.527*
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Married	0.116*	0.128*	0.112*	0.095*	0.086*	0.077*	0.080*	0.086*
	(0.015)	(0.015)	(0.014)	(0.014)	(0.014)	(0.013)	(0.013)	(0.013)
a1	-0.180*	-0.198*	-0.166*	-0.189*	-0.173*	-0.214*	-0.244*	-0.189*
	(0.041)	(0.039)	(0.037)	(0.042)	(0.04)	(0.042)	(0.042)	(0.037)
a2	-0.264*	-0.245*	-0.216*	-0.291*	-0.236*	-0.265*	-0.332*	-0.268*
	(0.040)	(0.039)	(0.035)	(0.041)	(0.039)	(0.041)	(0.041)	(0.037)
a3	-0.329*	-0.36*	-0.305*	-0.366*	-0.329*	-0.361*	-0.432*	-0.379*
	(0.041)	(0.039)	(0.036)	(0.041)	(0.039)	(0.042)	(0.041)	(0.037)
a4	-0.238*	-0.228*	-0.215*	-0.255*	-0.190*	-0.220*	-0.279*	-0.224*
	(0.039)	(0.038)	(0.034)	(0.04)	(0.038)	(0.041)	(0.040)	(0.036)
a6	-0.756*	-0.763*	-0.778*	-0.795*	-0.765*	-0.757*	-0.818*	-0.748*
	(0.057)	(0.053)	(0.055)	(0.053)	(0.053)	(0.053)	(0.050)	(0.051)
R1	0.295*	0.282*	0.329*	0.320*	0.329*	0.288*	0.287*	0.259*
	(0.021)	(0.020)	(0.002)	(0.020)	(0.022)	(0.019)	(0.019)	(0.020)
R2	0.154*	0.162*	0.151*	0.141*	0.147*	0.104*	0.127*	0.099*
	(0.020)	(0.019)	(0.020)	(0.020)	(0.021)	(0.019)	(0.018)	(0.019)
R3	0.108*	0.163*	0.154*	0.127*	0.149*	0.109*	0.134*	0.113*
	(0.018)	(0.018)	(0.017)	(0.018)	(0.018)	(0.017)	(0.017)	(0.017)
R5	-0.037	-0.024	-0.016	-0.010	-0.017	0.006	-0.024	-0.025
	(0.020)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.018)	(0.018)
N	20,468	21,081	21,082	20,083	19,173	20,019	19,785	19,575
R^2	0.232	0.240	0.239	0.231	0.244	0.224	0.250	0.227

Source: Own calculations using ENOE data from Q1 2005 to Q4 2012.
Note: Significant at 5%; robust standard errors in parentheses.

According to standard human capital theory, this differential increases steadily with age, but conventional models ignore differences in the age distribution of educational attainment. Unlike the results found in Table 2, the decomposition of age groups reveals the difference in wages between younger and older male workers. In general, the results support the standard statement of Mincer's (1974) equation: Older people have greater returns. Another important result is that by substituting age and its square for five-year age group dummies in a model such as in Table 2, it is possible to perform the Wald test. The results[6] indicate that the estimated coefficients of younger men (aged 25-29) in relation to older workers (aged 45-49, 50-54, and 55-59) are statistically different. In other words, it is necessary to dig deeper into the phenomenon by considering its relative supplies.

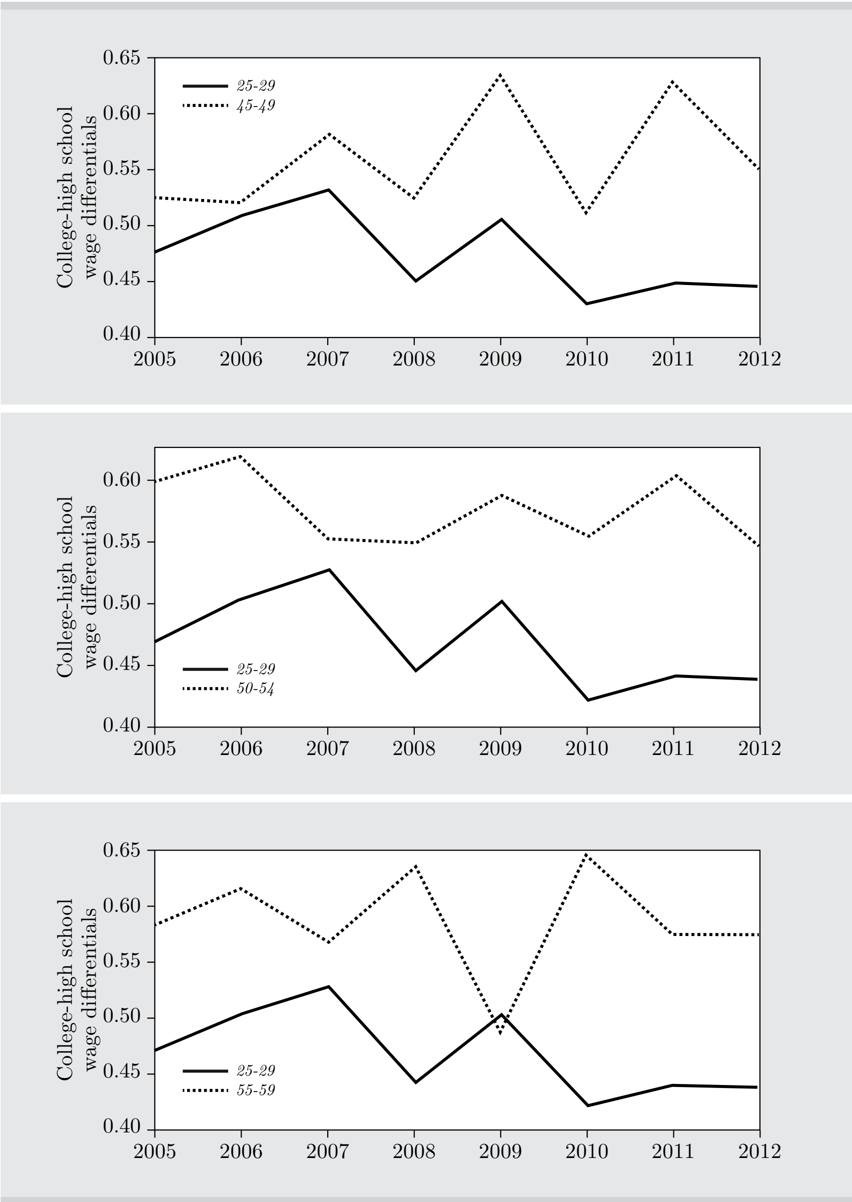
Figure 2 plots the college-high school wage gap for younger and older groups over the period analyzed. The overall patterns for the college premium for men aged 25-29 and 45-49 are very similar across years, but it appears that the evolution of the college premium for younger and older workers has tended to diverge since the financial and economic crisis of 2008. Once again, these results underscore the importance of studying differentials in the college premium among cohort groups. By examining the wage gaps for older workers aged 50-54 and 55-59, it is possible to infer that the difference has remained practically constant, with slight decreases toward the end of 2012. The employment stability earned by members of this age group may be a potential explanatory factor.

The more relevant change revealed by Figure 2 is the growing trend in the wage differential in the group of male workers aged 45-49. Conversely, the college premium among younger workers seems to have declined; it is important to highlight the fact that the financial crisis has had a particularly adverse effect on younger workers in Mexico (Villarreal, 2010). The importance of this lies in the fact that on the one hand, in the short term younger people have an incentive not to continue to higher levels of education. On the other hand, if young people decide to invest in higher levels of education, the main effect on the college premium will be observed in the long term.

To understand this change, we need to be aware not only of the gap between workers with the same education and different ages, but also of

6. The p-values in chi-square tests are $p < 0.01$ in all cases.

Figure 2. College-high school wage differentials by age group



Source: Table 3.

Table 3. Estimated college-high school wage differentials by age groups

Year/age	25-29	30-34	35-39	40-44	45-49	50-54	55-59
2005	0.470	0.530	0.534	0.528	0.521	0.599	0.584
	(0.032)	(0.032)	(0.032)	(0.034)	(0.038)	(0.047)	(0.074)
2006	0.504	0.469	0.598	0.508	0.517	0.619	0.613
	(0.030)	(0.031)	(0.030)	(0.031)	(0.039)	(0.048)	(0.070)
2007	0.528	0.518	0.531	0.515	0.579	0.552	0.568
	(0.030)	(0.032)	(0.030)	(0.032)	(0.035)	(0.046)	(0.061)
2008	0.444	0.526	0.541	0.566	0.521	0.549	0.634
	(0.031)	(0.033)	(0.033)	(0.032)	(0.036)	(0.044)	(0.062)
2009	0.501	0.574	0.491	0.528	0.633	0.588	0.488
	(0.034)	(0.029)	(0.032)	(0.035)	(0.037)	(0.044)	(0.065)
2010	0.422	0.501	0.517	0.502	0.506	0.554	0.645
	(0.030)	(0.031)	(0.031)	(0.034)	(0.034)	(0.044)	(0.067)
2011	0.441	0.500	0.545	0.516	0.627	0.602	0.575
	(0.029)	(0.030)	(0.032)	(0.033)	(0.035)	(0.044)	(0.059)
2012	0.438	0.506	0.561	0.594	0.546	0.545	0.573
	(0.030)	(0.031)	(0.031)	(0.033)	(0.039)	(0.042)	(0.056)

Source: Own calculations using ENOE data from Q1 2005 to Q4 2012.
Note: Robust standard errors in parentheses.

the composition of the relative supplies. Using Equation (11)[7], we can compute the average weekly hours worked by men aged 25-59 with any level of education; the results are contained in Table 4. According to the notation, C_{jt} corresponds to the number of hours per week worked by college graduates by age group j in year t including postgraduate workers, weighted by their wage gap. Likewise, H_{jt} is the total number of weekly hours worked by the labor force with incomplete college or less education, weighted by wage differentials.

The estimation of the relative supply of workers with a college degree shows an important growth trend in young workers, with an increase from 0.3 to 0.4 in less than a decade. This ratio may be explained because returns to education in Mexico are substantial and higher than those estimated for developed countries (see Psacharopoulos, *et al.* (1996), Psacharopoulos and Patrinos (2004), or Canton (2007) for details), whereas the natural laws of supply and demand would typically

7. The procedure is to compute Equation (11) estimating $\hat{\tau}^{s}_{tk}$ in separate regressions depending on s. A total of 24 panel data models are used to construct Table 4, taking into account differences in the effective supply of skilled and low-skilled workers. The results are not reported but are available from the author upon request.

Table 4. Relative college labor supply by age group

Year/age	25-29	30-34	35-39	40-44	45-49	50-54	55-59
				College			
2005	59,657	50,169	52,764	41,174	45,751	28,685	17,697
2006	60,193	55,440	55,785	50,308	47,343	32,394	17,109
2007	62,544	51,116	56,128	55,115	45,513	30,686	19,388
2008	62,478	48,750	49,749	52,419	49,444	33,686	21,281
2009	58,072	51,055	53,408	49,704	46,495	34,073	22,090
2010	60,201	49,328	50,085	51,993	48,848	34,541	20,516
2011	63,335	49,672	49,258	50,424	45,126	36,840	23,187
2012	64,801	50,613	49,830	49,923	44,150	38,609	22,775
				High school			
2005	186,224	169,608	167,428	155,711	126,039	103,986	77,399
2006	184,196	173,038	180,254	159,046	129,565	101,655	77,633
2007	177,967	169,256	169,789	163,969	132,831	107,536	78,591
2008	179,068	164,582	172,613	159,286	131,924	112,373	78,108
2009	173,832	152,572	166,466	153,568	131,646	109,908	78,071
2010	171,487	150,761	166,517	155,081	133,894	107,476	76,403
2011	176,408	158,454	175,127	162,847	138,741	113,022	82,016
2012	162,888	139,709	165,205	159,786	134,524	110,889	81,603
				Relative supply			
2005	0.320	0.296	0.315	0.264	0.363	0.276	0.229
2006	0.327	0.320	0.309	0.316	0.365	0.319	0.220
2007	0.351	0.302	0.331	0.336	0.343	0.285	0.247
2008	0.349	0.296	0.288	0.329	0.375	0.300	0.272
2009	0.334	0.335	0.321	0.324	0.353	0.310	0.283
2010	0.351	0.327	0.301	0.335	0.365	0.321	0.269
2011	0.359	0.313	0.281	0.310	0.325	0.326	0.283
2012	0.398	0.362	0.302	0.312	0.328	0.348	0.279

Source: Own calculations using ENOE data from Q1 2005 to Q4 2012.

yield a decrease in the college premium. Even more notably, the trend for the estimated supplies of older workers has a differentiated effect on the college premium. Individuals aged 45-49 recorded a decrease of around 4%, while conversely, individuals aged 50-54 and 55-59 increased their participation. In general, the results contained in tables 3 and 4 show a surprisingly mixed effect on wage differentials and relative supplies.

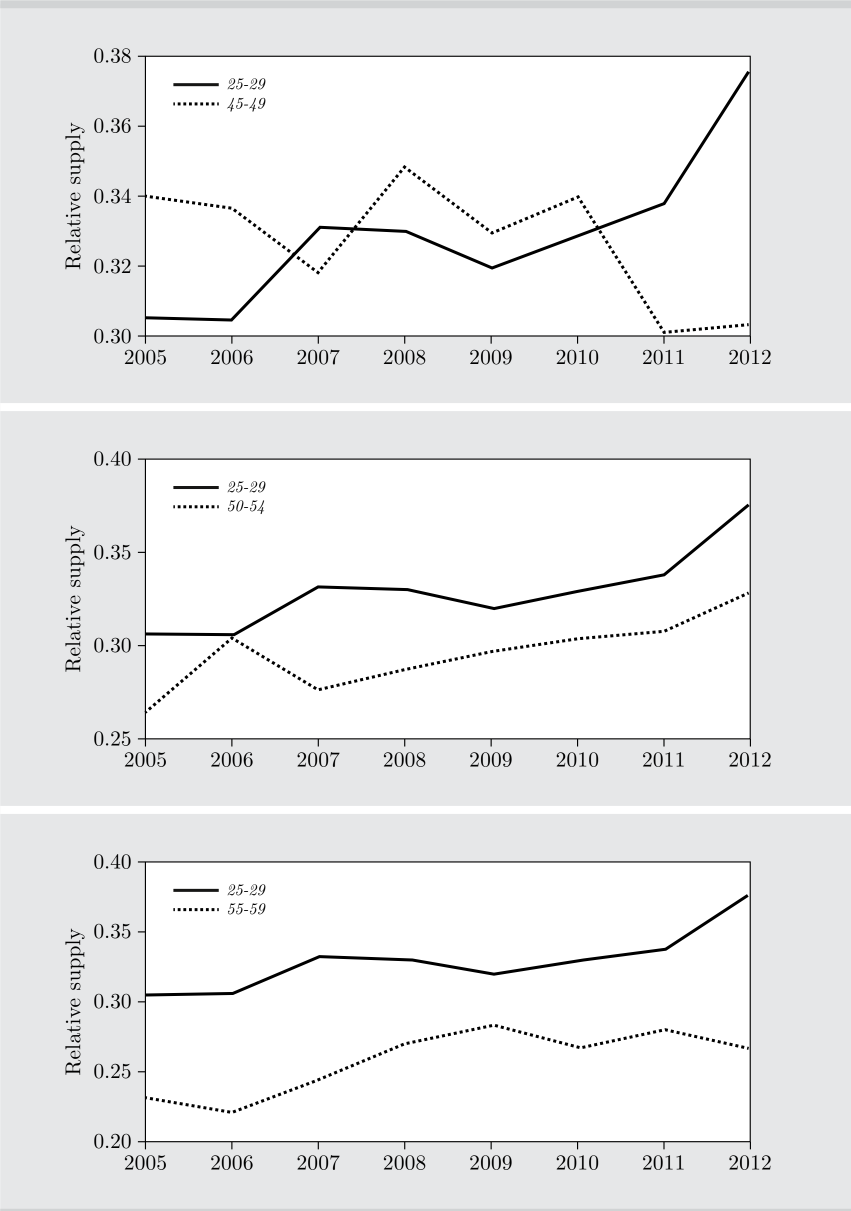
Figure 3 shows the behavior of relative supplies for each age group of interest. In all cases, with the exception of specific years for workers aged 45-49, the relative supply of the younger group is higher than the relative supply of older groups. One possible reason for this large difference is strongly suggested by the weighting parameter $\hat{\tau}$ in Equation (11). As previous work (Zepeda and Ghiara, 1999, Zamudio, 2001, Cañonero and Werner, 2002, and Hanson, 2003) has suggested, the recent increase in inequality in the income distribution is largely due to education. Thus, after trade liberalization occurred, the average relative wage of unskilled workers decreased and then, the effective supply may have been reduced for older cohorts.

Within this framework, the increase in the college-high school wage gap in workers aged 45-49 is attributable to steadily rising relative demand for college-educated labor, coupled with a dramatic slowdown in the rate of growth of the relative supply of college-educated workers. Card and Lemieux (2001) and Ferreira (2004) find similar behavior for groups of younger workers, but in the case of Mexico, younger workers are the segment of the population that is most affected, with unemployment rates twice as high as those of older adults. Furthermore, better-educated professionals have the highest unemployment rate (Villarreal, 2010).

The final step is to estimate two parameters of interest: the elasticity of substitution between the two education groups and the elasticity of substitution between different age groups. Using Equation (7) I establish the cohort born in 1956-1960 (aged 45-49) as the control group, hence the results presented correspond to differences with regard to older workers. The model presented in Table 5 suggests that a 1% increase in the relative supply of college labor causes a decrease of 0.33 percentage points in the college-high school wage differential, in the absence of non-neutral technology changes. Also, a 1% increase in the age-specific relative supply of college labor decreases the college premium by 0.6 percentage points, for that particular age group.

The empirical evidence provides an estimated elasticity of substitution between different age groups of around 1.7, while the parameter for the two education groups is about 3. We note that this specification does not seem to capture the annual wage gap (year effect) that would have, on average, increased in the absence of the age/cohort productivity factor and changes in supply according to educational level. In fact, the coefficients are not statistically different from 0 because the yearly dummies cannot capture technology shocks, which would require a long time-series sample and five-year interval dummies.

Figure 3. Relative college labor supply by age group



Source: Table 4.

Table 5 Estimated college-high school wage gap by cohort and year

Age-specific relative supply	-0.601*	(0.100)
Aggregate supply index	-0.332*	(0.121)
Age effect		
25-29	-0.121*	(0.012)
30-34	-0.090*	(0.011)
35-39	-0.063*	(0.009)
40-44	-0.050*	(0.007)
50-54	0.013	(0.008)
55-59	0.026*	(0.013)
Cohort effect		
1946-1950	-0.050*	(0.010)
1951-1955	-2E-04	(0.009)
1961-1965	0.015*	(0.006)
1966-1970	0.019*	(0.006)
1971-1975	0.024*	(0.007)
1976-1980	0.032*	(0.008)
1981-1985	0.029*	(0.009)
1986-1990	0.021	(0.012)
Year effect dummies		
2006	-0.010	(0.007)
2007	0.0010	(0.007)
2008	-0.010	(0.007)
2009	-0.003	(0.008)
2010	-0.003	(0.007)
2010	-0.024*	(0.008)
2012	-0.121*	(0.010)

Source: Own calculations using ENOE data from Q1 2005 to Q4 2012.
Note: Significant at 5%; robust standard errors in parentheses.

As a general rule, the elasticity of substitution captures the percentage change in relative demand for the two factors due to the change in relative factor prices at constant output. Also, the parameter of substitution between college and high school labor influences the impact of schooling on the education wage premium. In the traditional sense, the reduction is larger if the elasticity is low; i.e., they are far from perfect substitutes.

Similar studies for U.S, the U.K. and Canada (Card and Lemieux, 2001) and for Brazil (Ferreira, 2004) suggest that the elasticity of substitution between different age groups is 4.4. At the same time, the elasticity of substitution between college and high school labor types is about 2.5, and less than 2 in the case of Brazil. Nevertheless, these studies have analyzed long time-series data sets: 1959-1995 for the U.S., the U.K. and Canada, and 1976-1998 for Brazil. In marked contrast, this study of Mexico is focused on the short term, taking advantage of quarterly datasets. The results in Table 5 indicate that the two education groups are more easily substituted in Mexico than in other countries. In contrast, the substitution of workers between different age groups seems to be more complicated, based on the small estimated value for the elasticity of substitution.

This finding could be explained by the structural change occurring in Mexico's labor market. According to the literature review, empirical studies for Mexico have concluded that the wage gap between the two education groups has tended to widen with trade liberalization. The intuition behind previous studies is that employers reward skilled workers with an important college wage premium, making them hard to substitute. However, the combination in recent years of a high unemployment rate for skilled workers, slow economic growth, and the negative effects of the financial crisis could potentially explain this result.

6. Concluding remarks

This paper has introduced the first evidence in Mexico of the estimated evolution of the college wage gap by age groups, controlling for the relative supply of college graduate workers. Following the econometric methodology proposed by Card and Lemieux (2001) and using rotating panel data from 2005 to 2012, I present the partial elasticity of substitution between college and high school workers and across age groups. The results of the estimation indicate the existence of a large elasticity substitution (around 3) between male workers with different levels of education, which could mean that college- and high-school-educated workers are considered easily substitutable by employers. Furthermore, the small value of the estimated parameter (around 1.7) between the two education groups suggests that younger and older workers are viewed as different by employers and they are far from perfect substitutes. This result is very important because it suggests that in the current Mexican labor market it is easier to substitute skilled workers with unskilled workers rather than replace older workers with a younger labor force.

Overall, the model presents a negative and significant effect on the college premium for cohort variations in relative supply. This study highlights a decreasing trend in the college wage gap for younger workers (aged 25-29) combined with an increasing trend in the wage gap for older workers (aged 45-49). On the other hand, the behavior of the college premium for the oldest groups (aged 50-54 and 55-59) describes a constant trend in the wage gap.

The overall patterns for the college premium for men aged 25-29 and 45-49 are very similar across years but it appears that the evolution of the college premium for younger and older workers has tended to diverge after the financial and economic crisis of 2008. This underscores the need for wage gap studies in Mexico to consider the specific composition of workers by age group. According to this model, valid in a scenario of perfectly competitive equilibrium, the recent downward trend in the college premium for younger men depends mainly on the age effect. Although the results imply imperfect substitution between skilled and unskilled labor, it seems that in the case of Mexico, in contrast with the U.S., the U.K, Canada or Brazil, there is a small elasticity of substitution between the two age groups. An important implication of this finding is that, because of the aging of the population and increased levels of schooling, younger, educated workers are the segment of the population that is most affected.

The demographic and schooling transformation now underway in Mexico has the potential to both help and hinder its overall economic development agenda. Modifications of Mexico's Federal Labor Law enacted in 2012 brought important changes from an employer's perspective. One of the most interesting is the expansion of the types of employment relationships that are legally allowed. In addition to the already existing contracts for an indefinite term or a specific project, the reform introduced the seasonal employment category, which allows short-term employment to cover the need for additional workforce requirements during seasonal peaks, and the temporary employment contract, which permits short-term employment to cover immediate needs. In principle, the new recruitment scheme could have a positive impact on highly educated younger workers since seasonal or temporary employment could provide them with jobs and enable them to begin gaining experience quickly.

Finally, the results imply not only imperfect substitution between older and younger men but also between skilled and unskilled men. Future studies could include more than two categories of educated workers, including for example those with incomplete high school, incomplete college, and advanced degrees, as well as an extension for women in labor markets, controlling for self-selection.

REFERENCES

Airola, J. and C. Juhun (2008), "Wage inequality in post-reform Mexico," *Journal of Income Distribution* **17**(1): 110–34.

Ampudia, N. (2007), "Política monetaria, restricción crediticia y empleo," in Mantéy, M. and N. Levy, eds., *Políticas macroeconómicas para países en desarrollo.* Mexico: Miguel Ángel Porrúa-UNAM.

Becker, G. (1965), "A theory of the allocation of time," *The Economic Journal* **75**(299): 493–517.

Bia, M., R. Leombruni, and P. Messe (2009), "Young in-old out: A new evaluation based on generalized propensity score," *LABORatorio R. Revelli*, Working Paper No. 93.

Cañonero, G. and A. Werner (2002), "Salarios relativos y liberación del comercio en México," *El Trimestre Económico* **69**(273): 123–42.

Canton, E. (2007), "Social returns to education: Macro-evidence," *De Economist* **155**(4): 449–68.

Card, D. and T. Lemieux (2001), "Can falling supply explain the rising return to college for younger men? A cohort-based analysis," *Quarterly Journal of Economics* **116**(2): 705–46.

Chiquiar, D. (2008), "Globalization, regional wage differentials and the Stolper-Samuelson theorem: Evidence from Mexico," *Journal of International Economics* **74**(1): 70–93.

Cragg, M. and M. Epelbaum (1996), "Why has wage dispersion grown in Mexico? Is it the incidence of reforms or the growing demand for skills?" *Journal of Development Economics* **51**(1): 99–116.

Ferreira, S. (2004), "The provision of education and its impacts on college premium in Brazil," *Revista Brasileira de Economia* **58**(2): 211–33.

Gutiérrez, R. (2009). "La recesión y la medición de la tasa de desempleo de México ante una eventual reforma laboral," *Economía, Teoría y Práctica* **31**: 43–72.

Hairault, J., A. Chéron, and F. Langot (2007), "Job creation and job destruction over the life cycle: The older workers in the spotlight," IZA Discussion Paper No. 2597.

Hanson, G. (2003), "What has happened to wages in Mexico since NAFTA? Implications for hemispheric free trade," NBER Working Paper 9563.

Hanson, G. and A. Harrison (1995), "Trade, technology and wage inequality," NBER Working Paper 5110.

Hanson, G. and A. Harrison (1999), "Who gains from trade reform? Some remaining puzles," *Journal of Development Economics* **59**(1): 125–54.

Hernández, E. (2000), *Empleo y necesidades de capacitación en mercados regionales de trabajo en México.* Mexico: Secretaría del Trabajo y Previsión Social.

Huesca, L. (2004), "La rentabilidad de la escolaridad en los hogares asalariados en México durante 1984-2000," *Problemas del Desarrollo* **35**(138): 125–254.

Islas, A. and W. Cortez (2011), "How relevant is monetary policy to explain Mexican unemployment fluctuations?" MPRRA Paper No. 30027.

Kalwij, A., A. Kapteyn, and K. Vos (2010), "Retirement of older workers and employment of the young," *De Economist* **158**(4): 341–59.

Katz, L. and K. Murphy (1992), "Changes in relative wages, 1963-1987: Supply and demand factors," *Quarterly Journal of Economics* **107**(1): 35–78.

López, G. (2006), "Mexico: Two decades of the evolution of education and inequality," World Bank Policy Research Working Paper 3919.

Meza, L. (1999), "Cambios en la estructura salarial de México en el período 1988-1993 y el aumento en el rendimiento de la educación superior," *El Trimestre Económico* **66** (262): 189–226.

Mincer, J. (1974), *Schooling, experience and earnings.* New York: National Bureau of Economic Research.

Prskawetz, A., T. Tsachev, and V. Veliov (2012), "Optimal education in an age-structured model under changing labor demand and supply," *Macroeconomic Dynamics* **16**(2): 159–83.

Psacharopoulos, G. and H. Patrinos (2004), "Returns to investment in education: A further update," *Education Economics* **12**(2): 111–34.

Psacharopoulos, G., E. Velez, A. Panagides, and H. Yang (1996), "Returns to education during economic boom and recession: Mexico, 1984, 1989 and 1992," *Education Economics* **4**(3): 219–30.

Robertson, R. (2004), "Relative prices and wage inequality: Evidence from Mexico," *Journal of International Economics* **64**(2): 387–409.

Rodríguez-Oreggia, E. (2009), "La dinámica comparativa del sector informal en México: Temas actuales de política económica," in I. Soloaga, ed., *Sobre México.* Mexico: UPAEP.

Roger, M. and M. Wasmer (2009), "Heterogeneity matters: Labor productivity differentiated by age and skills," Paris School of Economics Working Paper No. 2009-51.

Rones, P. (1983), "Labor market problems of older workers," *Monthly Labor Review* May: 1–12.

Tello, C. and R. Ramos (2012), "Wage inequality and economic growth in Mexican regions," *Investigaciones Regionales* **24**: 115–52.

Varela, R., *et al.* (2010), "Determinantes de los ingresos salariales en México: Una perspectiva de capital humano," *Región y Sociedad* **22**(49): 117–42.

Villarreal, M. (2010), *The Mexican economy after the global crisis.* Mexico: UNT Libraries Government Documents Department.

Zamudio, A. (2001), "La escolaridad y la distribución condicional del ingreso: Una aplicación de regresión cuantil," *El Trimestre Económico* **68**(269): 39–70.

Zepeda, E. and R. Ghiara (1999), "Determinación del salario y capital humano en México: 1987–1993," *Economía, Sociedad y Territorio* **2**(5): 67–116.

APPENDIX

Table A1. Estimated college-high school wage differentials for men aged 25-59, specification 2

	2005	2006	2007	2008	2009	2010	2011	2012
Age	0.023* (0.006)	0.028* (0.006)	0.023* (0.006)	0.029* (0.006)	0.024* (0.006)	0.028* (0.006)	0.028* (0.006)	0.034* (0.006)
Age^2	-1.8E-04* (8.5E-05)	-2.5E-04* (7.6E-05)	-2E-04* (7.3E-05)	-2.8E-04* (7.5E-05)	-2.3E-04* (7.4E-05)	-2.6E-04* (7.1E-05)	-2.8E-04* (7E-05)	-3.6E-04* (7E-05)
College	0.376* (0.015)	0.399* (0.015)	0.409* (0.015)	0.393* (0.015)	0.400* (0.016)	0.365* (0.015)	0.402* (0.015)	0.387* (0.015)
Married	0.103* (0.014)	0.119* (0.014)	0.113* (0.013)	0.093* (0.013)	0.081* (0.013)	0.075* (0.012)	0.078* (0.012)	0.078* (0.012)
o2	0.226* (0.024)	0.185* (0.025)	0.210* (0.025)	0.187* (0.027)	0.218* (0.028)	0.237* (0.025)	0.202* (0.025)	0.282* (0.025)
o3	0.172* (0.027)	0.136* (0.027)	0.149* (0.026)	0.135* (0.029)	0.103* (0.032)	0.091* (0.027)	0.125* (0.031)	0.180* (0.031)
o4	-0.085* (0.017)	-0.073* (0.018)	-0.059* (0.016)	-0.094* (0.018)	-0.065* (0.018)	-0.093* (0.018)	-0.073* (0.016)	-0.065* (0.017)
o5	-0.191* (0.017)	-0.164* (0.017)	-0.140* (0.016)	-0.192* (0.017)	-0.188* (0.017)	-0.195* (0.017)	-0.176* (0.016)	-0.167* (0.016)
o6	-0.261* (0.021)	-0.282* (0.021)	-0.219* (0.019)	-0.253* (0.020)	-0.277* (0.021)	-0.297* (0.021)	-0.288* (0.020)	-0.271* (0.019)
o7	-0.305* (0.021)	-0.297* (0.022)	-0.277* (0.022)	-0.297* (0.022)	-0.319* (0.021)	-0.313* (0.022)	-0.318* (0.021)	-0.290* (0.021)
o8	-0.331* (0.024)	-0.301* (0.025)	-0.238* (0.023)	-0.326* (0.024)	-0.276* (0.024)	-0.281* (0.024)	-0.261* (0.023)	-0.290* (0.020)

Table A1. (continued)

	2005	2006	2007	2008	2009	2010	2011	2012
o9	-0.395*	-0.393*	-0.393*	-0.367*	-0.346*	-0.404*	-0.345*	-0.278*
	(0.028)	(0.029)	(0.027)	(0.027)	(0.027)	(0.027)	(0.026)	(0.025)
o10	-0.784*	-0.767*	-0.774*	-0.739*	-0.793*	-0.755*	-0.744*	-0.724*
	(0.044)	(0.043)	(0.047)	(0.039)	(0.042)	(0.038)	(0.035)	(0.040)
R1	0.285*	0.272*	0.316*	0.311*	0.313*	0.288*	0.273*	0.254*
	(0.02)	(0.019)	(0.019)	(0.019)	(0.021)	(0.018)	(0.019)	(0.019)
R2	0.156*	0.155*	0.150*	0.142*	0.143*	0.110	0.118*	0.103*
	(0.019)	(0.018)	(0.018)	(0.019)	(0.019)	(0.018)	(0.017)	(0.018)
R3	0.104*	0.153*	0.140*	0.122*	0.145*	0.108*	0.126*	0.117*
	(0.017)	(0.017)	(0.016)	(0.016)	(0.017)	(0.016)	(0.016)	(0.016)
R5	-0.039*	-0.026	-0.016	-0.007	-0.015	0.011	-0.018	-0.022
	(0.018)	(0.018)	(0.019)	(0.019)	(0.018)	(0.018)	(0.017)	(0.017)
N	20,468	21,081	21,082	20,083	19,173	20,019	19,785	19575
R^2	0.297	0.292	0.289	0.277	0.294	0.274	0.291	0.271

Source: Own calculations using ENOE data from Q1 2005 to Q4 2012.
Note: o1 = Professional, technical and craft workers; o2 = Education; o3 = Officers and directors; o4 = Clerks; o5 = Industrial craftsmen and assistants; o6 = Merchants; o7 = Transport operators; o8 = Personal service; o9 = Protection and surveillance; o10 = Farmworkers.
Significant at 5%; robust standard errors in parentheses.

Table A2. Estimated college-high school wage differentials for men aged 25-59, specification 3

	2005	2006	2007	2008	2009	2010	2011	2012
Age	0.021*	0.028*	0.023*	0.030*	0.028*	0.027*	0.025*	0.034*
	(0.007)	(0.007)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)
Age^2	-1.6E-04	-2.4E-04*	-1.9E-04*	-2.8E-04*	-2.5E-04*	-2.4E-04*	-2.3E-04*	-3.5E-04*
	(8.8E-08)	(8.3E-05)	(8E-05)	(8.1E-05)	(8E-05)	(7.8E-05)	(7.7E-05)	(7.5E-05)
College	0.542*	0.545*	0.551*	0.549*	0.564*	0.531*	0.572*	0.551*
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Married	0.113*	0.121*	0.111*	0.091*	0.081*	0.074*	0.078*	0.074*
	(0.015)	(0.015)	(0.014)	(0.014)	(0.014)	(0.013)	(0.013)	(0.013)
p2	0.212*	0.211*	0.178*	0.143*	0.129*	0.089*	0.118*	0.106*
	(0.021)	(0.021)	(0.021)	(0.021)	(0.022)	(0.022)	(0.023)	(0.022)
p3	-0.087*	-0.111*	-0.095*	-0.089*	-0.161*	-0.124*	-0.103*	-0.119*
	(0.018)	(0.018)	(0.017)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
R1	0.264*	0.267*	0.302*	0.312*	0.311*	0.275*	0.269*	0.240*
	(0.021)	(0.021)	(0.020)	(0.021)	(0.022)	(0.020)	(0.020)	(0.020)
R2	0.148*	0.161*	0.147*	0.148*	0.149*	0.108*	0.128*	0.107*
	(0.020)	(0.019)	(0.020)	(0.021)	(0.021)	(0.020)	(0.019)	(0.019)
R3	0.103*	0.154*	0.142*	0.124*	0.139*	0.104*	0.116*	0.104*
	(0.019)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.017)	(0.017)
R5	-0.047*	-0.029	-0.023	-3.03E-04	-0.016	0.005	-0.027	-0.028
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.019)	(0.019)	(0.019)
N	20,468	21,081	21,082	20,083	19,173	20,019	19,785	19,575
R^2	0.215	0.22	0.217	0.196	0.213	0.189	0.206	0.189

Source: Own calculations using ENOE data from Q1 2005 to Q4 2012.
Note: p1 = Employees and subordinates; p2 = Employers; p3 = Self-employed.
Significant at 5%; robust standard errors in parentheses.

Table A3. Estimated college-high school wage differentials for men aged 25-59, specification 4

	2005	2006	2007	2008	2009	2010	2011	2012
Age	0.022*	0.026*	0.020*	0.028*	0.023*	0.027*	0.027*	0.032*
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
Age^2	-1.8E-04*	-2.3E-04*	-1.8E-04*	-2.7E-04*	-2.1E-04*	-2.6E-04*	-2.7E-04*	-3.3E-04*
	7.80E-08	7.50E-05	7.20E-05	7.30E-05	7.30E-05	7.00E-05	6.80E-05	6.80E-05
College	0.362*	0.375*	0.391*	0.379*	0.386*	0.354*	0.384*	0.372*
	(0.015)	(0.015)	(0.014)	(0.015)	(0.015)	(0.015)	(0.014)	(0.014)
Married	0.097*	0.114*	0.107*	0.089*	0.075*	0.073*	0.078*	0.075*
	(0.014)	(0.013)	(0.013)	(0.013)	(0.013)	(0.012)	(0.012)	(0.012)
Economic activity dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Position dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	20,468	21,081	21,082	20,083	19,173	20,019	19,785	19,575
R^2	0.321	0.318	0.312	0.299	0.319	0.294	0.313	0.294

Source: Own calculations using ENOE data from Q1 2005 to Q4 2012.
Note: Significant at 5%; robust standard errors in parentheses.