MARCO A. BADILLA MAROTO
University of Costa Rica and University of California, Berkeley, USA

The Role of the Gender Wage Gap in Overall Wage Inequality: A Quantitative Exercise

ABSTRACT This article presents a novel wage inequality decomposition to analyze the impact of the gender wage gap on overall wage inequality. The decomposition determines the maximum relative wage between genders allowed before it begins to increase total inequality. In addition, I present a structural model of the labor market to evaluate the impact of establishing restrictions on intra-occupational gender pay gaps within each firm. Specifically, I introduce a restriction in which the average wage of one gender cannot exceed α times the average wage of the other gender. For $\alpha = 2$, the model predicts a 10 percent wage inequality reduction. However, with a tighter restriction of $\alpha = 1$, the inequality reduction dissipates and reverses into a wage inequality increase.

JEL Codes: J16, J31, D31, E64

Keywords: Gender, inequality, restriction, occupation

Despite some improvement over the past several decades, gender wage gaps persist in virtually all countries (Blau and Kahn, 2017; Hegewisch and Williams-Baron, 2016; Miller and Vaggins, 2018). According to the Organization for Economic Cooperation and Development (OECD), most of the member countries currently present a substantial gender wage gap.[1] Furthermore, these gaps are becoming increasingly difficult to explain (Blau and Kahn, 2017; Brynin, 2017). Goldin (2014) states that the explained portion of the gender wage gap decreased over time as human capital investments between men and women converged. To counter the persistence of the gender

ACKNOWLEDGMENTS Special thanks go to José Pablo Vásquez, Isabela Manelici, and Alonso Alfaro for all their mentoring and support while developing this paper. I also thank the editor, an anonymous referee, José Ignacio Gonzalez, Valeria Lentini, Milagro Saborío, Laura Blanco, and Luis Hall for their valuable comments and suggestions throughout the entire process. Finally, I acknowledge support from the Economic Sciences Research Institute (IICE) at the University of Costa Rica in providing the employer-employee administrative data set for Costa Rica's formal labor sector.

1. OECD, Gender wage gap (indicator) (doi: 10.1787/7cee77aa-en, accessed March 27, 2021).

wage gap, several direct approaches have been proposed or implemented, largely based on legal rights and compulsory regulation. Examples include establishing gender quotas (Bertrand and others, 2019), declaring pay secrecy illegal (Kim, 2015), banning questions on past earnings (Hansen and McNichols, 2020), introducing transparency laws (Baker and others, 2019), and publishing firms' gender wage gaps (Coghlan and Hinkley, 2018).

This paper explores the impacts of enforcing an intra-occupational gender wage gap limit within firms on overall wage inequality and on gender wage gaps in an upper level (occupations in the economy). Such direct restrictions must be strategically designed to counter the gap's persistence. Blau and Kahn (2017) find that occupation and industry are the most important variables in explaining the gender wage gap, despite the occupational upgrading of women relative to men. Moreover, Goldin (2014) adds that most of the current gender earnings gap comes from differences within occupations rather than between occupations. Card, Cardoso, and Kline (2016) point out that firm-specific pay policies may also be important for understanding the gender wage gap. Accordingly, I consider direct restrictions on intra-occupational gender wage gaps within each firm.

I evaluate the impact of such restrictions on inequality and occupational gender wage gaps in the economy for three specific years, 2008, 2013, and 2018, using a rich employer-employee administrative data set covering Costa Rica's formal employment sector. The effect on total inequality is especially important not only because Costa Rica has one of the highest Gini coefficients among OECD countries (OECD, 2020) but also because reducing the gender wage gap has played an important role in mitigating the rise of inequality in other countries, as Piketty, Saez, and Zucman (2018) find for the United States since the late 1960s.

The objective is achieved in three steps. First, I develop a novel wage inequality decomposition by gender, using wage dispersion as the measure of inequality. This mathematically materializes the link between gender wage gaps and total inequality that is empirically emphasized by Piketty, Saez, and Zucman (2018). Furthermore, it produces a candidate gender wage gap restriction, since it determines the maximum relative wage between genders that can exist without causing total inequality to increase. An advantage of using variance as the definition of inequality is that its decomposition allows quantifying the relative importance of possible sources of wage differences (Helpman and others, 2016). More specifically, this definition, together with the decomposition, is able to provide a clear mathematical expression of the gender wage gap's impact on the overall measure of inequality, buttressing previous motivations to incorporate a restriction on gender wage gaps. These connections and argumentation are not guaranteed to arise in an analytical form with other wage inequality measures. Moreover, as Magda, Gromadzki, and Moriconi (2019) state, the wage variance "is a common statistical measure of dispersion, and, unlike other popular measures of inequality such as the Gini coefficient and the 90–10 wage gap, the variance is additively decomposable into the between-firm component and within-firm component" (p. 8), which is important when comparing gender wages. Finally, there is a base literature that already uses variance decomposition to examine wage inequality (Bonhomme, Lamadon, and Manresa, 2019; Card and others, 2018; Goldschmidt and Schmieder, 2017); my new decomposition expands this literature in that it not only presents a relation between total inequality and gender wage gaps, it also incorporates within and between components for several characteristics simultaneously.

In this new decomposition, wage inequality can be broken down into gender-related components to control for several characteristics of the economy, namely, economic sector, type of firm (public or private), firm, and occupation. These components account for wage differences within the same gender across each controlling characteristic of the economy, as well as wage gaps between genders within each occupation for each firm. Furthermore, the gender wage gap component that arises from the decomposition sets the foundation for the gender wage gap restriction (α) since it establishes that inequality increases precisely when one gender's average wage within an occupation of a specific firm is more than two times the other gender's wage; therefore, the decomposition gives a restriction of $\alpha = 2$. Under this criterion, around 40 percent of all firms' departments (occupations within each firm) do not fulfill the gender wage gap restriction.

The gender-based wage decomposition shows that the growth in inequality between 2008 and 2018 is mainly explained by an increase in a segregation effect (30 percent of the total increase in inequality), by a rise in the wage differences within men in the same occupation of each firm (26 percent of the total increase), and by a similar rise for women (around 19 percent of the total increase). These past components not only account for most of the inequality growth, and therefore point to specific areas in which to focus corrective action, they also act as transmitters of effect of the gender wage gap restriction on total inequality. This transmission occurs because the restriction is imposed on wage averages, and when firms make decisions to comply with specific relationships of averages, it affects individual wages and gender workplace composition, which are the basis for the other components.

In the second step, I carry out a quantitative exercise using a novel heterogeneous-agent labor model. I use the model, together with the Costa Rican employer-employee data, to analyze the impacts of the gender wage gap restriction on the economy's occupational gender wage gaps and on total wage inequality. The model focuses on agents' (firms' and workers') reactions to exogenous and unexpected restrictions on each firm's intra-occupational gender wage gaps. The first main stage of the model concerns how each department (occupations within each firm) endogenously determines job growth based on three factors: its sales expectations, compliance with the restriction, and avoidance of higher labor costs. The second main stage incorporates the pressure on wages that the gender-occupation-differentiated labor markets generate through supply and demand and the matching of workers and employers through a series of characteristics, such as wage, gender, occupation, location, and experience. In general, the model focuses on how firms adapt to the restriction in a shocked labor market and how the pairing between agents in this market is readjusted owing to this distortion. In addition to heterogeneous (in wage) firms and workers, the model borrows search-and-matching intuition. However, the model differs from the common heterogeneous-agent and search-and-matching models (see, for example, Taber and Vejlin, 2020) in that it focuses on each firm's coping process and its incentives to deviate from plans when making its final decision on how to comply with the specific restriction; this results in a particular model with a numerical, but not an analytical, solution.

In the third step, the economy's wages that were introduced into the model and the model results are used to compute the change in the economy's occupational gender wage gaps and the components of the decomposition. The quantitative exercise shows that occupational gaps tend to deteriorate (move away from the equalization of averages) in favor of men for those occupations with little female participation, and they tend to deteriorate in favor of women if the occupation has low male participation. Some examples of the former occupations are mechanic, builder, and transporter; the latter include preschool and special education, domestic worker, beauty services, and nutritionist. The model also shows that the constraint produces a decrease in total inequality of 10 percent. This reduction is mainly achieved through lower wage differences among workers of the same gender within the same department. However, if the restriction is tightened too much ($\alpha = 1$), this overall inequality reduction dissipates and reverts to an increase.

This article relates mainly to two lines of research. First, I introduce a novel wage variance decomposition into the within-between inequality decomposition literature (Álvarez and others, 2018; Barth and others, 2016; Helpman and others, 2016; Song and others 2018). This new decomposition allows the within-between intuition to be present for several aspects simultaneously, such as sector, type of firm (public or private), firm, occupation, and gender. More specifically, the decomposition isolates a single component that quantifies the impact of gender wage gaps and occupational segregation on total wage inequality. This further elucidates the empirical findings of Piketty, Saez, and Zucman (2018) regarding the role of gender wage gaps in inequality. The computation of the decomposition using Costa Rican data shows that having intra-occupational gender wage gaps within firms under an $\alpha = 2$ limit has contributed to slowing down inequality (without any legal restriction being imposed). However, because there are still many cases in which this limit is violated, there is ample opportunity to reduce inequality by narrowing gender wage gaps.

Second, the study relates to the literature on the persistence of gender wage gaps (Bertrand, Goldin, and Katz, 2010; Blau and Kahn, 2017; Cook and others, 2021) and mechanisms to reduce it. With regard to the latter, Bertrand and others (2019) document that the gender wage gap within boards fell substantially with the introduction of gender quotas. Kim (2015) compares wages in six U.S. states that declared wage secrecy illegal before 2012 and states that did not; she finds that in the former, women's wages were 3 percent higher and the gender wage gap was reduced by 12 to 15 percent for women with a university degree. Finally, Baker and others (2019) argue that transparency laws have the potential to reduce the gender wage gap by approximately 30 percent. To contribute to this line of research, this article introduces a within-firm intra-occupational gender wage gap restriction that naturally arises from the novel wage inequality decomposition; more specifically, the decomposition isolates a gender wage gap component that determines the maximum relative wage between genders that can exist without increasing wage inequality. The paper does not argue that the direct restriction on gender average wage differences is more efficient than other legal regulations, such as gender quotas, but rather demonstrates that total inequality can be reduced if gender wage gaps are controlled at the intra-occupational within-firm level. Even so, occupational gender wage gaps may deteriorate (away from average equalization) at the economy level for occupations with a high gender concentration.

The rest of the article is organized as follows. After presenting the employer-employee administrative data set used, I explain the decomposition of wage inequality by gender and outline the quantitative exercise. I then present the results and close with some concluding remarks.

Administrative Data

I use the employer-employee administrative data set from the Costa Rican Social Security Fund (CCSS).[2] Observations in the database correspond to all people who contribute to social security, from either public or private firms. While the base is not a sample but a complete set of workers and firms, it excludes the informal sector, which represents approximately 30 percent of all jobs (OECD, 2017). Results are thus representative only of the country's formal sector.

The database contains information about the characteristics of workers and their employers. For workers, it has information about their salary (real January 2008 wages), gender, number of monthly contributions to the CCSS, date on which they started working for the current firm, occupation (in detail, based on the Costa Rican Occupation Classification, 2011 version), and respective employer (if any, as there are independent workers in the data set). For employers, the following information is available: location (province, canton, and district), number of workers, economic sector (in detail, based on the International Standard Industrial Classification of All Economic Activities, revision 3.1), and type of firm (public or private). The database has a monthly frequency and is available for 2008, 2013, and 2018.[3] I used complementary information from labor legislation to calibrate certain aspects of the labor market in the quantitative exercise, namely, the dynamics corresponding to severance pay and annual bonuses.

The final database used contains information on 857,729 workers and 21,279 firms for 2008, 1,009,630 workers and 24,687 firms for 2013, and 1,225,068 workers and 27,082 firms for 2018. The percentage of workers in the private sector ranges from 72 percent to 75 percent, depending on the year. Economic sectors are classified into twenty categories and occupations into forty-three classes.[4]

2. I obtained the database from the Economic Sciences Research Institute (IICE) at the University of Costa Rica (UCR), and its use is subject to authorization and confidentiality agreements with the IICE. The identity of workers and firms is protected through the use of identification codes created by the Costa Rican Social Security Fund.

3. For details on the cleaning process, see online appendix A. Supplementary material for this paper is available online at http://economia.lacea.org/contents.htm.

4. For more information on these classifications, see tables 5 and 6 in online appendix A. Table 5 includes descriptive statistics on average wages, gender wage gaps, and labor force gender composition by economic sector.

Wage Inequality Decomposition by Gender

There are many ways to measure wage inequality in an economy, including the distribution of wages, percentile curves, the ratio of salaries between specific percentiles of the labor force (for example, 50–10, 90–10, or 90–50), and the Gini coefficient (Magda, Gromadzki, and Moriconi, 2019; Sarlo, Clemens, and Emes, 2015). As discussed in the introduction, I use the dispersion of wages as the measure of inequality in this research. The variance of the wages in an economy with N workers in a specific period t is denoted as **T** and is defined as the average of squared deviations of each wage (ω_i) with respect to the average wage ($\bar{\omega}$); **T** describes how different or unequal workers' wages are from each other. Wage inequality is defined as

$$(1) \qquad \mathbf{T}(\omega) \equiv \frac{1}{N}\sum_{i=1}^{N}\left(\omega_i - \bar{\omega}\right)^2.$$

In the first phase of the decomposition of the variance of wages, total inequality is broken down into five parts to take into account the structure of the labor market (see online appendix B):

$$(2) \quad \mathbf{T}(\omega) = \mathbf{B}(s) + \mathbf{B}(T,s) + \mathbf{B}(f,T,s) + \mathbf{B}(\vartheta,f,T,s) + \mathbf{W}(i,\vartheta,f,T,s).$$

The **B**(·) components from the above decomposition show how different wages are between (1) sectors, s; (2) types of firm (public or private), T, controlling for sector s, (3) individual firms, f, controlling for type T and sector s, and (4) occupations, ϑ, in a given firm f of type T and in sector s. For example, the components of inequality between sectors and between types of firms are as follows:

$$(3) \qquad \mathbf{B}(s) \equiv \frac{1}{N}\sum_{s\in S} N_s\left(\bar{\omega}_s - \bar{\omega}\right)^2.$$

$$(4) \qquad \mathbf{B}(T,s) \equiv \frac{1}{N}\sum_{s\in S}\sum_{T\in s} N_{T,s}\left(\bar{\omega}_{T,s} - \bar{\omega}_s\right)^2.$$

The last component of the decomposition shows how unequal the workers are within a given occupation in a firm f, for each occupation of each firm:

$$(5) \qquad \mathbf{W}(i,\vartheta,f,T,s) \equiv \frac{1}{N}\underbrace{\sum_{s\in S}\sum_{T\in s}\sum_{f\in T}\sum_{\vartheta\in f}}_{\sum_{s,T,f,\vartheta}}\sum_{i\in\vartheta}\left(\omega_i - \bar{\omega}_{\vartheta,f,T,s}\right)^2.$$

From equation 2, the five components are broken down using the gender variable:[5]

$$\text{(6)} \qquad \mathbf{T}(\omega) = \mathbf{G}_M + \mathbf{G}_F + \mathbf{W}_M + \mathbf{W}_F + \mathbf{P}_M + \mathbf{P}_F + \mathbf{B}(s)_M + \mathbf{B}(s)_I + \mathbf{B}(s)_F + \mathbf{B}(T,s)_M + \mathbf{B}(T,s)_I + \mathbf{B}(T,s)_F + \mathbf{B}(f,T,s)_M + \mathbf{B}(f,T,s)_I + \mathbf{B}(f,T,s)_F + \mathbf{B}(\vartheta,f,T,s)_M + \mathbf{B}(\vartheta,f,T,s)_I + \mathbf{B}(\vartheta,f,T,s)_F.$$

In general, each $\mathbf{B}(\cdot)$ inequality component is broken down into three subcomponents; equation 7 presents an example with $\mathbf{B}(s)$. The first subcomponent, $\mathbf{B}(s)_M$, reflects (1) wage inequality between men from different sectors and (2) penalization (inequality increment) due to differences in gender composition (δ) between sectors (for example, $\bar{\omega}_{s,M} = \bar{\omega}_M$). The second subcomponent, $\mathbf{B}(s)_F$, for women is analogous to the first. Finally, the third subcomponent, $\mathbf{B}(s)_I$, penalizes the correlation between $\mathbf{B}(s)_M$ and $\mathbf{B}(s)_F$ for a given sector s; that is, if both men and women in sector s are simultaneously above or below the economy's average, in terms of wages and composition, it is taken as an inequality relative to the other sectors.

$$\mathbf{B}(s) = \underbrace{\sum_{s\in S}\frac{N_s}{N}\left(\delta_{s,M}\bar{\omega}_{s,M} - \delta_M\bar{\omega}_M\right)^2}_{\mathbf{B}(s)_M} + \underbrace{\sum_{s\in S}\frac{N_s}{N}2\left(\delta_{s,M}\bar{\omega}_{s,M} - \delta_M\bar{\omega}_M\right)\left(\delta_{s,F}\bar{\omega}_{s,F} - \delta_F\bar{\omega}_F\right)}_{\mathbf{B}(s)_I} + \underbrace{\sum_{s\in S}\frac{N_s}{N}\left(\delta_{s,F}\bar{\omega}_{s,F} - \delta_F\bar{\omega}_F\right)^2}_{\mathbf{B}(s)_F}$$

$$\text{(7)} \qquad \mathbf{B}(s) = \mathbf{B}(s)_M + \mathbf{B}(s)_I + \mathbf{B}(s)_F.$$

Next, the intra-occupational inequality component, $\mathbf{W}(i, \vartheta, f, T, s)$, is broken down into six subcomponents. The first two, shown in equations 8 and 9,

5. Figure 6 in online appendix B presents a tree graph of the decomposition.

are the within-group wage inequality among men in a given occupation in firm $f(\mathbf{W}_M)$ and among women in a given occupation in firm $f(\mathbf{W}_F)$. The next two subcomponents, in equations 10 and 11, encompass a penalty (inequality increment) assigned to each gender for not dominating the occupation completely: the more one group dominates in terms of composition, the more influence it has over the other, effectively penalizing the minority group. For example, the within-group inequality among women in a given occupation ϑ in firm f receives a penalty the more segregated they become ($\uparrow N_{\vartheta,M} \Rightarrow \delta_{\vartheta,M} \rightarrow 1$), and this penalty disappears as women dominate the composition ($\delta_{\vartheta,M} \rightarrow 0$); the penalty depends on the number of individuals affected and is quantified based on the respective gender's average salary.

(8) $$\mathbf{W}_M \equiv \sum_{s,T,f,\vartheta} \sum_{M \in \vartheta} \sum_{i \in M} \left(\omega_i - \bar{\omega}_{\vartheta,M}\right)^2.$$

(9) $$\mathbf{W}_F \equiv \sum_{s,T,f,\vartheta} \sum_{F \in \vartheta} \sum_{i \in F} \left(\omega_i - \bar{\omega}_{\vartheta,F}\right)^2.$$

(10) $$\mathbf{P}_M \equiv \sum_{s,T,f,\vartheta} N_{\vartheta,M} \bar{\omega}^2_{\vartheta,M} \delta^2_{\vartheta,F}.$$

(11) $$\mathbf{P}_F \equiv \sum_{s,T,f,\vartheta} N_{\vartheta,F} \bar{\omega}^2_{\vartheta,F} \delta^2_{\vartheta,M}.$$

Finally, equations 12 and 13 show the subcomponents associated with intra-occupational gender wage gaps. A peculiarity of these gaps is their dual benefit-penalty character, similar to the **P** components. Within a given occupation at a firm, there are two components associated with the gaps, $\mathbf{G}_{\vartheta,M}$ and $\mathbf{G}_{\vartheta,F}$:

(12) $$\mathbf{G}_M \equiv \sum_{s,T,f,\vartheta} N_{\vartheta,M} \delta^2_{\vartheta,F} \bar{\omega}_{\vartheta,F} \left(\bar{\omega}_{\vartheta,F} - 2\bar{\omega}_{\vartheta,M}\right) = \sum_{s,T,f,\vartheta} \mathbf{G}_{\vartheta,M}.$$

(13) $$\mathbf{G}_F \equiv \sum_{s,T,f,\vartheta} N_{\vartheta,F} \delta^2_{\vartheta,M} \bar{\omega}_{\vartheta,M} \left(\bar{\omega}_{\vartheta,M} - 2\bar{\omega}_{\vartheta,F}\right) = \sum_{s,T,f,\vartheta} \mathbf{G}_{\vartheta,F}.$$

The duality of the gender wage gap components is understood as follows. Suppose the average male wage of a given occupation in firm f is ten times the female wage. This causes $\mathbf{G}_{\vartheta,M}$ to be negative, so $\mathbf{W}_{\vartheta,M}$ decreases: $\mathbf{G}_{\vartheta,M} + \mathbf{W}_{\vartheta,M} < \mathbf{W}_{\vartheta,M}$. Thus, the within-group inequality among men decreases because it is offset by the fact that, on average, they are better off in terms of wage than the other gender group, and this reduces the inequality effect for men. At the same time, women are worse off because not only is there inequality within the group but their group is, on average, below the male group: $\mathbf{G}_{\vartheta,F} + \mathbf{W}_{\vartheta,F} > \mathbf{W}_{\vartheta,F}$. This type of double impact by gender is also found in some discrimination models in the literature on gaps and wage decompositions using econometric estimation.[6]

Strict inequality between averages is not punished. Rather, inequality is punished only above a threshold—in this case, two times the other wage average. This is convenient since the decomposition does not control for all individual characteristics; for example, variables such as education, age, and experience are excluded. Similarly, an equalization of wage averages for each gender would lead to an absence of a gender wage gap in the firm's respective occupation. This positive aspect is rewarded with a decrease in inequality for each gender; that is, $\mathbf{G}_{\vartheta,G} + \mathbf{W}_{\vartheta,G} < \mathbf{W}_{\vartheta,G}$, $\forall\ G \in \{M, F\}$.

Equation 6 and all the associated equations can be used to calculate various gender-specific origins of total wage inequality, as well as the gender-specific source of the growth of total inequality over time.

Quantitative Exercise

The above inequality decomposition formalizes the link between the firm's intra-occupational gender wage gap and total wage inequality. It also establishes specific limits within which the intra-occupational gender wage gap does not penalize with more inequality. This section presents a quantitative exercise to capture the effects of implementing a gender wage gap restriction that is aligned with the limits on the firm's intra-occupational gender wage gap established by the inequality decomposition. The model is solved in three stages: (1) firms' reaction to the gender wage gap restriction, (2) workers' reaction to firms' decisions and clearing of the labor market, and (3) computation of the effects using the database and final wages produced by the model. See algorithm 1 at the end of this section for a summary of the model's algorithm.

Environment

Before discussing the firms' and workers' reactions, I provide a brief description of the labor market's environment.

6. See Oaxaca and Ransom (1994) for an example of how discrimination can be divided into the advantage of one group and the disadvantage of the other.

TIMING. The model assumes two periods of interest. In the initial period, t, all firms are informed that the gender wage gap restriction has been approved and that they are required to comply before the next period of interest, $t + 1$. I also assume that the time difference between t and $t + 1$ is short, since a longer time frame would introduce inevitable speculation regarding the timing of firms' reactions. Firms' response to the restriction, workers' response to firms' decisions, and the clearing of the labor market are the specific components of the transition between t and $t + 1$.

WORKERS. There are three types of worker, i: employees, independents, and a new labor force that comes together between t and $t + 1$. Each worker has a specific occupation ϑ and a given wage $\omega_{i,\vartheta,f,t}$ at time t and is associated with a specific firm f. In the case of independent workers, they are kept in the model as an input for the labor supply, but they are then discarded because the measures of inequality and gender wage gaps are analyzed only for direct employees in the quantitative exercise. Other worker characteristics that are observable are their gender, the number of past monthly social security contributions, and the number of months they have worked in current firm.

FIRMS AND DEPARTMENTS. Each firm f has an economic sector, s, and type, T (public or private). All these firms are heterogeneous in their payroll structure and therefore in their gender wage gaps. Also, each firm has a specific location within the country (at the canton level) and a set of occupations. Firms that are already registered in the initial period t are denoted as established, and those that arise after the initial period are denoted as new. Each occupation ϑ within firm f is defined as a department and denoted as ϑ, f.

Equation 14 computes the absolute growth (between t and $t + 1$) of total job positions in each department, $h_{\vartheta,f,t+1}$, as the change in the number of workers over time; therefore, $L_{\vartheta,f,t}$ represents the number of workers (positions, L) in each department at time t. Equation 15 defines a labor cost, $\mathbf{C}_{\vartheta,f,t+1}$, for each department in each firm at time $t + 1$. It accounts for payroll costs, $\mathbf{CP}_{\vartheta,f,t+1}$, hiring costs, $\mathbf{CH}_{\vartheta,f,t+1}$, and firing costs, $\mathbf{CF}_{\vartheta,f,t+1}$. The payroll costs (equation 16) account for the wages, $\omega_{i,\vartheta,f,t+1}$, of all workers currently employed (including newly hired workers and excluding previously fired workers) and a normalized bonus given by $\eta \cdot \omega_{i,\vartheta,f,t+1}$, where η is a frequency-normalizing constant and where one month's pay, $\omega_{i,\vartheta,f,t+1}$, is the annual bonus by law. Firing costs (equation 17), derived from workers laid off between t and $t + 1$, are made up of the portion ι of the annual bonus earned up to point, $\iota \cdot \omega_{i,\vartheta,f,t}$, and a severance pay, $SV(\omega_{i,\vartheta,f,t})$, that depends on the time worked and the wage; the definition of the severance pay and the calibration of all parameters in the model are discussed in a later subsection. Hiring costs (equation 18) comprise all costs associated with interviews, initial productivity loss, and training of newly incorporated workers between t and $t+1$. According to Blatter, Muehlemann, and Schenker (2012), average hiring costs differ substantially depending on the firm's sector, its size, and the total number of hires. Hence, the hiring cost of each new worker is calculated as a percentage $a_{\vartheta,f}(s, L_{f,t}, h_{\vartheta,f,t+1})$ of the wage that depends on those three factors.

$$h_{\vartheta,f,t+1} \equiv L_{\vartheta,f,t+1} - L_{\vartheta,f,t}. \tag{14}$$

$$C_{\vartheta,f,t+1} \equiv CP_{\vartheta,f,t+1} + CF_{\vartheta,f,t+1} + CH_{\vartheta,f,t+1}. \tag{15}$$

$$CP_{\vartheta,f,t+1} \equiv \sum\nolimits_{i\in\vartheta,f} (1+\eta)\,\omega_{i,\vartheta,f,t+1}. \tag{16}$$

$$CF_{\vartheta,f,t+1} \equiv \sum\nolimits_{i\in\vartheta,f(fired)} \left[\iota \bullet \omega_{i,\vartheta,f,t} + SV\left(\omega_{i,\vartheta,f,t}\right)\right]. \tag{17}$$

$$CH_{\vartheta,f,t+1} \equiv \sum\nolimits_{i\in\vartheta,f(hired)} a_{\vartheta,f}\left(s, L_{f,t}, h_{\vartheta,f,t+1}\right) \bullet \omega_{i,\vartheta,f,t+1}. \tag{18}$$

THE INTRA-OCCUPATIONAL GENDER WAGE GAP RESTRICTIONS IN EACH FIRM. The exogenous and unexpected restrictions imposed in the initial period establish that for each occupation within a firm, each gender's average wage cannot be greater than α times the other gender's average wage. Equation 19 presents the restriction that must be fulfilled. The restriction arises from the $\mathbf{G}_M$ and $\mathbf{G}_F$ components, which emphasize that gender wage gaps increase inequality when a threshold of two is surpassed. Thus, $\alpha > 2$ would be considered a more flexible restriction since it allows gender wage gaps over the penalization threshold, and $\alpha < 2$ would be a more stringent constraint. Also, the restriction does not apply to firms of only one person (independent) or to occupations within firms that consist of only one person, whereas it does apply to occupations with more than one worker but only one gender (occupational segregation). Suppose that an occupation in firm f has no women workers. In that case, it could be argued that there is no gender wage gap (even one of the $\mathbf{G}_{\vartheta,G}$ components would disappear since one of the gender shares would be zero). However, in this wage inequality decomposition by gender, a value of zero will be assigned to a missing gender's average wage (as if this gender receives no compensation) with the specific objective of taking into consideration that occupational segregation is one of the documented causes of gender wage gaps.[7]

$$\text{(19)} \qquad \max\left\{\bar{\omega}_{\vartheta,f,M}, \bar{\omega}_{\vartheta,f,F}\right\} \leq \alpha \min\left\{\bar{\omega}_{\vartheta,f,M}, \bar{\omega}_{\vartheta,f,F}\right\}.$$

I assume the restriction is fully enforced; that is, all firms will comply with the restriction in the allotted time, based on (1) credible sanctions and strict vigilance on behalf of the regulatory agency and (2) the threat of repercussion on the part of consumers since the intra-occupational gaps are to be made public, following Coghlan and Hinkley (2018) (publicity factor). Some other assumptions are established to guarantee that the restrictions do not cease to have merit. Specifically, I assume that firms are not able to take the following illegal actions: (1) gender falsification, since the regulatory agency can verify the employer-employee information in the national registry; (2) falsification of the worker's occupation, since the agency can corroborate it with the employee's work history and membership in the respective professional association; (3) underreporting of the employee's wage, since the employer and employee have conflicting incentives on aspects such as taxes, social security contributions, and pension contributions; and (4) falsification of a position in the firm for which social security contributions are paid but wages are not, because of personal inspections made by the regulatory agency. Last, the current legislation prohibits firms from reducing wages unilaterally.

Furthermore, each department is classified into one of five categories according to its initial fulfillment of the restriction, $R \in \{0, 1, 2, 3, 4\}$. An R of type 0 indicates that the department meets the restriction from the start. Types 1 and 2 are departments in which the restriction is not met, favoring men or women, respectively. Finally, types 3 and 4 characterize departments that have no women or men, respectively.

Table 1 shows the percentage of departments (occupations within a firm, for all firms) that at the beginning of each period are not complying with the intra-occupational gender wage gap restriction; the statistics presented differentiate by those occupations in firms that experience a segregation problem (more than one worker and only one gender). The data show that many occupations that must make adjustments to meet the restriction reflect

7. See Blau and Kahn (2017); Coghlan and Hinkley (2018); and Khitarishvili, Rodríguez-Chamussy, and Sinha (2018).

TABLE 1. Percentage of Departments That Violate Gender Wage Gap Restriction: 2008, 2013, and 2018

	With segregation			*Without segregation*		
Alpha	*2008*	*2013*	*2018*	*2008*	*2013*	*2018*
$\alpha = 1.0$	63.4	60.77	61.59	26.09	25.57	27.14
$\alpha = 1.5$	42.95	40.53	39.89	5.64	5.34	5.44
$\alpha = 2.0$	39.48	37.22	36.57	2.17	2.03	2.12
$\alpha = 2.5$	38.34	36.07	35.41	1.03	0.88	0.97
$\alpha = 3.0$	37.85	35.68	34.98	0.54	0.49	0.54

Notes: The table presents the percentage of occupations within firms that are not initially complying with the intra-occupational gender wage gap restriction in 2008, 2013, and 2018. Alpha represents the restriction threshold: for each occupation within a firm, average wage of one gender cannot be greater than α times the other. Occupations in each firm that consist of more than one worker and contain only one type of gender (occupational segregation) are also subject to the intra-occupational gender restriction; the table differentiates by this effect.

a segregation issue. In general, around 38 percent of the occupations in firms are not meeting the restriction each year (with $\alpha = 2$), and of these, 94 percent have segregation problems.

Firms' Reaction to the Restriction

I assume that the gender wage gap restrictions represent, initially, an impact on firms' employee structure and the labor market, but not a shock to the demand for goods and services or on sales; that is, given that firms do not know how this novel shock will affect the demand for their product, they maintain the same sales expectation as before the restriction. In the beginning, firms' departments will have to review whether their payroll meets the restriction and, if it doesn't, think about how this novel policy will affect their initial optimal path of action regarding employment, $h^*_{\vartheta,f,t+1}$, where an asterisk denotes the optimal action (observable in the database) that a given department would have taken without the restriction.

The effective or final (with restriction) $h_{\vartheta,f,t+1}$ is endogenous to the model, since the effective labor cost, $\mathbf{C}(h_{\vartheta,f,t+1})$, when adapting to the restriction could change the department's decision to open a certain number of jobs. Also, $h_{\vartheta,f,t+1} > h^*_{\vartheta,f,t+1}$ is not possible insofar as each firm maintains its sales expectation, so no additional workers are needed for production. Therefore, $h_{\vartheta,f,t+1} \leq h^*_{\vartheta,f,t+1} + \epsilon$, where $\epsilon = 1[R \in \{3, 4\}]$, and it gives departments a hiring margin above their initial expectations in case they have no choice but to hire an additional worker (cases where there is only one gender and they need both).

Given the assumption that firms do not know how demand for their product will be affected—so they maintain the sales expectation for $t + 1$ that gave rise to $h^{*}_{\vartheta,f,t+1}$—the ideal scenario for the department is to be able to comply with the restriction without deviations from $h^{*}_{\vartheta,f,t+1}$. The reason for this is that the lower the number of new positions, the harder it will be for the firm to satisfy its sales expectations. Therefore, each department's decision can be summarized in choosing an $h_{\vartheta,f,t+1}$ that minimizes its distance with respect to $h^{*}_{\vartheta,f,t+1}$ (equation 20), while simultaneously complying with the gender wage gap restriction and not violating sales expectations (equation 21):

$$(20) \qquad \min_{h_{\vartheta,f,t+1}} \left\{ \left| h_{\vartheta,f,t+1} - h^{*}_{\vartheta,f,t+1} \right| \right\};$$

$$(21) \qquad h_{\vartheta,f,t+1} \leq h^{*}_{\vartheta,f,t+1} + \epsilon.$$

There is a third condition that must be satisfied when choosing $h_{\vartheta,f,t+1}$. I assume departments will be hesitant to choose an employment path $h_{\vartheta,f,t+1}$ in which $\mathbf{C}(h_{\vartheta,f,t+1}) > \mathbf{C}^{*}(h^{*}_{\vartheta,f,t+1})$, since further costs and adjustments could emerge during the clearing of the labor market (that is, rising wages owing to market pressure and a greater gender wage gap adjustment owing to the unexpected exit of workers going to other firms). Therefore, $\mathbf{C}(h_{\vartheta,f,t+1}) \leq \mathbf{C}^{*}(h^{*}_{\vartheta,f,t+1})$. Furthermore, when deciding if $h_{\vartheta,f,t+1} = h^{*}_{\vartheta,f,t+1}$ or lower, departments will take into account not only the labor costs in $t + 1$ associated with that decision but also the incentives they have to either stay in $h^{*}_{\vartheta,f,t+1}$ or move to the left (lower value). For instance, a lower $h_{\vartheta,f,t+1}$ entails lower payroll and hiring costs (if any), whereas higher values prevent revenue from decreasing (which would result if the firm had fewer workers and thus lower production) and avoid firing costs (if any). Equation 22 presents the firm's incentives to move based on the wages and compulsory bonus associated with hiring one fewer worker, $IMW_{\vartheta,f,t+1}$; the maximum of gender wage averages is chosen to compute the incentive since new positions with those specific averages as wages do not cause changes in the gender wage gap. Then, equation 23 presents the firm's incentives to move based on the hiring costs (if any) associated with hiring one fewer worker, $IMH_{\vartheta,f,t+1}$; note that this incentive makes sense only when $h_{\vartheta,f,t+1} > 0$, which is why the definition incorporates an indicator function.

$$(22) \qquad IMW_{\vartheta,f,t+1} \equiv (1+\eta)\max\left\{\overline{\omega}_{\vartheta,f,M,t+1}, \overline{\omega}_{\vartheta,f,F,t+1},\right\}.$$

$$(23)\qquad IMH_{\vartheta,f,t+1} \equiv a_{\vartheta,f}\left(s, L_{f,t}, h_{\vartheta,f,t+1}\right) \bullet \max\left\{\bar{\omega}_{\vartheta,f,M,t+1}, \bar{\omega}_{\vartheta,f,F,t+1}\right\} \bullet 1\left[h_{\vartheta,f,t+1} > 0\right].$$

Equation 24 presents the firm's incentives to stay based on the revenue reduction deriving from employing one fewer worker, $ISR_{\vartheta,f,t+1}$. The production of each department, $q_{\vartheta,f,t}$, is defined as a function of labor, $q_{\vartheta,f,t} = L_{\vartheta,f,t}^{\phi_s}$, in which ϕ_s is a sectoral output elasticity. The product price, $p_{\vartheta,f,t}$, is defined as a sectoral markup, μ_s, of the average costs of each product unit, $p_{\vartheta,f,t} = \mu_s \bullet \frac{C_{\vartheta,f,t}}{q_{\vartheta,f,t}}$. The latter is adjusted by $(h^*_{\vartheta,f,t+1} + 1 - h_{\vartheta,f,t+1})$ to account for increasing incentives when augmenting the deviation from $h^*_{\vartheta,f,t+1}$. Note that $IMW_{\vartheta,f,t}$ and $IMH_{\vartheta,f,t+1}$ do not need adjustment since $\mathbf{C}(h_{\vartheta,f,t+1})$ is updated with each $h_{\vartheta,f,t+1}$ value. Next, equation 25 presents the incentives to stay based on lower firing costs (if any), computed as an average, $ISF_{\vartheta,f,t}$. Similar to $ISR_{\vartheta,f,t}$, this incentive is multiplied by an adjustment factor. This incentive makes sense only when $h_{\vartheta,f,t+1} \leq 0$, which is why the definition incorporates an indicator function.

$$(24)\qquad ISR_{\vartheta,f,t+1} \equiv \left(h^*_{\vartheta,f,t+1} + 1 - h_{\vartheta,f,t+1}\right)\left(p_{\vartheta,f,t}\frac{\partial q_{\vartheta,f,t}}{\partial L_{\vartheta,f,t}}\right).$$

$$(25)\qquad ISF_{\vartheta,f,t+1} \equiv \left(h^*_{\vartheta,f,t+1} + 1 - h_{\vartheta,f,t+1}\right)\left[\overline{\iota \bullet \omega_{i,\vartheta,f,t} + SV\left(\omega_{i,\vartheta,f,t}\right)}\right] \bullet 1\left[h_{\vartheta,f,t+1} \leq 0\right].$$

Combining all incentives, I obtain the third restriction of each department's decision:

$$(26)\quad IMW_{\vartheta,f,t+1} + IMH_{\vartheta,f,t+1} + \mathbf{C}\left(h_{\vartheta,f,t+1}\right) \leq \mathbf{C}^*\left(h^*_{\vartheta,f,t+1}\right) + ISR_{\vartheta,f,t+1} + ISF_{\vartheta,f,t+1}.$$

The incentives to stay in the third restriction increase the right-hand side of the inequality, making it easier for $h_{\vartheta,f,t+1}$ values to meet the inequality without deviating to lower values. Similarly, incentives to move increase the left-hand side of the inequality, making it more challenging to meet the restriction and, therefore, to stay at values near $h^*_{\vartheta,f,t+1}$.

Finally, each department's decision can be expressed as follows:

$$\min_{h_{\vartheta,f,t+1}} \left\{ \left| h_{\vartheta,f,t+1} - h^{*}_{\vartheta,f,t+1} \right| \right\}, \tag{27}$$

subject to

$$h_{\vartheta,f,t+1} \leq h^{*}_{\vartheta,f,t+1} + \epsilon;$$

$$\max\left\{ \bar{\omega}_{\vartheta,f,M,t+1}, \bar{\omega}_{\vartheta,f,F,t+1} \right\} \leq \alpha \min\left\{ \bar{\omega}_{\vartheta,f,M,t+1}, \bar{\omega}_{\vartheta,f,F,t+1} \right\};$$

$$IMW_{\vartheta,f,t+1} + IMH_{\vartheta,f,t+1} + \mathbf{C}\left(h_{\vartheta,f,t+1}\right) \leq \mathbf{C}^{*}\left(h^{*}_{\vartheta,f,t+1}\right) + ISR_{\vartheta,f,t+1} + ISF_{\vartheta,f,t+1}.$$

The previous optimization problem holds even with $h^{*}_{\vartheta,f,t+1} < 0$. In this case, the pessimistic department's expectations on sales require a minimum decrease of $h^{*}_{\vartheta,f,t+1}$ without the restriction. This decision process is done only once by each department of firm f between t and $t + 1$ since it is only the initial reaction to the restriction. In the following section, I drop the subscripts of $h_{\vartheta,f,t+1}$ for convenience.

COMPLYING WITH THE RESTRICTION AND COMPUTING COSTS GIVEN A POTENTIAL NUMBER OF NEW JOBS (*h*). In general, for each possible h that the department considers it generates a viable way to deal with the restriction and an associated labor cost for the final period $t + 1$. Then the final h is chosen to minimize the distance with respect to h^{*}, subject to the restrictions in equation 27. This section emphasizes how a department manages to comply with the restriction and computes the final labor cost given a possible value of h. There are three potential cases to analyze, where h is positive, negative, or equal to zero. However, since there is heterogeneity in the circumstances each department experiences (types of R), the three cases will differ slightly according to the situation; since types 1 and 3 are analogous to types 2 and 4, respectively, I limit the discussion to types 0 (meets the restriction), 1 (favors men), and 3 (employs only men), according to the three possible cases of h. There is additional heterogeneity in terms of the firm's creation date, with new firms or departments within a firm that arise between the initial and final periods in the economy and established firms or departments that existed in the initial period before the restriction.

First, new departments or firms are viewed as payroll plans that have yet to be executed, and so h can only be positive. Since none of the workers has been hired yet, each department can determine the gender of the positions to be filled so that the restriction can be met ($R = 0$) and the optimal path regarding $h^*_{\vartheta,f,t+1}$ and $\mathbf{C}^*$ can be maintained; this unofficial discrimination in the form of gender-based hiring deriving from the firm's incentives to meet the restriction is discussed below. Initially, since none of the departments knows precisely the direction of the effect on wages from the forthcoming shock on the labor market, they determine the salaries of the positions based on their payroll plans and what they need to fulfill the restriction.

For departments that were already operating, a positive h implies that they open positions with wages such that R becomes or remains 0. New jobs, h, are divided among genders according to the current department's gender composition; if the department has only men or women, then h is divided according to the respective occupation's gender composition in that sector. Regarding the new positions' wages, if the department already has an R type 0, then the new positions will pay the average wage for the respective gender. If the department has an R type 1, men will be given the male average wage, and women's wages are set at the lowest amount that will move R to 0. Finally, if the department has an R type 3 (only men), then again men are paid the average male wage, while newly hired women are paid the lowest amount that will move R to 0.

An h of zero means that the department does not expect to hire or fire workers. If its R is type 0, it already complies with the restriction. If R is type 1, then the adjustment is made either by increasing specific (strategic) wages in order to transform R to type 0 or by increasing the female average wage and adjusting all associated female wages accordingly.[8] If the department has an R type 3 (all men), it has no option but to hire a woman with a wage such that the restriction is met. For a negative h, the department analyzes which workers to fire in order to move R to 0 while incurring the lowest firing costs possible. In the case of type 3 or 4, after the firing, the department still needs to hire a person of the opposite gender with a wage such that the payroll achieves an R of 0.

8. The cost ($n\gamma$) associated with altering either specific wages or the gender average wage is the same; the necessary change (γ) in the gender average salary to optimally satisfy R (that is, $\bar{\omega}^*$) can be redistributed arbitrarily (first option) or equally (second option):

$$\bar{\omega} = \frac{\sum_{i=1}^{n} \omega_i}{n} \Rightarrow \bar{\omega}^* = \bar{\omega} + \gamma = \frac{n\gamma + \sum_{i=1}^{n} \omega_i}{n}.$$

Labor Market

After the number of positions h within each department is endogenously computed, jobs will be eliminated or created, and wages will be adjusted and workers fired as needed. These changes, along with new people entering the labor force, affect supply and demand in the labor market and produce upward or downward pressure on the wages offered for new jobs by the departments of all firms. The movement in a specific occupation's demand curve depends on new and eliminated jobs in ϑ. Analogously, the supply curve movement depends on workers that enter and exit the specific labor market for ϑ.

Previously, I mentioned that firms discriminate by gender when filling a given job (that is, each position has an assigned gender). This informal discrimination in hiring is a unique characteristic of these labor markets since it is unrealistic to expect firms to consider both genders when they have incentives to hire either men or women to meet the intra-occupational restriction. Therefore, since the labor market's impact on a specific occupation could be different for each gender, the exercise on moving curves is performed once for men and once for women. The model has one labor market for each gender-occupation combination in the economy. Each of these markets will determine the average wage by gender that must prevail in that occupation in order to match the demand and supply of workers. I assume that supply and demand take a basic linear form, as follows:

$$\text{S: } \bar{\omega}_{\vartheta,G} = \theta_{S,\vartheta,G} \bullet L^{S}_{\vartheta,G}; \tag{28}$$

$$\text{D: } \bar{\omega}_{\vartheta,G} = -\sigma_{\vartheta,G} \bullet L^{D}_{\vartheta,G} + \theta_{D,\vartheta,G}. \tag{29}$$

The parameter $\sigma_{\vartheta,G}$ represents the slope of the labor demand.[9] In the initial period, I have the average wage and the number of workers employed for each gender-occupation; this information is used in each of the market equations to compute $\theta_{S,\vartheta,G}$ and $\theta_{D,\vartheta,G}$ for the initial equilibrium, where $\theta_{S,\vartheta,G}$ is the slope of the labor supply and $\theta_{D,\vartheta,G}$ is the labor demand y-axis intercept. These parameters are used, along with the demand and supply movements after the departments' reaction to the restriction, to recalculate the average gender

9. Since this parameter is not endogenous to the model, I calculate it following Alfaro, Campos, and Lankester (2019), who estimate the labor demand elasticity in Costa Rica for the comprehensive period 2005–17, resulting in a –0.358 percent impact on labor demand by a 1 percent increase in wages. The previous elasticity is algebraically transformed to calculate $\sigma_{\vartheta,G}$.

wage equilibrium in occupation ϑ, and, therefore, to compute the pressure on wages in that specific labor market for the new positions.

Equation 30 presents the pressure that wages may suffer during the new job negotiations because of the changes the labor market for each occupation, by gender:

$$\omega'_{z,\vartheta,G} = \omega_{z,\vartheta,G} + \Delta\bar{\omega}_{\vartheta,G} \tag{30}$$

where $\omega_{z,\vartheta,G}$ is the wage offered by the department; $\Delta\bar{\omega}_{\vartheta,G}$ is the change in the average wage determined by the labor market for occupation ϑ and gender G; and $\omega'_{z,\vartheta,G}$ is the updated wage for each position in the market. Therefore, the final wage for a new job is determined by a department's specific situation regarding the intra-occupational restriction and pressure in the labor market. This is an important clarification because in traditional competitive labor market models, wages are determined by market-level supply and demand factors rather than by the wage-setting policies of particular firms (Card, Cardoso, and Kline, 2016), and wages can vary across firms if there are market-based compensating differentials for firmwide amenities or disamenities, such as long work hours (Bertrand, Goldin, and Katz, 2010; Card, Cardoso, and Kline, 2016). In this article, I introduce a new circumstance that justifies wage differences across firms, namely, different situations in terms of compliance with the gender wage gap restriction (the type of R and the magnitude of the gap).

Workers' Decision and Labor Market Clearing

Once the new jobs have been incorporated into each labor market differentiated by occupation and gender, the matching between firms and workers starts. It is during this process that labor market pressure causes a change in the initial wages associated with each position. Some workers experience wage rigidity: because wages can be renegotiated only when there is a credible threat (Jarosch, 2021), workers who did not apply to any positions or did not receive any offers are not able to renegotiate the wage at their current job. Nevertheless, all workers in the economy, including independents and new labor market entrants, can apply and compete for the positions that arise, even if they already have a job.

There is no limit to the number of applications that a worker can send, but workers can apply only to jobs that are in the same occupation as their current job, offer the same or higher wages as currently earned (incorporating firms' reaction), and are located in the same canton where the worker currently works. The canton variable is used as a proxy for how far the workers are willing to travel work. For fired workers, the wage is compared to the previous one, and for new workers, it is the one they would have had according to data on the final period of analysis. Workers' applications indicate their gender, wage decile in their firm, and experience.

For each job opening, a firm receives all applications and applies an initial filter according to the gender they want to hire. Then each department assigns a priority number to the remaining applications based on the worker's experience and wage decile (a proxy of the worker's importance and position in the old or current firm). Each department simultaneously makes an offer to the highest priority application until the position is filled.

Workers receive all job offers at the same time, but they do not decide immediately; they have some space to wait for additional offers that might be rejected by other candidates. If several offers are received, the one with the highest wage is chosen. If a position is rejected, then the department offers the job to the next priority application in line. If no wage is enough to fill the position, then it is filled from an unobservable pool of unemployed and informal workers. These last positions cannot be left unfilled since that would imply that the firms can function without them and they were not really necessary in the first place. When the matching ends, all departments will be able to see how much they deviated from their initial strategy to meet the restriction in response to the pressure in the labor market and other conditions. At this point, if the deviation prevented the departments from achieving an R type 0, then they increase the lowest gender average wage and adjust all associated wages, using deviations from the mean, in order to finish complying with the restriction.

After this process is complete, all departments end up with an updated wage structure in $t + 1$ that complies with the intra-occupational restriction. The respective computations on decomposition components and gender wage gaps in the economy's occupations can then be recalculated in order to proceed with the analysis on the impacts of the restriction.

Calibration of Parameters

The components of the decomposition and the solution to the quantitative exercise are computed using a monthly frequency of the available data for the years 2008, 2013, and 2018. I am interested in the possible effects of gender wage gap restrictions in the short term, so I use the information for February

TABLE 2. Calibration of Model Parameters

Parameter	*Value*	*Description*
t	February	Initial period in which the restriction is announced
$t+1$	November	Final period in which all departments in a firm must comply with the restriction
η	1/12	Monthly portion of annual bonus
ι	1	Fraction of annual bonus paid when worker is laid off
$SV(\omega_{i,\vartheta,f,t})$	$\frac{1}{30}\omega_{i,\vartheta,f,t}\left(\begin{array}{c} 7\cdot\mathbf{1}_{[3<m\leq 6]}+14\cdot\mathbf{1}_{[6<m\leq 12]} \\ +20\frac{m}{12}\cdot\mathbf{1}_{[12<m\leq 96]}+20\cdot 8\cdot\mathbf{1}_{[96<m]}\end{array}\right)$	Severance pay; m months worked in firm (Source: Ministry of Labor)
$a_{\vartheta,f}(s,L_{f,t},h_{\vartheta,f})$	$\max\left\{a(s),\, a(L_{f,t})\cdot\frac{1}{4}\cdot\sqrt{h_{\vartheta,f,t+1}/\overline{h}_{\vartheta,f,t+1}}\right\}$	Hiring cost as a percentage of wage
α	{1, 1.5, 2, 2.5, 3}	Restriction on intra-occupational gender wage gap
ϵ	$1[R\in\{3,4\}]$	Each department's hiring margin over their initial expectations ($h^{*}_{\vartheta,f,t+1}$)
Experience	max{social security contributions, months employed by the firm}	Worker's work experience (Source: Administrative data computation)
ϕ_s	Mean: 0.84; standard dev.: 0.00	Output elasticity by sector (Source: Alfaro, Manelici, and Vásquez, 2021)
μ_s	Mean: 1.25; std. dev.: 0.00	Markup over cost by sector (Source: Alfaro, Manelici, and Vásquez, 2021)
$\sigma_{\vartheta,G}$	$2.79\cdot\overline{\omega}_{\vartheta,G}/L_{\vartheta,G}$	Slope of the labor demand curve (Source: Alfaro, Campos, and Lankester, 2019)
$\theta_{S,\vartheta,G}$	Mean: 817.6; std. dev.: 3,116.8 (unit: colones)	Slope of the labor supply curve (Source: Administrative data computation)
$\theta_{D,\vartheta,G}$	Mean: 1,996,094.0; std. dev.: 1,492,859.0 (unit: colones)	Labor demand y-axis intercept (Source: Administrative data computation)

and November of the respective years as the initial and final periods in the model. I do not consider January and December so as to exclude seasonal jobs to the extent possible.

Table 2 lists the model parameters and their definitions. The parameter $a_{\vartheta,f}(s, L_{f,t}, h_{\vartheta,f})$, which represents each firm's hiring costs for each worker as a percentage of the wage, is calibrated following Blatter, Muehlemann, and Schenker (2012), who calculate hiring costs, in number of weeks of wage payments, as a function of the economic sector $a(s)$, firm size $a(L_{f,t})$, and number of workers hired on average; only positive $h_{\vartheta,f,t+1}$ are considered in the average). Markup over cost (μ_s) and output elasticity (ϕ_s) were calibrated following Alfaro, Manelici, and Vásquez (2021), who use an expanded version of my administrative data set to compute these specific parameters.

ALGORITHM 1. Impact of an Intra-occupational Gender Wage Gap Restriction

```
In period t, every firm f is notified of the gender wage gap restriction (α).
Analysis of the situation:
If Department (ϑ, f) meets restriction (α), then
    R = 0;
else
    R ∈ {1, 2, 3, 4}
end
Deciding on the number of positions to be opened or closed (h):
for h = h* to 0 to −L do
    while R ≠ 0 do
        Generate payroll modifications to achieve R = 0 using h
    end
    Compute C
    Verify compliance of:
    if IMW + IMH + C ≤ C* + ISR + ISF then
        break (inside and outside loop)
    else
        next h value
    end
end
Using h, Department (ϑ, f) executes: pay modifications, firing, and job posting.
Workers' reaction:
Worker i (fired, employed, new labor force, independent) decides which jobs to apply for.
    Application factors: wage, occupation, location.
Labor market:
Wage pressure on each occupation-gender labor market.
Firms' decisions:
Department (ϑ, f) receives all applications.
    Decision factors: gender, experience.
Workers' decisions:
Worker i receives acceptances and takes the highest-paid job.
Final checks:
if Department (ϑ, f) meets restriction (α), then
    Computes final labor cost C
else
    Final adjustments:
    ↑ lowest gender average wage → R = 0 (wages readjust based on deviation from gender average)
    Computes final labor cost C
end
```

Note: This algorithm presents the decisions that firms and workers face and the clearing of the labor market after the gender wage gap restriction is announced. First, departments within firms find the optimal number of job positions to open or close based on sales expectations, labor costs, and compliance with the gender wage gap restriction. Second, workers (with or without jobs) decide which new job positions to apply for, based on occupation, location, and wages. Available job positions in the market suffer wage changes due to labor supply and demand pressure. Finally, firms decide whether to hire job candidates using the variables gender and experience, and workers decide whether to accept job offers based on the wage.

TABLE 3. Wage Inequality (T) Decomposition by Gender: 2008, 2013, and 2018

	Contribution to **T**			*Contribution to* Δ**T**		
Component	*2008*	*2013*	*2018*	*2008–13*	*2013–18*	*2008–18*
Δ**T** (%)	—	—	—	45.56	14.49	66.65
T	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
$\mathbf{W}_{gender}$	0.9999	0.9986	0.9984	0.9959	0.9966	0.9961
$\mathbf{B}_{gender}$	0.0001	0.0014	0.0016	0.0041	0.0034	0.0039
$\mathbf{W}(\vartheta, f, T, s)$	0.3923	0.4079	0.4246	0.4422	0.5400	0.4732
$\mathbf{B}(s)$	0.1028	0.1316	0.1058	0.1948	–0.0721	0.1104
$\mathbf{B}(T, s)$	0.0126	0.0244	0.0159	0.0502	–0.0426	0.0208
$\mathbf{B}(f, T, s)$	0.2352	0.1849	0.1990	0.0744	0.2969	0.1448
$\mathbf{B}(\vartheta, f, T, s)$	0.2571	0.2513	0.2546	0.2384	0.2779	0.2509
$\mathbf{W}_M$	0.2734	0.2674	0.2676	0.2543	0.2687	0.2589
$\mathbf{W}_F$	0.0956	0.1186	0.1323	0.1692	0.2263	0.1873
$\mathbf{P}_M$	0.1171	0.1311	0.1381	0.1619	0.1862	0.1696
$\mathbf{P}_F$	0.1074	0.1099	0.1145	0.1154	0.1460	0.1250
$\mathbf{G}_M$	–0.1080	–0.1220	–0.1270	–0.1521	–0.1656	–0.1564
$\mathbf{G}_F$	–0.0940	–0.0980	–0.1010	–0.1064	–0.1216	–0.1112
$\mathbf{B}(s)_M$	0.0333	0.0305	0.0223	0.0244	–0.0348	0.0057
$\mathbf{B}(s)_I$	0.0155	0.0392	0.0211	0.0912	–0.1042	0.0294
$\mathbf{B}(s)_F$	0.0540	0.0619	0.0625	0.0792	0.0669	0.0753
$\mathbf{B}(T, s)_M$	0.0056	0.0079	0.0041	0.0130	–0.0225	0.0018
$\mathbf{B}(T, s)_I$	0.0041	0.0103	0.0068	0.0240	–0.0170	0.0110
$\mathbf{B}(T, s)_F$	0.0029	0.0061	0.0050	0.0132	–0.0031	0.0080
$\mathbf{B}(f, T, s)_M$	0.1614	0.1143	0.1123	0.0111	0.0984	0.0387
$\mathbf{B}(f, T, s)_I$	0.0402	0.0358	0.0462	0.0263	0.1182	0.0554
$\mathbf{B}(f, T, s)_F$	0.0336	0.0347	0.0405	0.0371	0.0804	0.0508
$\mathbf{B}(\vartheta, f, T, s)_M$	0.1978	0.1760	0.1725	0.1281	0.1483	0.1345
$\mathbf{B}(\vartheta, f, T, s)_I$	0.0028	0.0166	0.0212	0.0468	0.0530	0.0488
$\mathbf{B}(\vartheta, f, T, s)_F$	0.0565	0.0587	0.0610	0.0635	0.0765	0.0676

Notes: The first section is a basic within-between decomposition by gender. The second section presents the first phase of the decomposition, that is, by sector, type of firm (public or private), firm, and occupation. The third section presents the second phase, that is, a decomposition of the first phase by gender. A positive sign in the contribution to Δ**T** indicates that the component's growth is in the same direction as inequality growth.

Results

Table 3 presents the calculation of the components of the inequality decomposition by gender, as a share of total inequality, for both the first and the second phase of the decomposition. It also shows how much each component contributes to the growth of inequality. Total inequality increased 46 percent between 2008 and 2013, 15 percent between 2013 and 2018, and 67 percent in the full period of 2008–18. The increase in wage inequality after 2008 coincides with the post-crisis period, when the Gini coefficient rose 3.9 percent, from 48.7 in 2008 to 50.6 in 2009, according to World Bank estimates.[10] Thereafter the Gini coefficient fell, reaching 48.0 in 2018. However, this measure is calculated through a survey and is based on household income, not wage income. Using the employer-employee administrative data set, I find Gini coefficients of 38.9 for 2008, 41.6 for 2013, and 40.9 for 2018; this shows that the inequality in the real wage distribution for formal workers grew 5.1 percent between 2008 and 2018. It is important to note that the wage variance and the wage-based Gini coefficient do not measure precisely the same aspect; the first focuses on the distance between wages and the second on proportions of the distribution (concentration). This is why their growth magnitudes do not match.

Finally, the table also presents the basic version of the decomposition according to gender. Almost all of the growth in total inequality can be attributed to wage differences within each gender (99.6 percent due to $\Delta\mathbf{W}_{gender}$) and the almost insignificant remainder to wage differences between genders (0.4 percent due to $\Delta\mathbf{B}_{gender}$). However, this distinction between components and, more specifically, the vagueness in the definition of the dominant component do not allow a deep understanding of the nature of the problem inherent in the wage differences within each gender group.

In the first phase of the decomposition, the inequality in each year of study depends mainly on the wage differences within workers of the same occupation in each firm, with around 41 percent due to $\mathbf{W}(\vartheta, f, T, s)$, followed by differences between occupations at the same firm, with 25 percent due to $\mathbf{B}(\vartheta, f, T, s)$. Likewise, much of the growth in wage inequality can be explained by an increase in wage differences within occupations in each firm—47 percent of the growth in inequality occurs through $\Delta\mathbf{W}(\vartheta, f, T, s)$—and an increase in the salary differences between occupations at the same firm—around 25 percent of the growth in inequality occurs by $\Delta\mathbf{B}(\vartheta, f, T, s)$.

In the second phase of decomposition, the key components of total inequality ($\mathbf{T}$) are (1) the differences within men of the same occupation in each firm (27 percent due to $\mathbf{W}_M$), (2) the penalty for occupational segregation in each firm (24 percent from $\mathbf{P}_M + \mathbf{P}_F$), and (3) the award for the fact that a certain number of occupations in each firm do not initially breach the restriction (around –22 percent from $\mathbf{G}_M + \mathbf{G}_F$). With regard to the latter, the award depends on how many departments were initially complying with the restriction

10. World Bank, Gini index (indicator) (https://data.worldbank.org/indicator/SI.POV.GINI?locations=CR, accessed March 27, 2021).

and the weights that magnify the award or penalty, for example, the number of workers in each firm's occupation. Finally, in this more precise decomposition, the growth in inequality is mainly explained by an increase in the segregation penalty, with 30 percent of the total increase due to $\Delta(\mathbf{P}_M + \mathbf{P}_F)$; by an increase in wage differences within men of the same occupation of each firm, with 26 percent due to $\Delta\mathbf{W}_M$; and by a similar increase in wage differences among women, with around 19 percent due to $\Delta\mathbf{W}_F$. The growth in total inequality is slowed by a decrease in the gender wage gaps, specifically around –27 percent from $\Delta(\mathbf{G}_M + \mathbf{G}_F)$.[11] This demonstrates the role that gender wage gaps play in reducing inequality, as also found by Piketty, Saez, and Zucman (2018).

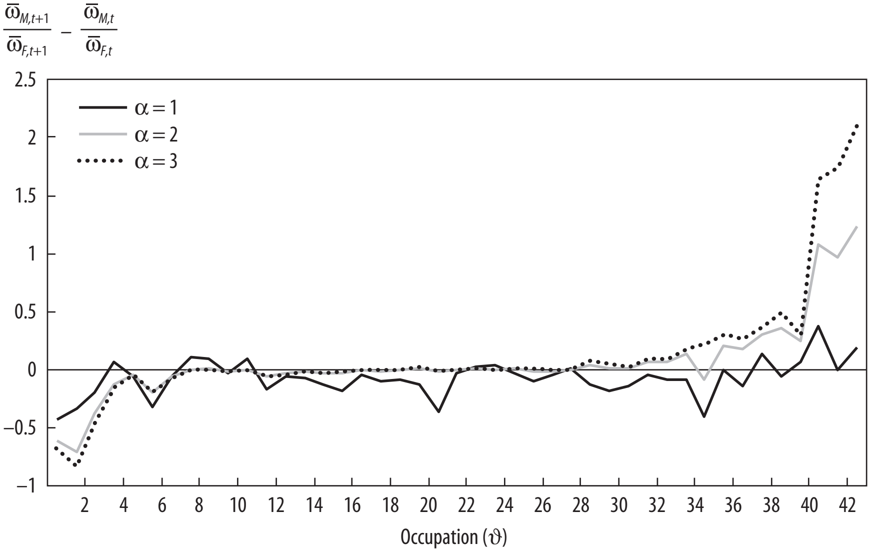
A question that naturally arises is why I impose a restriction on an element like $\mathbf{G}_M + \mathbf{G}_F$ if the differences within men of the same occupation in each firm seem to be more important. This question must be answered in two parts. First, the restriction is not only for the gaps between the gender average wages that characterize $\mathbf{G}_M + \mathbf{G}_F$ but also applies to the gender segregation and dominance in each department, which manifests in $\mathbf{P}_M + \mathbf{P}_F$. The latter represents an important component in inequality and its growth, while the former could become more negative (decreasing inequality) if remaining violations of the restriction are eliminated. Second, it is not necessary to implement an additional constraint on $\mathbf{W}_M$ and remove the one on gender wage gaps; although these components measure different aspects, an interdependence is inherently created when an incentive such as the restriction in question is introduced, mainly between the differences within each gender, $\mathbf{W}_M$ and $\mathbf{W}_F$, and the gender wage gaps, $\mathbf{G}_M + \mathbf{G}_F$. Specifically, this interdependence, which will allow the impact of the restriction to manifest in the differences within each gender, occurs when firms try to adapt their gender wage averages to the restriction, since they are inevitably modifying wages within each gender. I return to this insight in the discussion of the quantitative exercise results.

Impact of the Restriction on the Economy's Occupational Gender Wage Gaps

Using 2013 for illustration, figure 1 shows that the impacts on gender wage gaps of each occupation in the economy are particularly important for occupations with a low participation of either gender. Moreover, those with low male participation (left) tend to present a deterioration in the gaps in favor of women (that is, they move away from $\bar{\omega}_M/\bar{\omega}_F = 1$; see figure 2), and

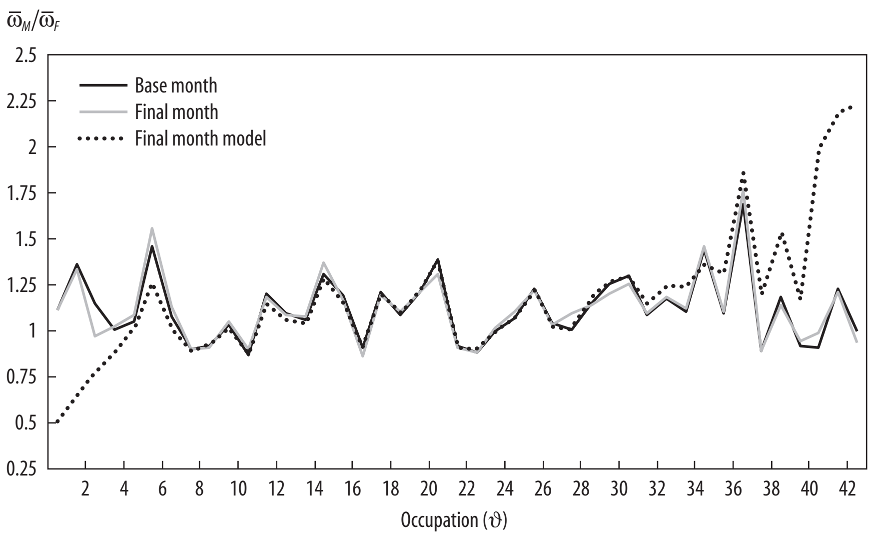
11. See table 5 in online appendix A.

FIGURE 1. Impact on Gender Wage Gap for Each Occupation, $\frac{\bar{\omega}_{M,t+1}}{\bar{\omega}_{F,t+1}} - \frac{\bar{\omega}_{M,t}}{\bar{\omega}_{F,t}}$: 2013



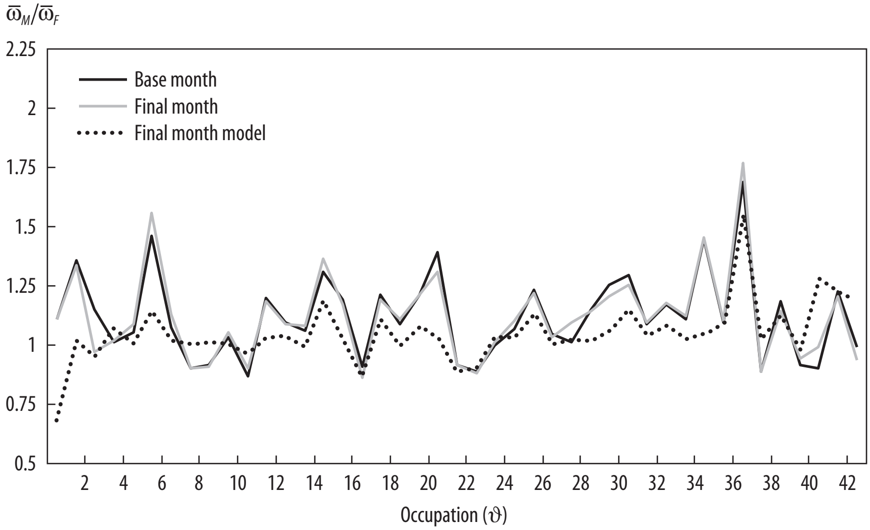
Note: Occupations are ordered from highest to lowest proportion of women in the composition. The first three are preschool and special education (1), domestic worker (2), and beautician (3); the last three are builder (41), mechanic (42), and transporter (43). See online appendix A (table 6) for the full list of occupations.

FIGURE 2. Gender Wage Gap for Each Occupation in Initial and Final Periods: 2013, $\alpha = 2$



Note: Occupations are ordered from highest to lowest proportion of women in the composition. The first three are preschool and special education (1), domestic worker (2), and beautician (3); the last three are builder (41), mechanic (42), and transporter (43). See online appendix A (table 6) for the full list of occupations.

FIGURE 3. Gender Wage Gap for Each Occupation in Initial and Final Periods: 2013, $\alpha = 1$



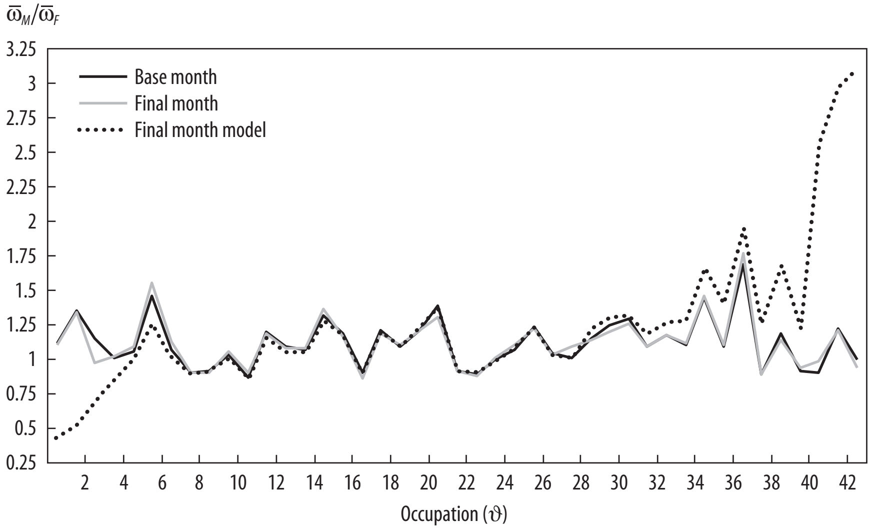
Note: Occupations are ordered from highest to lowest proportion of women in the composition. The first three are preschool and special education (1), domestic worker (2), and beautician (3); the last three are builder (41), mechanic (42), and transporter (43). See online appendix A (table 6) for the full list of occupations.

occupations with low female participation (right) present a deterioration of the gaps in favor of men.

Figure 3 shows that with a more stringent restriction ($\alpha = 1$), the deterioration effect ceases, because the restriction forces the departments to equalize these wage averages, and the counterfactuals thus approach $\overline{\omega}_M/\overline{\omega}_F = 1$. In contrast, figure 4 shows that greater flexibility ($\alpha = 3$) generates the opposite effect, since a greater margin is allowed in the differences between averages.

Figures 7 to 14 in online appendix C present the above effects for 2008 and 2018. Intuition suggests that for a certain occupation ϑ, if the value that the gaps can take is reduced or limited, there should be either a move toward the line of equal averages ($\overline{\omega}_M/\overline{\omega}_F = 1$) or at least some stability compared to initial gaps, as with occupations in the center of the graph. However, the figures reveal an important opposite effect in occupations with significant segregation. The fact that the gender wage gaps in each occupation of each firm cannot be greater than α does not imply the same is true for the economy

FIGURE 4. Gender Wage Gap for Each Occupation in Initial and Final Periods: 2013, $\alpha = 3$



Note: Occupations are ordered from highest to lowest proportion of women in the composition. The first three are preschool and special education (1), domestic worker (2), and beautician (3); the last three are builder (41), mechanic (42), and transporter (43). See online appendix A (table 6) for the full list of occupations.

as a whole.[12] Thus, one reason for the effect observed in occupations with a high concentration of one gender is that at the economy level, these occupations accumulate many departments with an *R* type 3 or 4. To comply with the restriction, these departments are forced to hire at least one member of the opposite gender, either by opening new positions or by using positions they expected to open due to growth. For example, for the set of departments without women, new jobs will be given to women, but the entry wage may be low relative to that of the few women who were already in that occupation at other firms in the economy. This causes the average salary of the few women in that occupation to decrease at the economy level and the gap to deteriorate in

12. The following example shows how two single-occupation firms can comply with the restriction without necessarily transferring that property to the occupation in the economy. Firm 1 has a woman with salary 1 and a man with salary 2, while firm 2 has a woman with salary 2 and a man with salary 4. Both comply with the restriction, and the global gender wage gap is equal to 2. However, if in the end firm 1 had three women with salary 1, both firms would still comply with the restriction, but the gap for this occupation in the economy is now 2.4.

TABLE 4. Wage Inequality (T) Decomposition by Gender: 2013 Simulation

Component	$\alpha = 1$	$\alpha = 2$	$\alpha = 3$
Contribution to Δ**T** (%)	25.18	−11.19	−8.38
T	1.0000	1.0000	1.0000
$\mathbf{W}_{gender}$	0.9945	0.9968	0.9951
$\mathbf{B}_{gender}$	0.0055	0.0032	0.0049
$\mathbf{W}(\vartheta, f, T, s)$	0.5709	0.9158	0.9352
$\mathbf{B}(s)$	0.0529	−0.0112	−0.0208
$\mathbf{B}(T, s)$	0.0123	−0.0115	−0.0188
$\mathbf{B}(f, T, s)$	0.1309	0.0872	0.0308
$\mathbf{B}(\vartheta, f, T, s)$	0.2331	0.0196	0.0735
$\mathbf{W}_M$	0.1631	0.6451	0.6612
$\mathbf{W}_F$	0.5517	0.1055	0.1441
$\mathbf{P}_M$	−0.0011	0.0811	0.1109
$\mathbf{P}_F$	0.1884	−0.0632	−0.0683
$\mathbf{G}_M$	−0.0591	−0.0071	−0.0446
$\mathbf{G}_F$	−0.2720	0.1544	0.1319
$\mathbf{B}(s)_M$	0.0090	−0.0038	−0.0052
$\mathbf{B}(s)_I$	0.0254	−0.0161	−0.0217
$\mathbf{B}(s)_F$	0.0185	0.0088	0.0062
$\mathbf{B}(T, s)_M$	0.0035	−0.0039	−0.0060
$\mathbf{B}(T, s)_I$	0.0044	−0.0048	−0.0084
$\mathbf{B}(T, s)_F$	0.0043	−0.0028	−0.0043
$\mathbf{B}(f, T, s)_M$	−0.0144	0.1378	0.0893
$\mathbf{B}(f, T, s)_I$	0.0992	−0.0526	−0.0672
$\mathbf{B}(f, T, s)_F$	0.0460	0.0021	0.0087
$\mathbf{B}(\vartheta, f, T, s)_M$	0.0201	0.0828	0.0877
$\mathbf{B}(\vartheta, f, T, s)_I$	0.1652	−0.1070	−0.0795
$\mathbf{B}(\vartheta, f, T, s)_F$	0.0477	0.0438	0.0653

Notes: The first section is a basic within-between decomposition by gender. The second section presents the first phase of the decomposition, that is, by sector, type of firm (public or private), firm, and occupation. The third section presents the second phase, that is, a decomposition of the first phase by gender. A positive sign in the contribution to Δ**T** indicates that the component's growth is in the same direction as inequality growth.

favor of men. The same would occur for occupations with low male participation. However, the effect is much greater in the case of female segregation (occupations to the right of the graph) because this problem of occupational segregation occurs mainly for women (that is, there are more *R* type 3 departments than type 4). Finally, this effect is consistent in the other simulations.

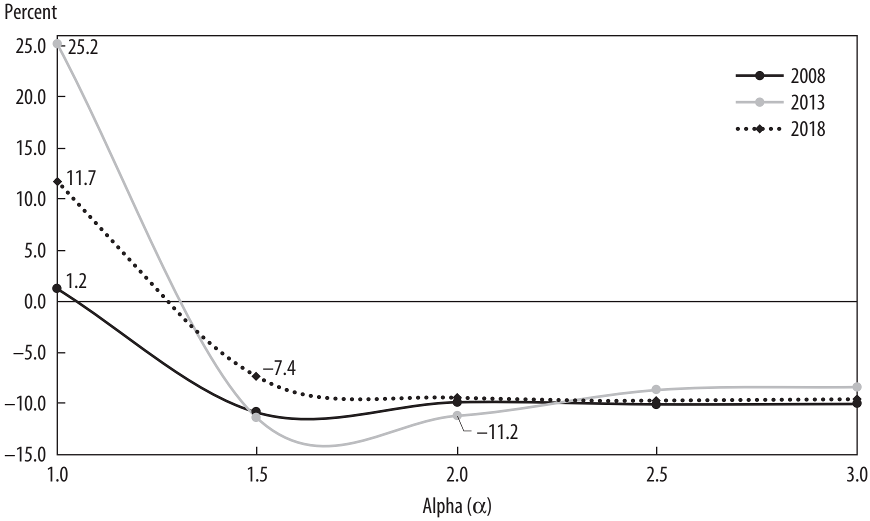
Impact of the Restriction on Wage Inequality

Table 4 presents the effects of the short-term restriction on the inequality components for the 2013 exercise. For $\alpha = 2$, total wage inequality decreases by 11.2 percent. The reduction is mainly driven by a decrease in wage differences within men of the same occupation in each firm (65 percent of $\Delta\mathbf{T}$ due to $\Delta\mathbf{W}_M$), as well as in differences within women (11 percent due to $\Delta\mathbf{W}_F$). The model suggests that when firms (more specifically, occupations within firms) are more vigilant of gender average wages, because of the restriction they must fulfill, wages within each gender tend to be relatively closer as a mechanism to maintain a better control of the respective average wage. Also, there is a decrease in the effect of segregation (penalization for not dominating, $\mathbf{P}_M + \mathbf{P}_F$) that contributes 2 percent to the reduction of inequality; however, the effect goes almost unnoticed because it is lowered by an increase in the female segregation effect ($\mathbf{P}_F$). The cause of this increase is not an increasing male dominance in each department but the fact that the restriction allowed the extreme female segregation to come to the surface: according to equation 11, the final penalization is dependent on how many women are affected and their wage, which were both 0 at the beginning for *R* type 3 cases.

Now, the initial objective of the intra-occupational gender wage gap restriction in each firm was to counter the penalties that arose from $\mathbf{G}_M$ or $\mathbf{G}_F$, for having gender average wages above the allowed limit. However, despite the fact that departments are no longer penalized for this reason, only rewarded for complying with the restriction, the sum of these components does not contribute as much as other components to the reduction of inequality; specifically, $\mathbf{G}_F$ contributes 15 percent and $\mathbf{G}_M$ zero. This occurs because the penalties for not complying with the restriction had a dual character (the benefit of one gender group was a disadvantage for the other), so reducing the penalties of one group eliminates the benefit of the opposite gender.[13] An extension of this point is shown in the $\alpha = 1$ case, in which the restriction requires gender wage gaps to be below the penalization threshold of 2. In this case, all departments receive a benefit for having gender wage gaps below the penalization threshold; furthermore, $\mathbf{G}_F$ contributes more to the inequality reduction since women not only eliminated the wage domination penalties

13. Suppose there is a single-occupation firm in which there is a man with a salary of 4 and a woman with a salary of 1. This generates a benefit for the man of $-7 = (1 - 2 \times 4)$, a penalty for the woman of $2 = (4 - 2 \times 1)$, and a global benefit of –5. For simplicity, I will ignore the gap weights in this example. Now, suppose that by abiding by the restriction, the man's salary falls to 2 (remember that it is actually a group of men whose average can drop through the use of scheduled job layoffs and new low-wage jobs) and the woman's wage remains at 1. This generates an impact of $0 = (2 - 2 \times 1)$ to the woman (punishment ceases because of the restriction), and for men the impact is $-3 = (1 - 2 \times 2)$ (they still draw a benefit, although smaller, for being above the woman's salary). In this situation, the aggregate benefit has dropped to –3.

FIGURE 5. Restriction's Impact on Total Wage Inequality Growth: 2008, 2013, and 2018 Simulations



(reached the threshold of 2) but are now receiving benefits because of the elimination of the gender wage gap (also reached a more difficult threshold of 1); on the other hand, $\mathbf{G}_M$ does not contribute much because men lost the benefits of being the wage-dominating gender and receive some new benefits from the gender wage gap elimination.

The decrease in total inequality is also observed when the constraint is relaxed ($\alpha = 3$), with a decrease of 8.4 percent (table 4), and when the constraint is tightened ($\alpha = 1.5$), with a decrease of 11.4 percent (see figure 5). However, a very tight gender wage gap restriction ($\alpha = 1$) dissipates the wage inequality reduction, causing an increase of inequality by 25.2 percent. The model's results suggest two main drivers for this effect, which can be seen in the same table 4. The first reason pertains to the within-gender differences, which may experience an increase of dispersion when more drastic changes in wages are needed in order to abide by the $\alpha = 1$ restriction; this occurs mainly for the gender with the lowest average wage, which in this case is $\mathbf{W}_F$. Second, since the $\alpha = 1$ constraint requires the equalization of both genders' average wage, differences between occupations, firms, type, and sector (based on wage averages) tend to increase; that is, the equalization of both genders' average wage within each occupation for each firm spotlights the average wage differences between those occupations within firms—that is, $\mathbf{B}(\vartheta, f, T, s)$. The past effects are also present, with variations in magnitudes, in the other simulations (see online appendix C).

Figure 5 shows the impact of the restriction on total wage inequality for all simulations. It shows how a more severe restriction is counterproductive, even reversing the potential benefit of inequality reduction.

The quantitative exercise results are especially interesting to analyze with respect to the initial situation and not with respect to the not modeled (real) final results. Although the simulations consider agent reactions, they are focused on a single event of interest, the restriction, and do not incorporate other events that may have occurred in the period. For this reason, they should not be compared directly with the final results that would have happened without the restriction on the intra-occupational gender wage gaps.

This precaution is of vital importance to avoid falling into excessive optimism or unfounded accusations. Specifically, there are going to be scenarios in which the measure of total inequality will increase or decrease from one period to another. If it decreases, and the restriction produces a labor scenario where inequality also decreases, but not as much as if the restriction had not been applied, voices will be raised against state intervention, claiming that this is proof that government involvement not only consumes resources but also increases distortions in the economy. On the other hand, if the inequality measure rises from one period to another and the simulation produces a scenario in which it decreases, other voices could praise the successful government intervention. In reality, what the simulations argue is that the restriction on the intra-occupational gender wage gaps in each firm can produce, by themselves (ignoring other events), a decrease in wage inequality, which is materialized through an interrelation of the components: occupational segregation, gender wage gaps, and inequality within the same gender.

Discussion

The previous sections have shown that the $\alpha = 2$ restriction could have promising results. The model suggests that it not only limits the magnitude of the gender wage gap within each occupation in each firm but also reduces wage inequality by about 10 percent. However, the model's results also show that in the end, the restriction would produce additional labor costs, as defined in equation 15, for approximately 50 percent of all departments.

Although the wage inequality decomposition points to a restriction of $\alpha = 2$, it might be desirable to set the restriction according to its benefits and downsides. The model's results suggest that toughening the restriction too much can backfire and increase overall inequality. The positive effect associated with an $\alpha = 1$ is that it causes the economy's occupational gender wage gaps to approach the equality line ($\bar{\omega}_M/\bar{\omega}_F = 1$). On the other hand, a more flexible restriction still limits the departments' gender wage gaps within a range and achieves a similar wage inequality reduction. As discussed, this $\alpha = 3$ has the downside of producing a gender wage gap deterioration in highly segregated occupations, but this issue could be solved with a differentiated α restriction, according to the gender concentration within each occupation in the economy. Moreover, occupations with a particularly high percentage of men or women, or entirely without one of the genders, could be given the most rigorous restriction of $\alpha = 1$, while the others are subject to $\alpha = 2$.

The results were computed based on a model that makes two essential assumptions. First, I assume a short time between the announcement of the restriction and compliance (approximately 1 year). This is important for the exercise, since giving more time would engender further assumptions and speculation about how firms would choose to behave regarding when to begin the process of compliance, which would affect the wage pressures in the labor market; the short-time assumption implies that all firms move to comply at about the same time. Furthermore, this feature is why the model is based on just two periods with a decision process in between, instead of several periods of adjustment. Second, I assume that firms maintain their sales expectations (prior restriction) once they are informed of the intra-occupational gender wage gap restriction, arguing that they do not hold enough information to estimate how this novel policy will affect demand. In general, the model centers on how firms introduce changes into their payroll and make decisions about hiring and firing based on intra-occupational gender gap requirements, and it takes into account the unofficial gender discrimination that may occur based on the firms' interest on filling positions with a specific gender to strategically meet the restriction. The model captures the labor market pressure on wages in specific occupations deriving from firms' needs for specific professions and genders to comply with the restriction. It also captures the fact that costs may increase in such a way that some firms will fire all workers and cease operation, while other firms are able to offset the new costs associated with the gender gap restriction through wage reductions during the labor market's pressure on specific occupations and genders. On the other hand, the model is unable to capture the fact that firms may renegotiate wages once a worker has received an offer for another, better-paying job. Finally, the model does not incorporate workers' decision between employment and leisure.

Finally, while the results may vary in other countries, the circumstances that motivated the restriction and were key in how the restriction ultimately affected overall wage inequality through certain components may also be present in other countries. These include, first, the presence of gender wage gaps (intra-occupational within each firm), so that there is space to increase their contribution in reducing or slowing inequality, as shown in table 4; second, wage inequality among men (or women) in a specific occupation in a given firm (non-zero contribution of $\mathbf{W}_M$ or $\mathbf{W}_F$); and third, the presence of gender segregation (occupations within each firm that are composed of more than one worker but only one gender). Finally, even though the decomposition presents a clear way to decrease overall inequality through the gender gap component (via a restriction on the gap), there could be cases in which a gender gap restriction has the opposite effect to the one intended; this is the case of limiting the wage gender gap with an $\alpha = 2$ versus forcing an elimination of the gender gap with an $\alpha = 1$ in the case of Costa Rica.

Concluding Remarks

This article evaluates the impact of a mandatory restriction on intra-occupational gender wage gaps within firms. To do so, a novel wage inequality decomposition is developed to relate gender-specific components, like differences within gender, gender wage gaps, and occupational segregation, to the overall measure of inequality. This decomposition is then used to set the basis of a labor market that models the firms' and workers' reactions to the restriction. This allows quantifying the impact of the restriction on wage inequality and on occupational gender wage gaps in the economy.

First, the wage inequality decomposition shows that the inequality measure grew 67 percent between 2008 and 2018; most of the growth is explained by an increase in the gender segregation and dominance component (30 percent of inequality growth) and higher wage differences within men of the same occupation in each firm (26 percent). Additionally, for the study years, around 40 percent of the departments in the economy (occupations in each firm) violate the standard intra-occupational gender wage gap restriction of $\alpha = 2$ by having a gender average wage over two times the other gender average; and 94 percent of these departments initially violate the restriction because of occupational segregation (departments with more than one worker and only one type of gender present).

The quantitative exercise shows that the effect of the restriction on gender wage gaps in the economy's occupations is particularly important for those with a high concentration of one gender; in general, these gaps tend to deteriorate (move away from the equalization of averages) in favor of men for those occupations with little female participation, and they tend to deteriorate in favor of women if the occupation has low male participation. This deterioration of the gap is exacerbated when the restriction is relaxed, and it is lessened with a more severe restriction (in which gender wage averages must be the same in each firm's occupation). The model also shows that the constraint of $\alpha = 2$ produces a decrease in total inequality of about 10 percent. This reduction is mainly achieved through smaller wage differences among workers of the same gender in the same department. However, this overall inequality reduction dissipates and reverts into an increase if the restriction is tightened too much ($\alpha = 1$).

Finally, although this article highlights the potential benefits and disadvantages of imposing limits on the intra-occupational gender wage gaps within each firm, it also raises two interesting points that should be embraced in future research. First, the quantitative exercise for the restriction centered its attention on the firms' labor decisions, but an extension of the model and data regarding each firm's production could provide further insights into the effects on their performance after the gender wage gap regulation. Second, it was shown that forcing (by mandate) wage gender equality with an $\alpha = 1$ restriction (under which gender wage averages are equal) increases overall wage inequality; however, it would be useful to explore whether that also occurs when the gender wage gap is eliminated through a less disruptive event. For instance, Roussille (2021) reports that changing the way candidates give the ask salary before hiring (namely, a change from an empty box to prefilled options based on similar candidates) drove the gender bid gap (wage offered by firm) to zero, with no penalty on the number of bids received after the change.

References

Alfaro, Alonso, Santiago Campos, and Valerie Lankester. 2019. "Labor Demand Dynamics in Costa Rica." Working Paper 006. San José, Costa Rica: Central Bank of Costa Rica.

Alfaro, Alonso, Isabela Manelici, and José P. Vásquez. 2021. "The Effects of Joining Multinational Supply Chains: New Evidence from Firm-to-Firm Linkages." Available online at www.isabelamanelici.com.

Álvarez, Jorge, Felipe Benguria, Niklas Engbom, and Christian Moser. 2018. "Firms and the Decline in Earnings Inequality in Brazil." *American Economic Journal: Macroeconomics* 10 (1): 149–89.

Baker, Michael, Yosh Halberstam, Kory Kroft, Alexandre Mas, and Derek Messacar. 2019. "Pay Transparency and the Gender Gap." NBER Working Paper 25834. Cambridge, Mass.: National Bureau of Economic Research.

Barth, Erling, Alex Bryson, James Davis, and Richard Freeman. 2016. "It's Where You Work: Increases in the Dispersion of Earnings across Establishments and Individuals in the United States." *Journal of Labor Economics* 34 (S2): 67–97.

Bertrand, Marianne, Sandra E. Black, Sissel Jensen, Adriana Lleras-Muney. 2019. "Breaking the Glass Ceiling? The Effect of Board Quotas on Female Labour Market Outcomes in Norway." *Review of Economic Studies* 86 (1): 191–239.

Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz. 2010. "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors." *American Economic Journal: Applied Economics* 2 (3): 228–55.

Blatter, Marc, Samuel Muehlemann, and Samuel Schenker. 2012. "The Costs of Hiring Skilled Workers." *European Economic Review* 56 (1): 20–35.

Blau, Francine, and Lawrence Kahn. 2017. "The Gender Wage Gap: Extent, Trends, and Sources." *Journal of Economic Literature* 55 (3): 789–865.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa. 2019. "A Distributional Framework for Matched Employer Employee Data." *Econometrica* 87 (3): 699–739.

Brynin, Malcolm. 2017. "The Gender Pay Gap." Research Report 109. Manchester, U.K.: Equality and Human Rights Commission.

Card, David, Ana R. Cardoso, Jörg Heining, and Patrick Kline. 2018. "Firms and Labor Market Inequality: Evidence and Some Theory." *Journal of Labor Economics* 36 (S1): S13–70.

Card, David, Ana R. Cardoso, and Patrick Kline. 2016. "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women." *Quarterly Journal of Economics* 131 (2): 633–86.

Coghlan, Erin, and Sara Hinkley. 2018. "State Policy Strategies for Narrowing the Gender Wage Gap." Policy Brief (April). University of California, Berkeley, Institute for Research on Labor and Employment.

Cook, Cody, Rebecca Diamond, Jonathan Hall, John List, and Paul Oyer. 2021. "The Gender Earnings Gap in the Gig Economy: Evidence from Over a Million Rideshare Drivers." *Review of Economic Studies* (forthcoming).

Goldin, Claudia. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review* 104 (4): 1091–119.

Goldschmidt, Deborah, and Johannes F. Schmieder. 2017. "The Rise of Domestic Outsourcing and the Evolution of the German Wage Structure." *Quarterly Journal of Economics* 132 (3): 1165–217.

Hansen, Benjamin, and Drew McNichols. 2020. "Information and the Persistence of the Gender Wage Gap: Early Evidence from California's Salary History Ban." NBER Working Paper 27054. Cambridge, Mass.: National Bureau of Economic Research.

Hegewisch, Ariane, and Emma Williams-Baron. 2016. "The Gender Wage Gap by Occupation 2016 and by Race and Ethnicity." Working Paper C456. Washington, D.C.: Institute for Women's Policy Research.

Helpman, Elhanan, Oleg Itskhoki, Marc-Andreas Muendler, and Stephen J. Redding. 2016. "Trade and Inequality: From Theory to Estimation." *Review of Economic Studies* 84 (1): 357–405.

Jarosch, Gregor. 2021. "Searching for Job Security and the Consequences of Job Loss." NBER Working Paper 28481. Cambridge, Mass.: National Bureau of Economic Research.

Khitarishvili, Tamar, Lourdes Rodríguez-Chamussy, and Nistha Sinha. 2018. "Occupational Segregation and Declining Gender Wage Gap: The Case of Georgia." Policy Research Working Paper 8583. Washington, D.C.: World Bank.

Kim, Marlene. 2015. "Pay Secrecy and the Gender Wage Gap in the United States." *Industrial Relations: A Journal of Economy and Society* 54 (4): 648–67.

Magda, Iga, Jan Gromadzki, and Simone Moriconi. 2019. "Firms and Wage Inequality in Central and Eastern Europe." IZA Discussion Paper 12214. Bonn, Germany: Institute of Labor Economics.

Miller, Kevin, and Deborah J. Vaggins. 2018. "The Simple Truth about the Gender Pay Gap." Fall 2018 edition. Washington, D.C.: American Association of University Women.

Oaxaca, Ronald L., and Michael R. Ransom. 1994. "On Discrimination and the Decomposition of Wage Differentials." *Journal of Econometrics* 61 (1): 5–21.

OECD (Organization for Economic Cooperation and Development). 2017. *OECD Reviews of Labour Market and Social Policies: Costa Rica.* Paris: OECD Publishing.

———. 2020. *OECD Economic Surveys: Costa Rica 2020.* Paris: OECD Publishing.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman. 2018. "Distributional National Accounts: Methods and Estimates for the United States." *Quarterly Journal of Economics* 133 (2): 553–609.

Roussille, Nina. 2021. "The Central Role of the Ask Gap in Gender Pay Inequality." University of California, Berkeley.

Sarlo, Christopher A., Jason Clemens, and Joel Emes. 2015. "Income Inequality Measurement Sensitivities." Vancouver, B.C.: Fraser Institute.

Song, Jae, David J. Price, Faith Guvenen, Nicholas Bloom, and Till von Wachter. 2018. "Firming Up Inequality." *Quarterly Journal of Economics* 134 (1): 1–50.

Taber, Christopher, and Rune Vejlin. 2020. "Estimation of a Roy/Search/Compensating Differential Model of the Labor Model." *Econometrica* 88 (3): 1031–69.