uc3m | Universidad Carlos III de Madrid

Bachelor in Economics

2023-2024

Bachelor's thesis

"Monopsony Power and Its Impact on Women"

Jaime de Andrés Velasco

Thesis Advisor

Juan Jose Dolado

Getafe, June 2024

ABSTRACT

This paper analyzes the impact of monopsony power on workers' wages in Spain differentiating by gender. To do so, data from the Continuous Sample of Working Lives (MCVL) provided by social security and an own-elaborated labor market concentration index are used. A theoretical monopsony model is set out to describe the relationship between labor concentration and labor supply elasticity. Additionally, an instrumental variables method is used to identify the wage elasticity response to changes in employment concentration. The results show how an increase in concentration leads to a reduction in wages, but the effects are ambiguous when differentiating by gender.

Keywords: Labor market, monopsony, gender wage gap, labor supply elasticity, labor market concentration.

RESUMEN

Este trabajo analiza el impacto del poder de monopsonio sobre los salarios de los trabajadores en España diferenciando por género. Para ello, se utilizan datos de la Muestra Continua de Vidas Laborales proporcionada por la seguridad social y un índice de concentración del mercado laboral de elaboración propia. Se desarrolla un modelo teórico de monopsonio para describir la relación entre la concentración laboral y la elasticidad de oferta laboral. Además, se usa un método de variables instrumentales para identificar la elasticidad del salario sobre la concentración. Los resultados muestran como el incremento de la concentración conlleva una reducción de los salarios, pero los efectos son ambiguos al diferenciar por género.

Palabras clave: Mercado laboral, monopsonio, brecha salarial de género, elasticidad de la oferta de trabajo, concentración laboral, Muestra Continua de Vidas Laborales

ÍNDICE GENERAL

ÍNDICE DE FIGURAS

ÍNDICE DE TABLAS

1. INTRODUCTION

The minimum wage in Spain has increased by 54 % since 2018. One of the main arguments used to justify these successive increases in recent years is the theory of monopsony.

Monopsony is a type of market where there is only one buyer of goods or services. In the labor market, this occurs if there is only one company demanding labor. The labor supply in such markets has a positive slope, unlike in competitive markets where it is flat, and the employer (monopsonist) can only hire more workers if they offer higher wages.

In this context, with marginal costs exceeding the wage, the monopsonist ends up paying an equilibrium wage below the worker's marginal productivity, resulting in lower levels of employment and wages compared to a competitive situation. Therefore, assuming a monopsonistic market, the wage can be increased (up to the competitive wage) without affecting, or even increasing, employment (Card & Krueger, 1994).

In parallel, both in the Spanish labor market and in most countries around the world, there is a gender pay gap between men and women; even after controlling for various observable variables —such as experience, children, education, or economic sector— this differential remains significant.

The hypothesis of this study is that women, due to their specific characteristics —such as lower labor mobility, caregiving responsibilities, or challenges in salary negotiation— have lower labor supply elasticity, and as a result, the effect of monopsony power may be greater on them. If this is true, the monopsonistic exploitation they face could explain part of the gender pay gap.

The aim of this study is to analyze the effect of monopsony power on wages in general and, specifically, to identify how it affects women in Spain. To measure monopsony power, a concentration measure of companies by sector in the labor market will be used, calculated from data from the Continuous Sample of Working Lives (MCVL) provided by Social Security. An instrumental variables strategy will be used to estimate causal parameters.

The results from the model estimated with instrumental variables indicate that a 1 % increase in labor concentration decreases wages by an average of 1.93 %. However, when including the gender interaction in the analysis, it is observed that although men's wages are negatively affected, women's wages increase slightly. This latter finding does not align with existing theory or empirical studies in other countries. This could be due to the lower female labor participation in Spain, which implies higher female labor supply elasticity compared to males, particularly concerning the extensive margin.

The structure of the paper is detailed as follows. Section 2 provides a review of the existing literature, including previous studies on monopsony power, gender pay inequality, and labor market concentration measures. Next, Section Section 3 develops a simple theoretical framework integrating these concepts. Section **??** presents the empirical specification and methodology used. Following this, Section Section 5 describes the data utilized. Section **??** presents the results obtained. Finally, Section Section 6 concludes and discusses the limitations of the study. The Appendix includes additional Tables and Figures discussed in the main text.

2. LITERATURE REVIEW

The pioneering study that developed the idea of monopsony in the labor market dates back to Robinson, 1933, in a book where the author discusses existing theories on market structures and business behavior, introducing novel concepts for the time, such as monopsony and price discrimination. Robinson explains monopsony as a theory adjacent to the monopoly theory developed up to that point. In this sense, while the monopolist faces a downward-sloping demand curve for goods, the monopsonist's situation involves an upward-sloping labor supply curve.

Since then, the ideas introduced by Joan Robinson have been extensively developed by numerous economists from various perspectives. Langella y Manning, 2021 highlight two distinct approaches.

The traditional approach describes monopsony with static models, assuming there are unique aspects that influence how the worker values their employment. The labor supply curve's elasticity, and specifically, concentration indices are used as a measure of monopsony. Boal y Ransom, 1997 is a good example, expressing Pigou's exploitation rate as a function of labor supply elasticity, which I will further develop in this paper. Some publications using this model with concentration measures are Azar et al., 2019 and Tenreyro et al., 2018.

Manning, 2003 proposes studying monopsony using dynamic models. In particular, he uses the search and matching model developed by Burdett y Mortensen, 1998, where the measure of monopsony is the elasticity of separation and recruitment with respect to wages.

Ransom y Oaxaca, 2010 use this strategy to approximate labor supply elasticity based on the elasticity of separation with respect to wages. Their results show that women's elasticity is lower than that of men and that a static model can adequately describe the differences in monopsony power between men and women.

Both Hirsch et al., 2006 and Barth y Dale-Olsen, 2009 also use a dynamic model to estimate labor supply elasticity by gender with a panel of firms and workers in Germany and Norway, respectively. The former finds results similar to Ransom y Oaxaca, 2010, but with smaller elasticities. The latter notes that "between 20 and 90 % of the wage difference among individuals with less education can be attributed to labor market frictions."

3. THEORETICAL FRAMEWORK

3.1. Theoretical Specification

3.1.1. Model Assumptions

- There are M firms and n workers in the labor market.
- Each firm independently decides the amount of labor it will hire; there is no cooperation.
- Firms compete in the goods market and face a labor supply curve.
- Firms are price takers in the goods market.
- Each firm maximizes its profits considering the responses of other firms in the market.
- There is perfect information about labor supply and prices in the market.

3.1.2. Terminology

- M: Number of firms in the goods market.
- $Y_i(N_i)$: Revenue of firm i as a function of the number of workers in firm i.
- $w(N)$: Wage rate as a function of the total number of workers N in the labor market.
- N_i: Number of workers in firm i.
- N: Total number of workers in the labor market.
- π_i: Profit of firm.

3.2. Model derivation

Profit Function:

The profit function for the oligopsonist i represents the difference between the revenue generated by hiring labor and the labor cost incurred, that is:

$$\pi_i = Y_i(N_i) - w(N)N_i \tag{3.1}$$

The first order condition (f.o.c) of the profit function with respect to the employment is:

$$\frac{d\pi_i}{dN_i} = \frac{dY_i(N_i)}{dN_i} - \frac{d(w(N)N_i)}{dN_i} = \frac{dY_i}{dN_i} - N_i\frac{dw(N)}{dN} - w(N) = 0$$

$$\pi_i' = Y_i'(N_i) - w'(N)N_i - w(N) = 0$$

$$Y_i' = w'(N)N_i + w(N)$$

The labor supply elasticity is:

$$\epsilon_N = \frac{\partial N}{\partial w}\frac{w}{N}$$

And therefore, the inverse of the labor supply elasticity is given by:

$$\frac{1}{\epsilon_N} = \frac{\partial w}{\partial N}\frac{N}{w} = \frac{w'(N)\cdot N}{w} =$$

Solving for $w'(N)$ and substituting into the f.o.c, the **monopsonistic exploitation rate of firm** i is:

$$E_i = \frac{Y'(N) - w}{w} = \frac{N_i}{N}\cdot \epsilon_N^{-1} \tag{3.2}$$

where E_i measures the extent to which each firm can extract surplus from its workers.

Weighted Average Exploitation Rate

The weighted average exploitation rate E captures the overall level of exploitation within the labor market, taking into account the distribution of employment among firms. It is calculated as:

$$E = \sum_{i=1}^{M}\left(\frac{N_i}{N}\cdot E_i\right) \tag{3.3}$$

Herfindahl-Hirschman Index (HHI):

The HHI is a measure of market concentration used to assess the level of competition within a market. It is defined as:

$$HHI = \sum_{i=1}^{M}\left(\frac{N_i}{N}\right)^2 \tag{3.4}$$

The HHI ranges from 0 to 1, where higher values indicate greater market concentration. A value close to 1 suggests a highly concentrated market, where a few firms dominate employment, while a value close to 0 indicates a more competitive market with a more uniform distribution of employment.

To express the Herfindahl-Hirschman Index (HHI) as a function of the labor supply elasticity (ϵ) and the weighted average exploitation rate by employment (E), we first need to define the relationship between the exploitation rate (E) and elasticity (ϵ).

We know that the weighted average exploitation rate by employment (E) is calculated as:

$$E = \sum_{i=1}^{M} \left(\frac{N_i}{N} \cdot E_i \right)$$

and the exploitation rate of each company E_i is defined as:

$$E_i = \frac{N_i}{N} \cdot \epsilon_N^{-1}$$

Since E_i is constant for all firms ($E_i = \frac{N_i}{N} \cdot \epsilon_N^{-1}$), we can substitute it into the formula for E to obtain:

$$E = \epsilon_N^{-1} \sum_{i=1}^{M} \left(\frac{N_i}{N} \right)^2$$

$$E = \epsilon_N^{-1} \cdot HHI \tag{3.5}$$

Therefore, the Herfindahl-Hirschman Index (HHI) is expressed as a function of the labor supply elasticity (ϵ) and the weighted average exploitation rate by employment (E) as $HHI = \epsilon_N \cdot E$.

If we assume two completely segregated groups, the group with the higher elasticity will be able to obtain higher wages as it will be less subject to monopsonistic exploitation, as shown by equation Equation 3.2. Additionally, for heterogeneous worker groups, if the firm observes differences in mobility costs by group, the firm will offer the lowest wage to the group with the highest mobility costs.

Equation Equation 3.5 implies that, for a constant elasticity value, an increase in labor market concentration leads to an increase in the weighted exploitation rate and, consequently, a decrease in wages.

4. METHODOLOGY AND EMPIRICAL SPECIFICATION

To describe the effect of labor concentration and monopsony power on workers' wages, the $\log(Salary)$ is regressed on $\log(HHI)$ and a set of control variables.

$$\log(Salary) = \beta_0 + \beta_1 \cdot \log HHI_{t,j,p} + \beta_2 \cdot X_i + \mu_t + \mu_j + \mu_p + \varepsilon \tag{4.1}$$

[1]

The variables *Salary* and *HHI* have been transformed using the logarithmic transformation to establish the relationship in terms of elasticity; that is, the coefficient β_1 represents the percentage change in wages with respect to a 1

The control vector X_i includes the following variables, which will be analyzed in detail in the data section."

- *Female*: indicator of being female.
- *Education*: the *Basic Education* variable is used as the reference.
- *Children*: number of children.
- $Children \times Female$: Captures how the effect of having children on wages varies depending on the person's sex. The effect on women is expected to be negative.
- *Experience*: years in the labor market.
- $Experience^2$: The relationship between worker experience and wages traditionally describes a parabola with a negative slope. For the initial years of experience, wages are low; then, with more years of experience, wages increase, and towards the end of the worker's career, wages decrease.

Additionally, the regression includes fixed effects for year μ_t, economic activity μ_j, and province μ_p. The year fixed effects μ_t represent differences that occur from year to year, regardless of the other variables in the model. They are used to control for time-varying factors that affect wages. They can capture the effects of changes in the economic cycle or specific policies that impact all individuals equally. The economic activity fixed effects μ_j are used to capture sector-specific differences. For example, in economic sectors such as mining, there is a component of compensating wage differentials, where workers may earn higher wages to compensate for greater risk. Finally, the province fixed effects μ_p characterize economic differences between provinces. Agglomeration economies tend to result in higher wages in provinces with large cities and greater populations, De la Roca y Puga, 2017; they also capture differences in labor policies between regions.

[1] t, j, p indicating time, economic activity, and province, respectively.

To describe the relationship between concentration and gender, an interaction variable between $\log(HHI)$ and the variable *Female* is included in the regression.[2]

$$\log(Salary) = \beta_0+\beta_1\cdot\log HHI_{t,j,p}+\beta_2\cdot\text{Female}+\beta_3\cdot(\log HHI_j\times\text{Female})+X_i\cdot\beta_4+\mu_t+\mu_j+\mu_p+\varepsilon \tag{4.2}$$

The coefficient β_3 is the difference in the effect of concentration on wages between women and men.

$$\frac{\partial\log(Salary)}{\partial\log(HHI_{t,j,p})} = \beta_1 + \beta_3 \cdot \text{Female}$$

- When Female = 0 (Male):

$$\left.\frac{\partial\log(Salary)}{\partial\log(HHI_{t,j,p})}\right|_{\text{Female}=0} = \beta_1$$

For men, the effect of the concentration index on wages is β_1.

- When Female = 1 (Female):

$$\left.\frac{\partial\log(Salary)}{\partial\log(HHI_{t,j,p})}\right|_{\text{Female}=1} = \beta_1 + \beta_3$$

For women, the effect of the logarithm of the concentration index on wages is $\beta_1+\beta_3$.

The difference in the effect of concentration on wages between women ($Female = 1$) and men ($Female = 0$) is β_3.

4.1. Identification of Elasticity

The coefficients of the regressions described above may be biased due to omitted variables. In this case, the coefficient of the variable under study, $\log(HHI)$, may not be correctly identified if a third variable not included in the model simultaneously affects wages and the number of firms participating in the labor market, and thus the concentration index.

An instrumental variables strategy is used to solve this problem. Following Azar et al., 2022 and Bassanini et al., 2022, $\log(HHI_{t,j,p})$ is instrumented with the mean of $\log(1/N_{t,j,p'})$, where $N_{t,j,p'}$ is the number of firms in year t, in sector j, in the rest of the provinces p'. In effect, the number of firms in other regions is closely related to labor concentration and thus satisfies the relevance condition. At the same time, it is exogenous because it is independent of productivity changes within a specific labor market.

[2]The variable "Female"no longer appears in the control vector X_i.

However, productivity changes at the national level, between sectors, are not captured by the instrument. If these changes are related to other variables in the regression, the instrument may not be sufficient to identify elasticity.

For the case of Equation 4.2, $\log(HHI_{t,j,p}) \times Female$ is instrumented with the interaction of $\log(1/N_{t,j,p'}) \times Female$.

As an alternative, $\log(HHI_{t-1,j,p})$ has also been used as an instrument. This variable introduces a lag in the HH index. Given the persistence in labor concentration, $\log(HHI_{t-1,j,p})$ is correlated with $\log(HHI_{t,j,p})$. The justification for this instrument is that, in general, the factors affecting concentration in the previous period are different from the shocks that occur contemporaneously.[3]

To summarize, four models are estimated per equation. First, a model estimated by Ordinary Least Squares (OLS); then, three models estimated by Two-Stage Least Squares (2SLS) using as instruments $\log(1/N_{t,j,p'})$, $\log(HHI_{t-1,j,p})$, and both instruments simultaneously.

In all models, it is assumed that variance is not constant between provinces, but it is within them. Following Abadie et al., 2017, the method described by Arellano, 1987 is used, and standard errors clustered at the province level are calculated for all estimations.

[3]For the model estimated with the instrument $\log(HHI_{t-1,j,p})$, one year of observations is lost.

5. DATA

5.1. Data Source

The data come from the Continuous Sample of Working Lives (MCVL), which is provided by the Spanish Social Security and corresponds to a random sample of 4 percent of social security affiliates and pensioners. The information combines data extracted from Social Security databases and is supplemented with fiscal information and data from the Continuous Register of the National Statistics Institute (INE). Additionally, in this case, the MCVL with fiscal data provides additional information on the individual's income and other relevant data, such as the number of children or dependents.

5.1.1. Sample

The wave of the MCVL used is not the latest available and contains data from the year 2006 to 2019. Additional filters have been applied to the raw data provided by the Social Security in order to obtain a representative sample for the case study:

- Only those workers who remain in the same company during the reference year are retained in the database, excluding those who change companies, are fired, or are hired during the reference year.
- Individual and year are jointly keys in the database, meaning that multiple employment is not allowed. For those with more than one simultaneous job, the row with the lower wage income is removed from the database.
- Only full-time workers in the private sector are included.

Using these conditions, a data panel was created with all individuals who appear every year in the database (2006 - 2019), resulting in a total of 430,576 observations. The data panel including the public sector and full-time workers contained 648,654 observations.

5.2. Variables Used

5.2.1. Dependent Variable: Salary

The dependent variable used is salary. This represents the worker's labor income and is calculated as the sum of monetary and non-monetary compensation received by the worker in the reference year.

5.2.2. Herfindahl-Hirschman Index (HHI)

In the theoretical framework, the concentration index is defined as:

$$HHI_l = \sum_{i=1}^{M} \left(\frac{N_i}{N}\right)^2$$

where N_i is the number of workers in company i and N is the number of workers in the labor market l.

The labor market is defined at the level of industry-year-region, where the industry is measured at the 2-digit level of CNAE (2009), and the regions are provinces. This results in a total of 89 industries, 52 regions, and 14 years. Not all industries are present in all regions. Consequently, the number of observations per year varies, as shown in Table 6.1. The MCVL provides information on the number of workers in each secondary and primary social security account per worker. These data are used to calculate the concentration index.

With a random sample of workers rather than companies, it is expected that the concentration measures will be biased. To explore the sign and magnitude of this bias, I follow Tenreyro et al., 2018 and use the Monte Carlo method, which involves selecting random subsamples of companies and calculating the HHI for different values of M, and then comparing this value with that obtained using the entire sample. The results, shown in Figure 6.6, indicate that the bias is positive and not very large when more than 10 companies are used. The direction of the bias is as expected: the fewer the number of companies per labor market, the higher the concentration index.

To prevent this from affecting the regressions, all concentration indices calculated using workers from fewer than 10 companies per sector, province, and year are removed.

Notably, there is an increase in concentration during the 2008 crisis, as shown in Figure 6.4, and how it varies with population size by province, as explained by Luccioletti, 2022, is illustrated in Figure 6.2.

5.2.3. Control Variables

- Experience: a continuous variable measuring the years elapsed since the first affiliation to social security.
- Gender: a binary categorical variable with female as the reference category.
- Education: an ordinal categorical variable reflecting the level of education attained, with basic level below secondary education (Compulsory Secondary Education), intermediate level between secondary education and high school or second-degree vocational training or equivalent degrees, and higher education requiring at least a bachelor's degree. Basic education is taken as the reference.

- Number of Children: a discrete variable counting the number of children a person has.

5.3. Descriptive Analysis

The Table 5.1 presents the descriptive statistics for all the control variables, salary, and the instrument.

Variable	Qualitative	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Age	-	16	38	44	43.88	50	90
Experience	-	1	17	23	23.27	30	60
Gender	Men: 62.98 %	-	-	-	-	-	-
	Women: 37.02 %	-	-	-	-	-	-
Education	Basic: 29.60 %	-	-	-	-	-	-
	Intermediate: 55.67 %	-	-	-	-	-	-
	Higher: 14.78 %	-	-	-	-	-	-
Children	-	0	0	1	0.8578	2	15
Salary	-	300	17,121	23,599	30,124	34,670	4,342,523
HHI	-	0.00031	0.01420	0.03207	0.05884	0.07121	0.83135
N[4]	-	2	16	64	94.22	157	322

Tabla 5.1
Descriptive Statistics

Before filtering the sample to include only the private sector, the descriptive statistics were similar. Age and experience are, on average, slightly lower, the proportion of men increases slightly, and the average salary decreases by about €1000.

5.4. Model 1

First, the effect of concentration is estimated with the regression Equation 4.1 without including the interaction term $\log(HHI) \times Female$. Table 5.2 shows that in the model with the instrument $\log(1/N)$, a significant elasticity (p-value below 1 %) of $-1{,}93$ % is estimated. An increase in concentration of 1 % decreases, on average, the salary by approximately 1,93 %.

The first stages of the instrumental variable regressions are included in appendix **??**. It is concluded that the instruments are relevant and explain a large portion of the variance in the endogenous variable.

[4] The logarithm of the Herfindahl-Hirschman Index $\log(HHI)$ in the labor market $l = (j, p)$ at time t is instrumented with the average of $\log(N_{t,j,p'})$. In this context, "Nrepresents the average number of firms in the other geographic areas p' for the same occupation j and time period t.

For the remaining models, the estimated coefficients of $\log(HHI)$ are smaller in absolute value, very similar to each other, and are not significant for α values below 5 %. The estimated elasticity for the models without an instrument and with the lagged HHI instrument is −1,36 %; and −1,37 % in the model with both instruments. The lack of difference between the two models using the lagged HHI instrument and the simple model estimated with OLS may be due to insufficient variation between the instrument and the endogenous variable.

The coefficient estimated with the first instrument does show a significant variation compared to the coefficient estimated with OLS.[5] Additionally, it is consistent with the theoretical model proposed in Section 3, which predicted that an increase in labor concentration would, on average, lead to lower wages.

The other partial effects estimated with the control variables are significant and consistent across regressions. Women earn, on average, almost 25 % less than men. Workers with intermediate and higher education earn 16 % and 52 %, respectively, more than workers with basic education. The quadratic relationship between experience and salary is significant.

The total percentage change in salary for each additional child is positive both for men, $\beta_3 = 7{,}37\,\%$, and for women, $\beta_3 + \beta_8 = 3{,}16\,\%$. However, the coefficient β_8 indicates that there is a 4,21 % difference in the effect of children on the salary of women and men. The fact that both coefficients are positive indicates that the income effect of having children (greater need for income) outweighs the substitution effect (reduction in hours worked) in terms of its impact on salary.

The effect of children on women, $\beta_3 + \beta_8$, is expected to be negative; however, it is possible that, given the sample is restricted to individuals who have worked continuously from 2006 to 2019, women in this sample have specific characteristics that offset the negative impact of having children on their salary. For example, if the effect of having children includes reduced labor participation in the form of part-time contracts, such individuals would not be included in the sample.

Next, Table 5.2 presents the results of the regression (4.1) with the dependent variable $\log(Salary)$.

[5] $\log(1/N_{t,j,p'})$

Dependent Variable:	log($Salary$)			
	OLS	IV log($1/N_{p'}$)	IV log(HHI_{t-1})	2 IVs
Variables				
log(HHI)	-0.0136**	-0.0193***	-0.0136*	-0.0137**
	(0.0065)	(0.0045)	(0.0069)	(0.0067)
Female	-0.2498***	-0.2498***	-0.2493***	-0.2493***
	(0.0096)	(0.0096)	(0.0099)	(0.0099)
Children	0.0737***	0.0737***	0.0728***	0.0728***
	(0.0069)	(0.0069)	(0.0069)	(0.0069)
Experience	0.0168***	0.0168***	0.0178***	0.0178***
	(0.0013)	(0.0013)	(0.0014)	(0.0014)
Experience^2	-0.0002***	-0.0002***	-0.0002***	-0.0002***
	($2{,}3 \times 10^{-5}$)	($2{,}28 \times 10^{-5}$)	($2{,}39 \times 10^{-5}$)	($2{,}39 \times 10^{-5}$)
Intermediate Education	0.1600***	0.1600***	0.1574***	0.1574***
	(0.0155)	(0.0155)	(0.0154)	(0.0154)
Higher Education	0.5274***	0.5272***	0.5280***	0.5280***
	(0.0424)	(0.0423)	(0.0427)	(0.0427)
Female × Children	-0.0421***	-0.0421***	-0.0414***	-0.0414***
	(0.0068)	(0.0068)	(0.0068)	(0.0068)
Fixed Effects				
Province	✓	✓	✓	✓
Economic Activity	✓	✓	✓	✓
Year	✓	✓	✓	✓
Fit Statistics				
Observations	430,576	430,576	399,362	399,362
R^2	0.41075	0.41072	0.40544	0.40544
Adjusted R^2	0.17134	0.17130	0.17014	0.17014

Standard errors clustered at the Province level in parentheses
*Significance Codes: ***: 0.01, **: 0.05, *: 0.1*

Tabla 5.2
Marginal Effects Results

5.5. Model 2

The results of the model Equation 4.2 with the interaction term log(HHI) × $Female$, presented below, do not include all the control variables due to the similarity of these coefficients with those from the first model. Fixed effects at the province, economic activity, and year levels have been used again; the complete model is attached in the appendix (Table 6.3) along with the first-stage results of the models.

Dependent Variable:	$\log(Salary)$			
	OLS	IV $\log(1/N_{t,j,p'})$	IV $\log(HHI_{t-1,j,p})$	2 IV
Variables				
$\log(HHI)$	-0.0121	-0.0258***	-0.0118	-0.0120
	(0.0079)	(0.0052)	(0.0083)	(0.0081)
Female	-0.2657***	-0.1425**	-0.2673***	-0.2665***
	(0.0350)	(0.0569)	(0.0360)	(0.0358)
$\log(HHI) \times Female$	-0.0045	0.0306**	-0.0052	-0.0049
	(0.0099)	(0.0147)	(0.0102)	(0.0102)

Standard errors clustered at Province level in parentheses
*Significance Codes: ***: 0.01, **: 0.05, *: 0.1*

Tabla 5.3
Results for Equation 4.2

The coefficients β_1 and β_3 for the variables $\log(HHI)$ and $\log(HHI) \times Mujer$, respectively, are not significant at the 10 % level in all models using the lagged HHI as an instrument. Additionally, as in Model 1, they are very similar to each other. For the model that uses only $\log(1/N)$ and $\log(1/N) \times Mujer$ as instruments, it is found that the effects of labor concentration on $\log(Salario)$ are significant; negative for men, with $\beta_1 = -2{,}54\,\%$, and positive for women, with $\beta_1 + \beta_3 = 0{,}48\,\%$. Furthermore, the difference in the effect of labor concentration between women and men, that is, the coefficient of the interaction variable between $\log(HHI)$ and $Mujer$, is positive and significant at the 5 % level, with $\beta_3 = 3{,}06\,\%$.

This last result is contrary to the theory presented in **??** and to the empirical findings of several studies on labor concentration in other countries. For example, Bassanini et al., 2022, using a similar strategy, estimates the interaction parameter for several European countries. In Denmark, there was no difference in the effect of concentration by gender; in Portugal and France, the parameter was negative; and only in Germany was it positive. However, for all countries, $\beta_1 + \beta_3$ is negative. The following section discusses what this might be due to.

6. CONCLUSION

In this work, I have developed a theoretical model that explains the differences in wages between groups with different elasticities and its relationship with labor concentration. Subsequently, I have estimated this model using data from the Continuous Sample of Working Lives. The results confirm the negative relationship between concentration and wages; a 1 % increase in the HHI index decreases wages on average by -1.93 %. However, the results of the interaction $\log(HHI) \times Mujer$ do not align with the expected outcomes of the proposed theoretical framework.

There could be several reasons for this discrepancy. First, it may be due to the data source. The MCVL is a random sample of the population of workers and retirees, not of the population of firms.[6] This implies that the probability of workers coming from large firms is higher than if it were a random sample of firms. This type of bias is different from that described in **??**, in that the latter affects the concentration index while the former affects the composition of the studied worker sample.

Data filtering has also changed the composition of the sample. Only private sector workers who work full-time and who have been continuously employed from 2006 to 2019 are studied. Individuals who could be more affected by monopsony, such as part-time workers, are not included in the sample. Moreover, the effects of monopsony on women might be more pronounced if the separation elasticity were measured with a dynamic model similar to the one proposed by Burdett y Mortensen, 1998. Finally, another conjecture is that the lower labor participation in Spain compared to other countries analyzed in previous studies might lead to a higher elasticity of female labor supply, especially on the extensive margin (i.e., participating versus not participating in the labor market), which could imply a lower degree of monopsonistic exploitation of women (see e.g., Couceiro de León y Dolado, 2023).

A possible direction for future work would be to use a dynamic model and data from the Firms and Workers Panel (PET) provided by the Spanish Social Security. This panel is a random sample of firms that allows for the identification of all workers within each firm, showing their employment histories. Additionally, it would be interesting to find new valid instruments with sufficient variability.

In summary, this work identifies the negative relationship between labor concentration and wages and demonstrates the difficulty in measuring the causal effects of concentration across different demographic groups.

[6] An attempt was made to use the Firms and Workers Panel with information from both sides of the labor market, but this data source could not be obtained due to bureaucratic obstacles from the General Treasury of Social Security.

BIBLIOGRAFÍA

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. (2017). When Should You Adjust Standard Errors for Clustering? (24003).

Arellano, M. (1987). Computing Robust Standard Errors for Within-Groups Estimators. *Oxford Bulletin of Economics and Statistics*, *49*(4), 431-434.

Azar, J., Marinescu, I., & Steinbaum, M. (2019). Measuring Labor Market Power Two Ways. *AEA Papers and Proceedings*, *109*, 317-321.

Azar, J., Marinescu, I., & Steinbaum, M. (2022). Labor Market Concentration. *Journal of Human Resources*, *57*(S), 167-199.

Barth, E., & Dale-Olsen, H. (2009). Monopsonistic discrimination, worker turnover, and the gender wage gap. *Labour Economics*, *16*(5), 589-597.

Bassanini, A., Araki, S., Green, A., Marcolin, L., & Volpin, C. (2022). Labor Market Concentration and Competition Policy across the Atlantic. (15641).

Boal, W. M., & Ransom, M. R. (1997). Monopsony in the Labor Market. *Journal of Economic Literature*, *35*(1), 86-112. http://www.jstor.org/stable/2729694

Burdett, K., & Mortensen, D. T. (1998). Wage Differentials, Employer Size, and Unemployment. *International Economic Review*, *39*, 257-273.

Card, D., & Krueger, A. B. (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *American Economic Review*, *84*(4), 772-793.

Couceiro de León, A., & Dolado, J. J. (2023). Revisiting the Public-Private Wage Gap in Spain: New Evidence and Interpretation [Special Issue in Honour of Manuel Arellano]. *SERIEs Journal of the Spanish Economic Association*, *14*(3-4), 353-377.

De la Roca, J., & Puga, D. (2017). Learning by working in big cities. *Review of Economic Studies*, *84*(1), 106-142.

Hirsch, B., Schank, T., & Schnabel, C. (2006). Gender Differences in Labor Supply to Monopsonistic Firms: An Empirical Analysis Using Linked Employer-Employee Data from Germany. (47).

Langella, M., & Manning, A. (2021). Marshall Lecture 2020: The Measure of Monopsony [Monopsony in the UK]. *Journal of the European Economic Association*, *19*(6), 2929-2957.

Luccioletti, C. (2022). Labor Market Power Across Cities. (wp2022_2214).

Manning, A. (2003). *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton University Press. https://doi.org/10.2307/j.ctt5hhpvk

Ransom, M. R., & Oaxaca, R. L. (2010). New Market Power Models and Sex Differences in Pay. *Journal of Labor Economics*, *28*(2), 267-289.

Robinson, J. (1933). *The Economics of Imperfect Competition*. Macmillan.

Tenreyro, S., Abel, W., & Thwaites, G. (2018). Monopsony in the UK. (13265).

APPENDIX

Data

HHI

Year	HHI Number
2006	3845
2007	3867
2008	3862
2009	3892
2010	3881
2011	3872
2012	3859
2013	3858
2014	3865
2015	3873
2016	3875
2017	3869
2018	3893
2019	3882

Tabla 6.1
HHI values per year (2006-2019)

Figura 6.1
Distribution of concentration by industry

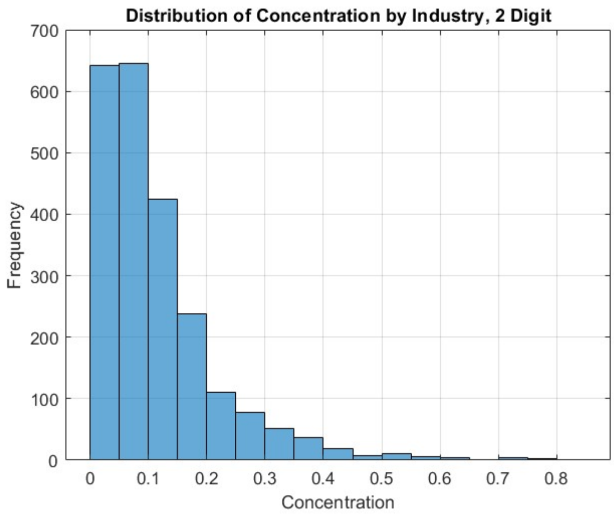


Figura 6.2

Average HHI by province

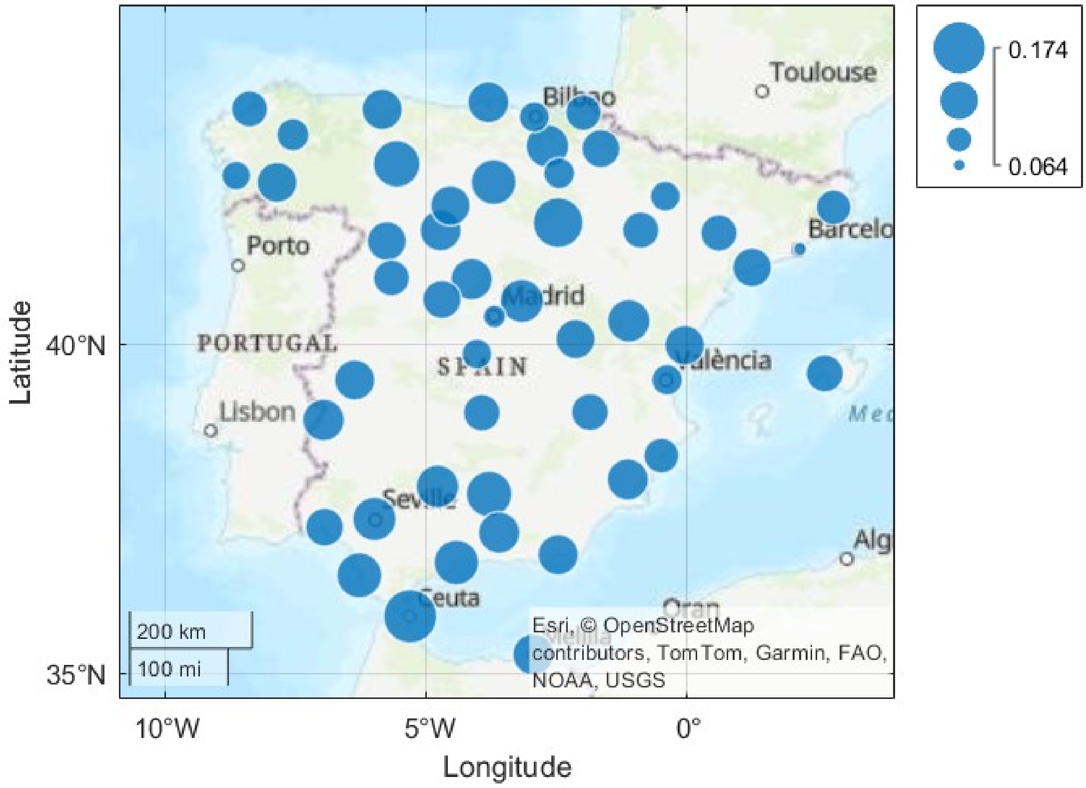


Figura 6.3

Average HHI by industry

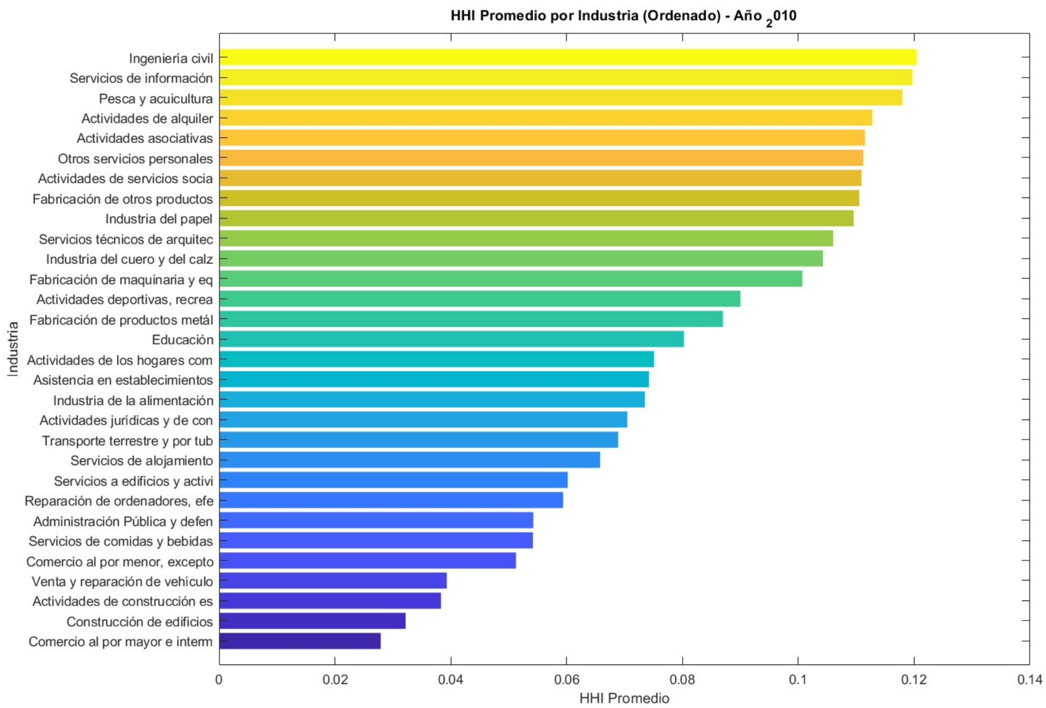


Figura 6.4

Evolution of the average HHI

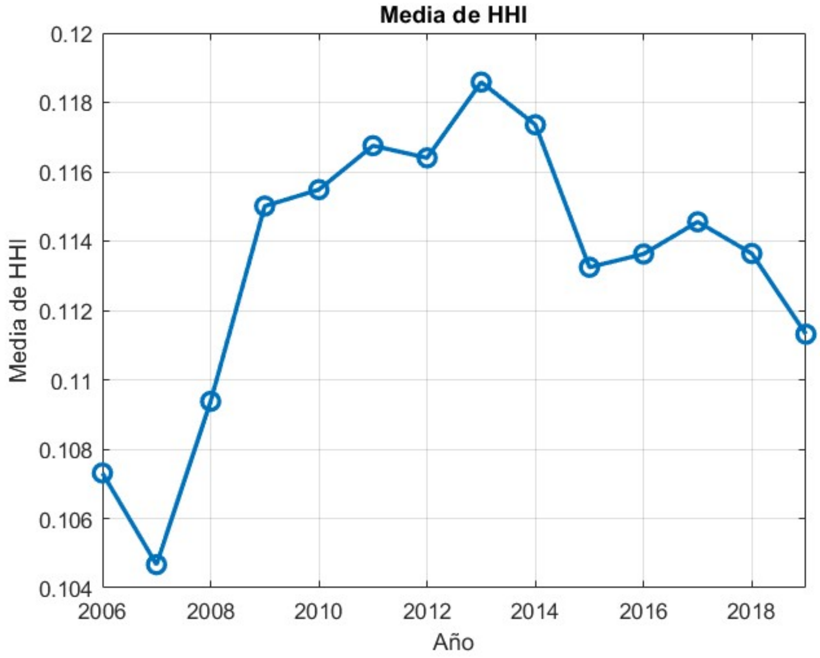


Figura 6.5

Evolution of the HHI Standard Deviation

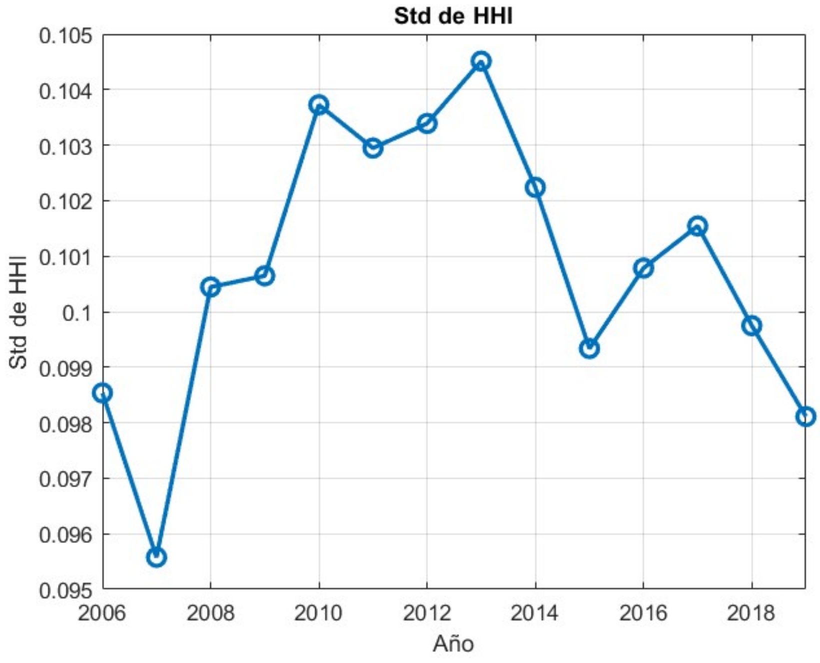


Figura 6.6

Monte Carlo

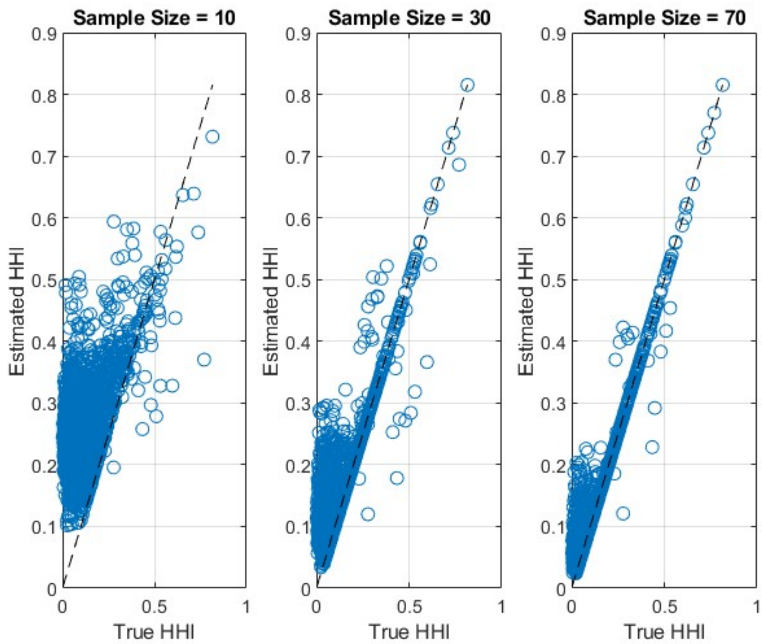
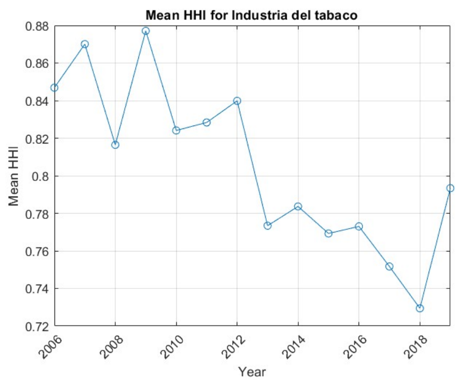
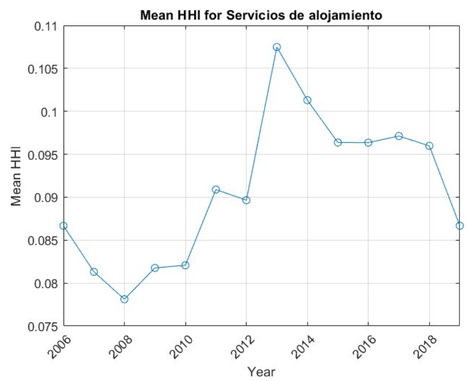


Figura 6.7

Comparison of Average HHI for different industries



(a) *Average HHI for the Tobacco Industry*



(b) *Average HHI for Accommodation Services*

Results

Dependent Variable:	log($Salary$)		
	(1)	(2)	(3)
Variables			
log(N)	0.7960***		0.2101***
	(0.0567)		(0.0312)
log(HHI_{t-1})		0.8789***	0.8607***
		(0.0150)	(0.0207)
Female	-0.0029	0.0009	0.0004
	(0.0084)	(0.0013)	(0.0013)
Children	-0.0020	-0.0005	-0.0009
	(0.0026)	(0.0005)	(0.0005)
Experience	0.0014	0.0003	-0.0001
	(0.0011)	(0.0003)	(0.0002)
Experience^2	$-4{,}38 \times 10^{-5}$**	$-7{,}72 \times 10^{-6}$	$-6{,}17 \times 10^{-9}$
	$(2{,}03 \times 10^{-5})$	$(5{,}13 \times 10^{-6})$	$(4{,}31 \times 10^{-6})$
Secondary Education	0.0062	-0.0012	-0.0013
	(0.0154)	(0.0021)	(0.0024)
Higher Education	-0.0448***	-0.0067***	-0.0075***
	(0.0111)	(0.0022)	(0.0026)
Female × CHILDREN	-0.0056	-0.0013*	-0.0009
	(0.0044)	(0.0008)	(0.0008)
Fixed Effects			
Province	✓	✓	✓
Economic Activity	✓	✓	✓
Year	✓	✓	✓
Goodness-of-fit Statistics			
Observations	430,576	399,362	399,362
R^2	0.79498	0.96100	0.96231
Adjusted R^2	0.09995	0.82512	0.83098

Standard errors clustered at the Province level in parentheses
*Significance Codes: ***: 0.01, **: 0.05, *: 0.1*

Tabla 6.2

First stages of the regressions with Equation 4.1

Dependent Variable:	log($Salary$)			
	OLS	IV log($1/N_{t,j,p'}$)	IV log($HHI_{t-1,j,p}$)	2 IV
Variables				
log(HHI)	-0.0121	-0.0258***	-0.0118	-0.0120
	(0.0079)	(0.0052)	(0.0083)	(0.0081)
Female	-0.2657***	-0.1425**	-0.2673***	-0.2665***
	(0.0350)	(0.0569)	(0.0360)	(0.0358)
log(HHI)× Female	-0.0045	0.0306**	-0.0052	-0.0049
	(0.0099)	(0.0147)	(0.0102)	(0.0102)
Children	0.0737***	0.0735***	0.0728***	0.0728***
	(0.0069)	(0.0069)	(0.0069)	(0.0069)
Experience	0.0168***	0.0168***	0.0178***	0.0178***
	(0.0013)	(0.0013)	(0.0014)	(0.0014)
Experience^2	-0.0002***	-0.0002***	-0.0002***	-0.0002***
	($2{,}31 \times 10^{-5}$)	($2{,}29 \times 10^{-5}$)	($2{,}4 \times 10^{-5}$)	($2{,}4 \times 10^{-5}$)
Secondary Education	0.1599***	0.1604***	0.1574***	0.1574***
	(0.0155)	(0.0158)	(0.0154)	(0.0154)
Higher Education	0.5274***	0.5278***	0.5279***	0.5279***
	(0.0425)	(0.0423)	(0.0428)	(0.0428)
Female × Children	-0.0421***	-0.0419***	-0.0414***	-0.0414***
	(0.0067)	(0.0067)	(0.0068)	(0.0068)
Fixed Effects				
Province	✓	✓	✓	✓
Economic Activity	✓	✓	✓	✓
Year	✓	✓	✓	✓
Goodness-of-fit Statistics				
Observations	430,576	430,576	399,362	399,362
R^2	0.41076	0.40975	0.40545	0.40545
Adjusted R^2	0.17137	0.16994	0.17017	0.17017

Standard errors clustered at the Province level in parentheses
*Significance Codes: ***: 0.01, **: 0.05, *: 0.1*

Tabla 6.3

Results of Marginal Effects for Equation 4.2 (Complete)

Dependent Variable:	$\log(HHI)$	$\log(HHI) \times Female$
Model:	(1)	(2)
Variables		
$\log(1/N)$	0.7939^{***}	0.1036^{***}
	(0.0563)	(0.0085)
$\log(1/N) \times Female$	0.0099^{*}	0.3009^{***}
	(0.0056)	(0.0248)
Female	0.0368^{*}	-2.303^{***}
	(0.0219)	(0.1808)
Children	-0.0019	0.0116^{**}
	(0.0026)	(0.0045)
Experience	0.0014	-0.0027^{*}
	(0.0011)	(0.0015)
Experience ^{2}	$-4{,}32 \times 10^{-5**}$	$2{,}18 \times 10^{-5}$
	$(2{,}02 \times 10^{-5})$	$(3{,}36 \times 10^{-5})$
Secondary Education	0.0063	-0.0094
	(0.0155)	(0.0115)
Higher Education	-0.0449^{***}	-0.0320^{**}
	(0.0110)	(0.0129)
Female × Children	-0.0060	-0.0214
	(0.0045)	(0.0141)
Fixed Effects		
Province	✓	✓
Economic Activity	✓	✓
Year	✓	✓
Goodness-of-fit Statistics		
Observations	430,576	430,576
R^2	0.79500	0.90824
Adjusted R^2	0.10005	0.88923

Standard errors clustered at the Province level in parentheses
*Significance Codes: ***: 0.01, **: 0.05, *: 0.1*

Tabla 6.4

First Stage, $\log(1/N)$ *as Instrument*

Dependent Variables:	$\log(HHI)$	$\log(HHI) \times Female$
Model:	(1)	(2)
Variables		
$\log(HHI_{t-1})$	0.8772***	-0.0323***
	(0.0152)	(0.0053)
$\log(HHI_{t-1}) \times Female$	0.0049***	0.9795***
	(0.0011)	(0.0035)
Female	0.0180***	-0.0654***
	(0.0043)	(0.0103)
Children	-0.0005	0.0001
	(0.0005)	(0.0003)
Experience	0.0003	$-6{,}03 \times 10^{-5}$
	(0.0003)	(0.0001)
Experience^2	$-7{,}84 \times 10^{-6}$	$-7{,}34 \times 10^{-8}$
	$(5{,}25 \times 10^{-6})$	$(2{,}73 \times 10^{-6})$
Secondary Education	-0.0011	0.0005
	(0.0021)	(0.0010)
Higher Education	-0.0066***	-0.0004
	(0.0023)	(0.0011)
Female × Children	-0.0013*	-0.0004
	(0.0007)	(0.0007)
Fixed Effects		
Province	✓	✓
Economic Activity	✓	✓
Year	✓	✓
Goodness-of-fit Statistics		
Observations	399,362	399,362
R^2	0.96101	0.99374
Adjusted R^2	0.82514	0.99243

Standard errors clustered at the Province level in parentheses
*Significance Codes: ***: 0.01, **: 0.05, *: 0.1*

Tabla 6.5

First Stage, $\log(HHI_{t-1})$ *as Instrument*

Dependent Variables:	log(HHI)	log(HHI) $\times$ $Female$
Model:	(1)	(2)
Variables		
log($1/N$)	0.2107***	0.0499***
	(0.0316)	(0.0056)
log($1/N$) $\times$ $Female$	0.0002	0.0158***
	(0.0014)	(0.0017)
log(HHI_{t-1})	0.8580***	-0.0355***
	(0.0212)	(0.0063)
log(HHI_{t-1}) $\times$ $Female$	0.0076***	0.9748***
	(0.0022)	(0.0037)
Female	0.0279***	-0.0190**
	(0.0050)	(0.0078)
Children	-0.0009*	0.0003
	(0.0005)	(0.0003)
Experience	-0.0001	-0.0002*
	(0.0002)	(0.0001)
Experience^2	$-1{,}51 \times 10^{-7}$	$2{,}99 \times 10^{-6}$
	$(4{,}26 \times 10^{-6})$	$(3{,}19 \times 10^{-6})$
Secondary Education	-0.0012	0.0006
	(0.0024)	(0.0011)
Higher Education	-0.0074***	-0.0009
	(0.0026)	(0.0011)
Female $\times$ Children	-0.0008	-0.0010
	(0.0008)	(0.0007)
Fixed Effects		
Province	✓	✓
Economic Activity	✓	✓
Year	✓	✓
Goodness-of-fit Statistics		
Observations	399,362	399,362
R^2	0.96232	0.99380
Adjusted R^2	0.83103	0.99250

Standard errors clustered at the Province level in parentheses
*Significance Codes: ***: 0.01, **: 0.05, *: 0.1*

Tabla 6.6

First Stage, Both Instruments