LABOUR ECONOMICS
WORKING PAPERS

Working Paper Series
LEWP No. 2

Gender differences in the duration of sick leave: Economics or biology?

Martín-Román, A.L.
Moral, A.
Pinillos-Franco, S.

December 24

Available online at: https://www.aeet.eu/labour-economics-wp-2/

AEET
ASOCIACIÓN ESPAÑOLA
ECONOMÍA DEL TRABAJO

Gender Differences in the Duration of Sick Leave: Economics or Biology?

Ángel L. Martín-Román
Universidad de Valladolid
Global Labor Organization (GLO) Fellow
almartin@uva.es

Alfonso Moral
Universidad de Valladolid
amoral@uva.es

Sara Pinillos-Franco
Universidad Autónoma de Madrid
sara.pinillos@uam.es

Abstract
This study addresses the gender gap in workplace sick leave duration, focusing on the underlying economic and biological factors that contribute to this disparity. Using a novel methodological approach, we combine the stochastic frontier technique with an Oaxaca-Blinder-type decomposition to separate sick leave into medically justified and "opportunistic" days. Our analysis, based on detailed administrative data of workplace accidents in Spain, reveals that men and women recover at different rates for the same injuries, with biological differences explaining the majority of the observed gender gap. Additionally, we identify that men tend to use more sick leave days for reasons unrelated to health recovery. The findings offer valuable insights for policymakers and employers, providing an empirical foundation for targeted policies that reduce gender-based discrimination in the workplace and ensure fairer resource allocation. This research contributes to a deeper understanding of the gender gap in occupational health and offers implications for improving workplace equality.

Keywords: Gender Gap; Sick Leave Duration; Workplace Accidents; Stochastic Frontier Analysis; Occupational Health

JEL Codes: I12; I13; J16; J28; C21

Correspondence to Ángel L. Martín-Román
Ángel Martín-Román acknowledges the financial support from the Government of Spain Ministry of Science and Innovation under project PID2020-112509 GB-I00.

Ángel L. Martín-Román. Dpto. Análisis Económico, Facultad de Ciencias Sociales jurídicas y de la Comunicación. Universidad de Valladolid, Plaza universidad, 40005 Segovia, Spain. E-mail: almartin@uva.es
Alfonso Moral. Dpto. Análisis Económico, Facultad de Ciencias Sociales jurídicas y de la Comunicación. Universidad de Valladolid, Plaza universidad, 40005 Segovia, Spain. E-mail: amoral@uva.es
Sara Pinillos-Franco. Dpto. Análisis Económico, Facultad de Ciencias Económicas y Empresariales. Universidad Autónoma de Madrid, Calle Francisco Tomás y Valiente 5, Madrid, 28049, Spain. E-mail: sara.pinillos@uam.es

1. Introduction

The number of workplace accidents still reaches worrisome figures across Europe. Despite the number of fatal accidents has been reduced by approximately 70% in the period 1994-2018, more than 3,000 workers died due to a workplace accident in 2019, and more than 2.4 million workers suffered from a non-fatal workplace accident in Europe (Eurostat, 2023). After non-fatal accidents at the workplace, workers need an inactivity period to completely recover their health, and normally receive some type of benefit based on their wage to help them during this hard life event. Therefore, studying workplace accidents is not only a matter of health and safety, but also an economic issue concerning developed economies. Indeed, workplace accidents and occupational diseases entailed 3.3% of the European GDP in 2019 or, in other words, €460,000 million (European Commission 2021).

In the majority of Western countries, women take more sick days than men. Women take on average 7.6 sick days more than men in Europe, 3.1 days more in the US, and 5.2 days more in Canada (Ichino and Moretti 2009). This gender gap might entail additional discrimination against women if it is not deeply examined and fully understood (e.g., employers could avoid hiring women looking at these figures). The different prevalence of illnesses between men and women, sex differences in labor participation, or the different roles that men and women occupy in society are amongst the most popular theories to explain this gender gap. However, no study has empirically examined the reasons behind this gender difference accounting for the biological and economic drivers separately.

The aim of this paper is to answer the following research questions: What determines the differences observed between men and women on sick leave? Are these mainly the result of physiological factors or are they due to different behaviours between the sexes? Determining whether economics or biology is the main factor causing the differences has significant consequences for the implementation of efficient policies in the field of public health.

The contribution of this research work is twofold. On the one hand, we set the conceptual setting to analyze the underlying economic and biological factors behind the gender gap in workplace-related sick leaves. To do that, we develop a methodological approach that, as far as we know, has never been implemented in the way we do. On the other hand, we obtain fresh empirical evidence on the differences in sick leave duration between sexes that are extremely helpful for policymakers to take action. Since we address that part of sick leave due to physiological factors differently than the other part as a result of economic behaviour, we provide policymakers with an effective tool to implement targeted policies.

The methodology we use combines two different techniques that, to the best of our knowledge, have not been applied together in the way we do so far. We make use of the stochastic frontier approach to estimate the economic and the biological components of the duration of a workplace-related sick leave, in the same vein as Martín-Román and Moral (2017) and Martín-Román, Moral, and Pinillos-Franco (2024). Then, we estimate two different empirical relationships, one for male workers and another for female workers, to capture the idiosyncratic features of both sexes. Finally, we apply an Oaxaca-Blinder-type decomposition so as to obtain the parameters of interest to carry out the empirical exercise[1].

Our findings suggest that men and women recover at different rates for the same illness, i.e., their biological responses to face a specific injury differ and explain most of the observed gender gap. Additionally, we were able to split sick days depending on whether they are associated with physiological/medical reasons (standard duration) or they are "opportunistic", i.e., additional days that a worker might enjoy maintaining his/her sick leave status and not returning to work even if he/she is completely recovered in health terms. This precision analysis allows to know the optimal usage of these sick days for men and women separately or, in other words, to grasp whether the full duration of sick leave is justified by

[1] See Oaxaca (1973) and Blinder (1973).

medical reasons, or there also exists an unjustified part (we attach it to an opportunistic behavior from the worker side). The results suggest that men use more sick days than women for other commitments different from the pure health recovery. This is the first study in identifying and quantifying the different components behind the gender gap in the duration of sick leave, after its estimation with a frontier stochastic technique. This allows us to precisely know the type of days hidden during a worker's sick leave.

We used the universe of workplace accidents in Spain that contains exact information about the number of days a worker is on sick leave, the type of injury suffered, its severity, etc., and other personal information such as workers' age, sex, or occupational class. With this rich administrative data, we shed light on what is actually behind the gender gap in the duration of sick leave, supporting with empirical evidence some of the proposed theories in the literature. Additionally, these findings may have policy implications that help governments to allocate their scarce resources; employers to properly reward their employees; or reduce workplace discrimination for female workers. The results of this study might help governments decide to devote more resources to monitoring workers when they are on sick leave, or instead, invest more in security and safety at the workplace. Employers may reward low-absence employees with promotions and/or wage incentives (Markussen 2012), widening the gender wage gap if they only look at the raw gap in sick leave duration. Therefore, our findings might help them to properly reward workers that are more engaged with the company, that is, that effectively use their sick leave for pure health recovery and not for other commitments. Finally, the longer sick leave duration of women might be seen as an economic burden for companies, so it is necessary to alleviate this cost to favor gender equality in the labor market. Using a female wage subsidy financed by general taxation might be a solution, as proposed by Ichino and Moretti (2009). That is, the goal of this market intervention would be redistribution and not efficiency, transferring part of the absenteeism cost attached to biological reasons from women to men.

The paper is organized as follows. In section two, we present a review of the literature explaining the different hypotheses about the gender gap in the duration of sick leave.

Section three describes the Spanish institutional setting and the database employed. Section four presents the empirical strategy adopted in this research. Section five shows the main results, section six presents some policy implications, and section seven concludes.

2. Background

Gender differences in sick leave have been detected in the extensive (incidence of absences) and intensive (duration of absences) margins, and women are those who present worse sick leave outcomes. Women's odds of being on sick leave are higher than those of men (Markussen et al. 2011) and their recovery rates are slower, leading them to present longer durations of sick leave compared with males (Antczak and Miszczyńska 2021; Casini et al. 2013; Coutu et al. 2021; Fontaneda et al. 2019; Laaksonen et al. 2010; Mastekaasa 2014; Mastekaasa and Melsom 2014; Østby, Mykletun, and Nilsen 2018). In this paper, we focused our attention on the intensive margin, as we were able to distinguish between sick days attached to medical/physiological reasons, and sick days attached to behavioral factors.

One of the reasons for differences in the length of sick leave between men and women is that they do not present the same prevalence of injuries. Women present higher odds of suffering from musculoskeletal diseases compared to men, such as anterior cruciate ligament tears, multidirectional shoulder instability, ankle instability or osteoporosis, among others (Wolf et al. 2015). Besides, women are more likely to present daily disability illnesses such as rheumatism, anemia, thyroid, eczema, headaches or mental illnesses, whereas men present a higher prevalence of diseases related to survival such as cardiovascular diseases, stroke, lung and kidney diseases, or liver cirrhosis (Case and Paxson 2005; Macintyre, Hunt, and Sweeting 1996). This different prevalence of illnesses leads women to longer recovery periods after injury and surgery compared to men, and higher work disability rates that complicate their return to work (Coutu et al. 2021). This different prevalence of injuries is in line with our results, as we found a positive effect that leads women to present longer standard durations of sick leave and, consequently, to extend the gender gap. The analysis proposed in the methodological section identifies part of the observed gender gap in the sick leave duration with differences in the type of injuries suffered by men and women (dubbed as *composition effect)*.

The literature also identifies the existence of differences in anatomy and hormones between both sexes that lead to different periods of recovery (Case and Paxson 2005; Krenz and Strulik 2021; Martin-Roman and Moral 2016; Martín-Román et al. 2024; Mukuria et al. 2017; Oaxaca 1973; Spierdijk, van Lomwel, and Peppelman 2009). Women not only present higher morbidity rates compared to men due to their reproductive biology (Kananurak 2014) but also their anatomy might complicate their recovery after injury/surgery. For instance, the efficacy of bypass and the grafts after a cardiac surgery is lower among women due to smaller coronary arteries and lower surface area, raising their odds of returning hospital (Bechtel and Huffmyer 2020). Additionally, women are more likely to develop further diseases after surgery or traumatic events (e.g. a workplace accident) such as dysphoria, anxiety, and depression, complicating again their recovery (Freedman et al. 2002; Kempen et al. 2003; Modica et al. 2014; Oksuzyan, Gumà, and Doblhammer 2018). The decomposition applied in the empirical section also identifies the part of the gap explained by different recovery periods for the same illness between both sexes (referred to as *biological effect and reference injury effect/average injury effect)*.

Another reason to explain the gap in the sick leave duration between men and women is the different roles traditionally attached to men and women in society. Employed women devote more hours to unpaid work than men, implying a double burden for them, translating into poorer health outcomes and delays in their return to work (Côté and Coutu 2010). For instance, a woman at home recovering from a work accident will probably perform different domestic chores (e.g., cooking, cleaning, childcare, etc.) while recovering from her injury, implying a longer recovery period. Something similar happens when explaining the widening of the gender earnings gap. Despite men present greater ability than women to occupy highpaid positions and better facility to prosper within companies, when women present higher family commitments, these two factors are intensified and, consequently, the gender earnings gap widens (Goldin et al. 2017). Finally, although women's situation changes when they are on sick leave, the allocation of household tasks might not change in traditional families, as gender norms are more relevant in explaining the distribution of domestic chores than other factors such as time availability (Farré et al. 2022). Indeed, a recent paper by Depalo and Pereda-Fernández (2023) found that female workers experienced an increase in worked hours or working remotely during the COVID-19 pandemic, which may be due to an increase in hours devoted to household tasks. This source of differences may be behind another of the effects identified in the methodological section (dubbed as *induced inefficiency effect*).

A final group of factors that may explain the gender gap in sick leave duration is due to differences purely attached to anatomically be a man or a woman in society. Men and women do not present the same personality and behavior, which can affect the length of sick leave. For instance, women tend to be more neurotic and extrovert than men (Weisberg, De Young, and Hirsh 2011) and individuals who display these personality traits present higher odds of increasing sick leave (Løset and Soest 2022). Additionally, women exhibit more *cautious* behavior, which may explain why this effect was positive. Croson and Gneezy (2009) already documented that women are more risk-averse, less competitive, and more context-sensitive to make their decisions compared to men. Besides, women also present higher social risk aversion and inequality aversion compared to their male counterparts (Friedl, Pondorfer, and Schmidt 2020). Thus, these well-documented gender differences in behavior (Schünemann, Strulik, and Trimborn 2017), affect individuals' health outcomes differently (Bauer, Göhlmann, and Sinning 2007; Nelson 2014). These latter explanations for the gender difference in absence days can also be identified in the proposed decomposition (termed *behavioral inefficiency effect)*.

3. Institutional setting

All Spanish workers are entitled to receive economic compensation when they are on sick leave, but the amount of this benefit depends on whether their temporary incapacity (TI) was due to a work-related (occupational) accident or not. During the TI of the worker, the Social Security Administration covers the medical expenses of the injured worker for 365 days, potentially extendable 180 days more. If a worker is not able to recover his/her health after this period, the National Institute of Social Security decides whether the worker is transferred to the permanent disability system or receives a medical discharge.

Injured workers after a work-related accident receive 75% of the reference wage[2] the day after his/her general practitioner issues the sick leave certificate, and this benefit is paid by the mutual insurance company. However, if the worker suffers from a non-work-related accident, the first three days after the accident, the worker receives no amount of sick leave. From the 4th to the 20th day the worker receives 75% of the reference wage and this is paid by the employer until the 15th day, finally, from the 20th day to onwards, the injured worker receives 60% of the reference wage and this is paid by the Social Security Administration (from day 16th and onwards).

In this paper, we analyze a database of Spanish private-sector workers who are unable to work due to a work-related accident. The Social Security Law (SSGL) of 1994 regulates the amount of sick leave benefits for this type of accident and it has not undergone any regulatory change to date. Additionally, the SSGL establishes the concept of occupational accident (Art. 156) as any bodily injury suffered by the worker due to or as a consequence of developing his/her paid job, i.e., this includes all accidents suffered by

[2] This is calculated from the wage that the worker has earned in the last month before the accident. This reference wage has upper limits that are established in the State General budget each year and they are equal for all professional categories and contingencies (Art. 148 of the SSGL); and lower limits, whose amount depends on the minimum wage of each year increased by one-sixth. The estimation of the contributory base has been unchanged during the period analysis of our study.

employees within their workplace, to and from work (*in itinere* accidents) and/or owing to carrying out a union position or any work-related task demanded by the employer (*in mission* accidents). Although this definition excludes occupational diseases (Art. 157 of the SSGL) and common illnesses (Art. 158 of the SSGL), there are some illnesses considered occupational accidents as well. Precisely, there are three types of illnesses excluded from the term "occupational disease" that are considered work-related accidents: diseases in the strict sense, pre-existing or latent diseases, and intercurrent diseases. The diseases in the strict sense are those contracted by workers as a result of carrying out their work duties [Art. 156.2.e of the SSGL]. The pre-existing or latent diseases correspond to those that the worker already holds and that have been aggravated as a consequence of developing his/her job [Art. 156.2.f of the SSGL]. Finally, intercurrent diseases refer to those that are not directly related to the accident but have been exacerbated due to the accident or contracted during the recovery process [Art. 156.2.g of the SSGL].

4. Data

We used the universe of workplace accidents in Spain, provided by the Statistics of Accidents at Work (SAW). This is an annual administrative register of all occupational accidents that occurred in Spain that includes rich information about injured workers (age, sex, occupational class, injured part of the body, severity of the injury, etc.) and conditions of the suffered accident such as the characteristics of the company.

For the estimations, we used a dataset for the period 2011-2019 restricted to private sector workers who work on a full-time basis. This restriction is due to the self-employed follow a different law scheme, and part-time workers receive a lower amount of sick leave benefits that may make them behaviorally react differently when they are on sick leave (e.g., they might try to return to work earlier than full-time workers as their household income loss during their sick leave might be higher). We did not consider fatal accidents either as, unfortunately, it is not possible to monitor any workers' behavior after them. We also eliminated "*in itinere*" accidents because their incidence is different from the rest of occupational accidents[3]. Additionally, we removed some detected registered errors such as ages incompatible with labor market or compensations out of the legally established limits. Our final database consisted of 3,916,249 injured workers due to a work-related accident.

5. Methodology

This paper combines two different empirical approaches in a novel way. As far as we know, this is the first time this has been done. Firstly, we estimate the duration of sick leave to identify the part explained by physiological/medical factors and another that may be attached to an opportunistic behavior from the worker's side. Secondly, we use decomposition techniques to analyze which factors explain the mean differences between sick leave durations of men and women.

5.1. Stochastic frontier estimation.

To analyze the duration of sick leave, we use an approach based on the stochastic frontier technique (Martín-Román and Moral 2014, 2017; Martín-Román, Moral, and Pinillos-Franco 2024). Following this method, after an injury, there exists a recovery period only attached to medical or physiological factors. That period is identified as 'standard duration' (D_i^s) and it is a lower boundary that can be defined as follows:

$$d_i^s = X_i \beta + v_i \quad \text{with} \quad d_i^s = ln(D_i^s) \qquad (1)$$

With X_i a vector of characteristics, β a vector of coefficients and v_i a random error of mean zero and variance σ_v^2.

[3] In general, men present a higher incidence of occupational accidents than women but, in the case of *in itinere* accidents this situation is reversed, requiring further analyses that are far from the aims of this study. Indeed, in 2023, a 54% of *in itinere* accidents in Spain were suffered by female workers, whereas the remaining 46% by male workers (SAW, 2023).

However, insurers do not normally perceive this duration, as they only have information about the actual sick leave duration (D^r). Hence, this actual duration is not only a consequence of medical and physiological factors but also of the worker's capacity to increase his/her period of recovery. It is thus a problem of asymmetric information which generates opportunistic behaviors from workers covered by accident insurance. This increase in duration might be related to an inefficient situation when monitoring sick leaves.

In formal terms, the actual duration results from adding to the standard duration another random disturbance (u_i) with a positive mean and variance σ_u^2. It can be expressed as follows:

$$d_i^r = d_i^s + u_i \text{ with } d_i^r = ln(D_i^r) \qquad (2)$$

From equations (1) and (2), the final model can be obtained:

$$d_i^r = X_i\beta + v_i + u_i \qquad (3)$$

Two technical aspects are needed to conduct this estimation. On the one hand, it is worth mentioning that equation (3) describes a composed error model, so an OLS estimation would generate inconsistency in the constant term and impede splitting the variance of both disturbances. In these cases, a maximum likelihood estimation is more accurate when using the stochastic frontier technique. Specifically, it is a low or cost frontier due to the standard duration (frontier) being below the actual one.

On the other hand, having an additional disturbance also implies supposing for it a statistical distribution. Some examples of distributions that may be employed are Half-Normal (Aigner et al. 1977), Exponential (Meeusen and van Den Broeck 1977), Truncated Normal (Stevenson 1980), or Gamma (Greene 1980b, 1980a).

Finally, the frontier estimation also allows us to calculate the level of inefficiency of any sick leave through the following expression:

$$EF_i=\frac{f(X_i\beta)exp(v_i+u_i)}{f(X_i\beta)exp(v_i)}=exp(u_i) \qquad (4)$$

5.2. Decomposition of the gender gap

Once stochastic frontiers estimations have been obtained separately for men and women, the next step is to decompose the difference between the average duration of leave in each group into its different components.

To conduct this type of analysis, the seminal works by Oaxaca (1973) and Blinder (1973) developed a methodology that has been widely applied in economic literature, particularly in the case of wage discrimination. In its original version, the decomposition assumes a linear relationship between the dependent variable (the duration of sick leave denoted by *d*) and the explanatory variables (*X*), which must also be independent of the error term (ϵ).[4]

At this point, we can follow Yun's approach (Yun 2004, 2005) to perform the decomposition of stochastic frontier estimation for two reasons. Firstly, it allows identification issues to be addressed in the detailed decomposition associated with the use of dummy variable groups in estimation. Secondly, it puts forward a generalization for any functional relationship that can be extended to the frontier estimation applied in this study (See **Appendix 1** for technical details). The other option is to use a reference injury determined by the dummy variables removed from the estimation to avoid multicollinearity.

[4] The literature also includes extensions of these types of decompositions for nonlinear models where the dependent variable (continuous or discrete) is a function of a linear combination of regressors. Some examples of these studies are Even and Macpherson (1990), Fairlie (1999, 2005), and Nielsen (1998) for logit and probit models, or Ham, Svejnar, and Terrell (1998) for duration models.

In formal terms, the decomposition proposed by Oaxaca and Blinder starts with two estimations of the dependent variable (i.e., *d* in this case), one for each of the groups to be compared (female and male workers, or *f* and *m* in this case), as expressed below:

$$d_i^h = \beta_o^h + \sum_{k=1}^{K} X_{ik}\beta_k^h + \varepsilon_i^h \quad \text{with h: m, f} \qquad (5)$$

Where β_o^h and β_k^h are the coefficients resulting from the estimates in each population group, X_{ik} is the corresponding vector of K explanatory variables, and $E\left(\varepsilon_i^h | X_{ik}\right)=0$. In this point, the mean predicted values can be obtained for each group and, with them, a counterfactual estimation. Such a counterfactual results from using the mean values of the variables in the women's group together with the coefficients obtained for the men's group. Its formal specification would be as follows:

$$\overline{d^{fm}} = \widehat{\beta_o^m} + \sum_{k=1}^{K} \overline{X_k^f}\, \widehat{\beta_k^m} \qquad (6)$$

By adding and subtracting the counterfactual within the difference in means of both groups, the duration gap can be decomposed as follows:

$$\overline{d^f} - \overline{d^m} = \widehat{\beta_o^f} + \sum_{k=1}^{K} \overline{X_k^f}\widehat{\beta_k^f} - \widehat{\beta_o^m} - \sum_{k=1}^{K} \overline{X_k^m}\widehat{\beta_k^m} + \widehat{\beta_o^m} + \sum_{k=1}^{K} \overline{X_k^f}\,\widehat{\beta_k^m} - \widehat{\beta_o^m} - \sum_{k=1}^{K} \overline{X_k^f}\,\widehat{\beta_k^m} = \underbrace{\left(\widehat{\beta_o^f} - \widehat{\beta_o^m}\right) + \sum_{k=1}^{K} \overline{X_k^f}\left(\widehat{\beta_k^f} - \widehat{\beta_k^m}\right)}_{\textit{Unjustified effect}} + \underbrace{\sum_{k=1}^{K}\left(\overline{X_k^f} - \overline{X_k^m}\right)\widehat{\beta_k^m}}_{\textit{Justified effect}} \qquad (7)$$

The first component (unjustified) reflects that similar characteristics affect each group differently. For the exercise proposed in this study, this component would imply, for example, that similar injuries lead to different recovery processes in men and women. As for the second term, it is considered justified because it is reasonable to assume that different values of the explanatory variables generate differences in the dependent variable. In other words, if the severity level of the injury or the ages of the workers are not the same, the duration of leave may also differ.

Figure 1 shows a graphical representation of the decomposition, although with certain underlying assumptions. Firstly, a single explanatory variable (*X*) is considered, which determines the duration of the sick leave. It is also assumed that the duration of leave increases with the values of this variable in both groups at a constant rate $\left(\beta^h\right)$ starting from an initial value indicated by the intercept $\left(\beta_o^h\right)$. Finally, it is assumed that all components of the decomposition operate in the same direction, meaning that the group with a higher initial leave duration also experiences a greater increase in their recovery period as X_i increases. The blue line represents the estimated duration for males, the black line for females, and the green ones are auxiliary lines used to identify the different effects. Both the gap to be explained and the obtained components are shown in bold, and if they are on the right (left) of the bracket, they are considered to have a positive (negative) value.

[Figure 1]

As an alternative to **Figure 1**, **Figure 2** presents another version of the same decomposition where the difference in the intercept terms operates in the opposite direction to the rest of the unjustified component, while maintaining the remaining assumptions.

[Figure 2]

Based on this well-known decomposition, we can now move on to the case at hand by incorporating the modifications brought by the estimation of stochastic frontiers into the model. As previously explained, the first change generated by the frontier estimation is the inclusion of an additional term associated with inefficiency. Furthermore, if we assume the linearization of the objective function through logarithmic transformation, the individual estimations are expressed as follows:

$$d_i^h = \beta_o^h + \sum_{k=1}^{K} X_{ik}\beta_k^h + u_i^h + \varepsilon_i^h \quad \text{with h:m, f} \qquad (8)$$

Where a new term $\left(u_i^h\right)$ is included, which corresponds to a random disturbance that always takes a positive value and captures the part of sick leave associated with worker behavior. According to this, the new difference in means could be expressed as follows:

$$\overline{d}^f - \overline{d^m} = \widehat{\beta_o^f} + \sum_{k=1}^{K} \overline{X_{ik}}\widehat{\beta_k^f} + \overline{u}^f - \widehat{\beta_o^m} - \sum_{k=1}^{K} \overline{X_{ik}}\widehat{\beta_k^f} - \overline{u^m} \qquad (9)$$

Following this same line of reasoning, the corresponding counterfactual construction that determines the expected duration for women if the variables affected them in the same way as men would be:

$$\overline{d^{fm}} = \widehat{b_o^m} + \sum_{k=1}^{K} \overline{X}_k^f \widehat{\beta_k^m} + \overline{u^{fm}} \qquad (10)$$

Where $\overline{u^{fm}}$ represents the expected average duration associated with the inefficiency that women would have if the estimated coefficients for the equation of men are applied to them. By adding and subtracting this counterfactual within the difference in means of both groups, we obtain the following expression for the now modified decomposition:

$$\overline{d}^f - \overline{d^m} = \left(\widehat{\beta_o^f} - \widehat{\beta_o^m}\right) + \sum_{k=1}^{K} \left(\overline{X}_k^f - \overline{X_k^m}\right)\beta_k^m + \sum_{k=1}^{K} \overline{X}_k^f \left(\widehat{\beta_k^f} - \widehat{\beta_k^m}\right) + \left(\overline{u}^f - \overline{u^{fm}}\right) + \left(\overline{u^{fm}} - \overline{u^m}\right) \qquad (11)$$

[Figure 3]

Figure 3 shows this decomposition graphically. Now, the solid lines (blue and black) represent the standard duration values for men and women, respectively. The dashed lines (blue and black) refer to the estimated total sick leave duration for each group. Therefore, the vertical difference between them reflects the inefficiency that men and women exhibit in the labor market. As in the previous graphs, it is assumed that there is a single explanatory variable positively related to the standard duration of sick leave. It is also considered that the difference in intercepts operates in the same direction as the rest of the unjustified component (**Figure A1** in **Appendix 2** shows the decomposition when intercept differences act in the opposite direction). Furthermore, an additional assumption is included: in this case, it is assumed that the inefficiency component increases as the sick leave duration increases, and thus, the difference between the solid and dashed lines becomes larger. This assumption is sensible when considering efficiency as the ratio between the actual sick leave duration and the minimum expected duration for the recovery of a specific work-related accident. In such circumstances, the longer the sick leave, the more days associated with inefficiency.

After obtaining the mathematical expression of the decomposition and its graphical representation, the next step is to provide theoretical content to each of its components:

- $\left(\widehat{\beta_o^f}-\widehat{\beta_o^m}\right)$: This component can be interpreted in two ways. If the normalized version of the estimation proposed by Yun (2005) is used, it could be interpreted as the difference attributable to the fact that the sick leave generated by an average injury is different for men and women (**average injury effect**). If the normalized regression is not used, it would measure the different recovery rates between men and women after a reference injury, which is determined by the dummy variables removed from the estimation to avoid multicollinearity (**reference injury effect**).

- $\sum_{k=1}^{K}\left(\overline{X_k^f}-\overline{X_k^m}\right)\beta_k^m$: This term refers to the fact that men and women may have different physiological characteristics and also experience different types of injuries. The literature considers these differences as justified because differences in characteristics should generate differences in sick leave duration. In our case, we will refer to it as the **composition effect**, which indicates that men and women may have different characteristics or experience different accidents.

- $\sum_{k=1}^{K}\overline{X}_k^f\left(\widehat{\beta}_k^f\text{-}\widehat{\beta}_k^m\right)$: In the canonical decomposition, this term is considered as the unjustified difference[5]. However, in the present case, it can be given another interpretation. Specifically, if men and women have different coefficients for the same characteristics, it implies a different recovery period for the same injury depending on the sex of the injured person. This result could be associated with the different biology of both sexes justifying distinct sick leave durations. For this reason, we dub it the **biological effect**. Both **average injury effect** and **reference injury effect** can also be considered biological effects because they involve differences in duration between men and women for the same injury.

- $\left(\overline{u^f}\text{-}\overline{u^{fm}}\right)$: This component captures what we consider a purely statistical outcome, and we dub it **induced inefficiency effect** accordingly. When applying the coefficients from one group to the other leads to a modification in their standard duration and, consequently, their inefficiency. Hence, this creates a difference between the average inefficiency of one group and the inefficiency that would happen if the coefficients of the other group were applied to it.

- $\left(\overline{u^{fm}}\text{-}\overline{u^m}\right)$: This is what we can deem the true **behavioral inefficiency effect**. We apply the same coefficient specification (in this case, that from the male group) to both men and women and compare the values we obtain for the inefficiency measurement in both cases. Therefore, the differences between these magnitudes would only be a consequence of the sex of the injured victims.

[5] This term was originally considered as a wage discrimination measurement in the early literature.

6. Results

As it is explained before, to obtain the mean difference expressed in Equation 9 it is needed to estimate Equation 3 for men and women separately. We conducted two different regressions to calculate the standard duration of sick leave. First, to correct the potential multicollinearity and identification problems that arise from using dummy variables (i.e., leave one of them out from the model as a reference category), we calculated a normalized regression (see **Appendix 1**) for men and women separately. **Table 1** depicts the results of that regressions. Within each group, the first column shows the estimated coefficients, column two its significance, and column three the normalized coefficients. It can be observed that, for both sexes, injuries such as fractures, traumatic amputations, multiple injuries, and heart attacks implied longer sick leave durations (positive coefficient) than other type of injuries such as superficial injuries, burns, or choking (negative coefficient.). The duration of sick leave was also longer for men and women when the injured part of the body was the neck, the shoulder, the arm, the leg, the ankle, or it affected multiple parts of the body. Additionally, if the accident involved hospital care, hospitalization, a serious injury, and/or was a relapse of a previous injury, the duration of sick leave lasted more time for both sexes. The same was also true for workers from the age of 40 and manual workers.

[Table1]

After estimating the duration of sick leave for male and female workers, the gender difference in the duration of sick leave is decomposed as expressed in Equation 11. **Table 2** shows the decomposition of the gender gap in percentage and number of days. The results depict that the difference in coefficients (i.e., our labelled *biological effect*) is the one that contributes the most to explain the gender gap in sick leave duration (208%), followed by the c*omposition effect* (32%), that is due to the different characteristics that men and women present at the baseline, the *average injury effect* (20%) and the *behavioral inefficiency effect* (18%). The only effect operating in the opposite direction, that is, contributing to narrowing the gender gap, was the *induced inefficiency effect,* i.e., the difference between women's inefficiency and women's inefficiency estimated after applying male coefficients.

[Table 2]

Although the normalized regressions to obtain the duration of sick leave provide a single estimation, the interpretation of results might be difficult as the reference category is an "average injury" that must be interpreted as a notional concept which does not exist in real life. Therefore, the duration of sick leave is also estimated using a regular regression. To do so, the first step is to identify a comparable reference group to interpret the results. This reference group is selected by the combination of the injury and injured part of the body that appears in the most records within our database, and this was the case of "ankle sprains" with more than 180,000 registers. It is also selected that the injury is minor, neither with hospitalization nor hospital care, that this is not the consequence of a recurrent injury, and that the worker is between 30 and 40 years old and belongs to the low-skill professional category.

Table 3 shows the results of the regular regression for men and women separately. The estimated coefficient and its significance are depicted for both groups. Suffering from a fracture, dislocation, traumatic amputation, multiple injuries, or a heart attack leads to longer sick leave durations compared to suffering from a sprain for men and women alike. This is also true for women suffering from psychological trauma and men suffering from concussions. When the injured part of the body is the neck, shoulder, arm, wrist, or legs, or it affects multiple parts, the duration of sick leave is longer than when the ankle is the injured part. Again, manual and old workers present longer sick leave durations compared to unskilled workers aged from 30 to 40 respectively. Finally, injuries that required hospital care, hospitalization, were serious and/or were a relapse of a previous injury, implied longer sick leave durations compared to a slight injury that did not require hospitalization nor hospital care and was not a relapse of a previous injury.

[Table 3]

Once estimated the duration of sick leave for both groups, the observed gender gap is decomposed (**Table 4**). As previous results presented in **Table 2**, the differences in the standard duration of sick leave explain most of the gap, and all identified effects operate in the same direction. However, after estimating the duration of sick leave with a normal regression and decomposing the gender gap, the allocation of days changed in the unjustified part of the gap (i.e., our reference injury and biological effects). Now, the part that explains the most the gender gap was the difference in the constant terms, i.e., the *reference injury effect* (216%), followed by the *composition effect* (32%), the *behavioral inefficiency effect* (18%), and the *biological effect* (12%). Again, the *induced inefficiency effect* contributed to reducing the observed gender gap (-179%).

[Table 4]

As a robustness check, a normalized regression is also conducted including age and its square as continuous variables, instead of using categorical variables. Results are depicted in **Table A1** and **Table A2** in **Appendix 3**. The coefficient of age in the frontier estimation is positive, indicating that as workers grow older, their probability of presenting a long standard duration of sick leave is higher (**Table A1**). The rest of the normalized coefficients are similar to those obtained in **Table 1**. However, when decomposing the gender gap, we observe that not only the *statistical effect* but also the *average injury effect* contributed to reducing the gap (**Table A2**). Additionally, the effect that explained the most the gender gap was the *biological effect* (296%), followed by the *composition effect* (34%), and the *behavioral inefficiency effect* (17%).

7. Policy implications

Our paper provides useful evidence for decision-making at the governmental and firm level. Both institutions devote time and resources to monitor and compensate absent workers during their sick leave, so it is essential for them to 1) know whether injured workers strictly use their absent days for pure health recovery and not for other commitments, and 2) gain insight into the existent gender gap in sick leave duration to compensate male and female workers accordingly. In this sense, the findings might help governments decide to allocate their scarce resources to safety and prevention at the workplace or, instead, to monitor absent workers during their recovery.

Additionally, considering the main drivers behind the gender gap in the duration of sick leave in Spain, public administration might compensate female workers with a wage subsidy financed by general taxation. The different period of recovery for the same injury ("biological effect") is not a human choice and individuals should not be discriminated by the manner they were born. Thus, this type of subsidy might help to transfer part of this unavoidable cost from the most affected part in biological terms (female workers) to the least affected one (male workers) without adding negative discrimination in the job market.

Our findings also help firms to compensate their workers properly. On the one hand, firms might know which workers are truly engaged with them through the optimal use of sick leave. That is, workers who only devote absent days to health recovery and return to work in due time should be rewarded for their commitment to the firm, regardless of the actual duration of their sick leave. On the other hand, firms may also reward workers' efforts to self-security. Among high-risk jobs, the seriousness of an occupational accident tends to be higher than among low-risk jobs and, consequently, the number of absent days after this type of injury is greater. Then, those workers that invest in their security to avoid serious injuries should be compensated as well (this does not mean that they have to spend their economic resources but instead, that they present cautious and safe behavior at work – e.g., a proper driving or responsible use of dangerous substances or machinery). Normally, firms compensate with wage premiums for the fact of assuming risks at work from the employee's side, although they might be compensated for reducing this risk on their own as well (Guardado and Ziebarth 2019). This can make work safer if companies are unable to provide adequate security for workers, or if it is more cost-effective for workers to invest in their safety.

8. Conclusions

To our knowledge, this is the first study that empirically examines the effects that are behind the observed gender gap in the duration of sick leave. Applying a novel approach that allows us to distinguish between medical/physiological and behavioral absent days, we were able to precisely identify the proportion of days that are due to one of each reason. This outcome is important since we are not only quantifying the gender gap but also providing an economic meaning to the whole gap without leaving any part "unjustified", as it is commonly observed in studies that use decomposition techniques.

Additionally, despite the results being sensitive to different specifications (the proportion of days explained by each of the identified components changes when using different methods and/or variables), we observe that biological differences between men and women matter the most in explaining the gender gap in the duration of sick leave in general. This distinction is essential to avoid the gender discrimination that women may suffer in the labor market if the issue is not adequately addressed.

Although the rest of the identified effects explain a lower proportion of the observed gender gap, they are not negligible and should also be considered for policymaking. Non-observable characteristics at baseline are included in the unexplained part of the duration of sick leave (i.e., what we attached to an opportunistic behavior from workers' side to further extend their leave), and the estimated gap of this part may hide further discrimination for women.

Spanish society has experienced significant progress toward gender equality in recent years, but there persist traditional gender norms that may be reflected in the results. Women are still the main responsible of domestic duties and dependent care, and this role is not normally altered when their situation changes (e.g., when they become unable to work out of home due to a workplace accident). Thus, women are more likely to experience double pressure in opposite directions when they are on leave. On one side, employers demand their injured female workers return to work as soon as possible, whereas, on the other side, families may require their female members to spend more time at home.

Therefore, addressing the gender differences in the duration of sick leave requires a combination of public policies considering all the effects identified in this paper and the context in which these policies are developed.

References

Aigner, Dennis, C. A. Kno. Lovell, and Peter Schmidt. 1977. “Formulation and Estimation of Stochastic Frontier Production Function Models.” *Journal of Econometrics* 6(1):21–37.

Antczak, Elżbieta and Katarzyna M. Miszczyńska. 2021. “Causes of Sickness Absenteeism in Europe—Analysis from an Intercountry and Gender Perspective.” *International Journal of Environmental Research and Public Health 2021, Vol. 18, Page 11823* 18(22):11823.

Bauer, Thomas, Silja Göhlmann, and Mathias Sinning. 2007. “Gender Differences in Smoking Behavior.” *Health Economics* 16(9):895–909.

Bechtel, Allison J. and Julie L. Huffmyer. 2020. “Gender Differences in Postoperative Outcomes After Cardiac Surgery.” *Anesthesiology Clinics* 38(2):403–15.

Blinder, Alan S. 1973. “Wage Discrimination: Reduced Form and Structural Estimates.” *The Journal of Human Resources* 8(4):436.

Case, Anne and Christina H. Paxson. 2005. “Sex Differences in Morbidity and Mortality.” *Demography* 42(2):189–214.

Casini, Annalisa, Isabelle Godin, Els Clays, and France Kittel. 2013. “Gender Difference in Sickness Absence from Work: A Multiple Mediation Analysis of Psychosocial Factors.” *European Journal of Public Health* 23(4).

Côté, Daniel and Marie France Coutu. 2010. “A Critical Review of Gender Issues in Understanding Prolonged Disability Related to Musculoskeletal Pain: How Are They Relevant to Rehabilitation?” *Disability and Rehabilitation* 32(2):87–102.

Coutu, Marie-France, Marie-José Durand, Daniel Coté, Dominique Tremblay, Chantal Sylvain, Marie-Michelle Gouin, Karine Bilodeau, Iuliana Nastasia, and Marie-Andrée Paquette. 2021. “How Does Gender Influence Sustainable Return to Work Following Prolonged Work Disability? An Interpretive Description Study.” *Journal of Occupational Rehabilitation 2021* 1–18.

Depalo, Domenico and Santiago Pereda-Fernández. 2023. “The Differential Impact by Gender of the Covid-19 Pandemic on the Labor Outcomes of Older Adults.” *SERIEs* 14:503–53.

European Commission. 2021. *EU Strategic Framework on Health and Safety at Work 2021-2027 Occupational Safety and Health in a Changing World of Work*.

Even, William E. and David A. Macpherson. 1990. "Plant Size and the Decline of Unionism." *Economics Letters* 32(4):393–98.

Fairlie, R. W. 1999. "The Absence of the African-American Owned Business: An Analysis of the Dynamics of Self-Employment." *Journal of Labor Economics* 17(1).

Fairlie, R. W. 2005. "Entrepreneurship and Earnings among Young Adults from Disadvantaged Families." *Small Business Economics* 25:223–36.

Farré, Lídia, Yarine Fawaz, Libertad González, and Jennifer Graves. 2022. "Gender Inequality in Paid and Unpaid Work During Covid-19 Times." *Review of Income and Wealth* 68(2):323–47.

Fontaneda, Ignacio, Miguel A. Camino López, Oscar J. González Alcántara, and Dale O. Ritzel. 2019. "Gender Differences in Lost Work Days Due to Occupational Accidents." *Safety Science* 114:23–29.

Freedman, Sara A., Natali Gluck, Rivka Tuval-Mashiach, Dalia Brandes, Tuvia Peri, and Arieh Y. Shalev. 2002. "Gender Differences in Responses to Traumatic Events: A Prospective Study." *Journal of Traumatic Stress* 15(5):407–13.

Friedl, Andreas, Andreas Pondorfer, and Ulrich Schmidt. 2020. "Gender Differences in Social Risk Taking." *Journal of Economic Psychology* 77:102182.

Goldin, Claudia, Sari Pekkala Kerr, Claudia Olivetti, and Erling Barth. 2017. "The Expanding Gender Earnings Gap: Evidence from the LEHD-2000 Census." *American Economic Review* 107(5):110–14.

Greene, William H. 1980a. "Maximum Likelihood Estimation of Econometric Frontier Functions." *Journal of Econometrics* 13(1):27–56.

Greene, William H. 1980b. "On the Estimation of a Flexible Frontier Production Model." *Journal of Econometrics* 13(1):101–15.

Guardado, José R. and Nicolas R. Ziebarth. 2019. "Worker Investments in Safety, Workplace Accidents, and Compensating Wage Differentials." *International Economic Review* 60(1):133–55.

Ham, John C., Jan Svejnar, and Katherine Terrell. 1998. "Unemployment and the Social Safety Net during Transitions to a Market Economy: Evidence from the Czech and Slovak Republics." *American Economic Review* 88(5):1117–42.

Ichino, Andrea and Enrico Moretti. 2009. "Biological Gender Differences, Absenteeism, and the Earnings Gap." *American Economic Journal: Applied Economics* 1(1):183–218.

Kananurak, Papar. 2014. "Healthcare Use and Voluntary Health Insurance after Retirement in Thailand." *Applied Health Economics and Health Policy* 12(3):299–313.

Kempen, G. I. J. M., R. Sanderman, W. Scaf-Klomp, and J. Ormel. 2003. "Gender Differences in Recovery from Injuries to the Extremities in Older Persons. A Prospective Study." *Disability and Rehabilitation* 25(15):827–32.

Krenz, Astrid and Holger Strulik. 2021. "The Impact of Menstruation Hygiene Management on Work Absenteeism of Women in Burkina Faso." *Economics & Human Biology* 43(101067).

Laaksonen, Mikko, Arne Mastekaasa, Pekka Martikainen, Ossi Rahkonen, Kustaa Piha, and Eero Lahelma. 2010. "Gender Differences in Sickness Absence – the Contribution of Occupation and Workplace." *Scandinavian Journal of Work, Environment & Health* 36(5).

Løset, Gøril Kvamme and Tilmann von Soest. 2022. "Big Five Personality Traits and Physician-Certified Sickness Absence:" *European Journal of Personality*.

Macintyre, Sally, Kate Hunt, and Helen Sweeting. 1996. "Gender Differences in Health: Are Things Really as Simple as They Seem?" *Social Science & Medicine* 42(4):617–24.

Markussen, Simen. 2012. "The Individual Cost of Sick Leave." *Journal of Population Economics* 25(4):1287–1306.

Markussen, Simen, Knut Røed, Ole J. Røgeberg, and Simen Gaure. 2011. "The Anatomy of Absenteeism." *Journal of Health Economics* 30(2):277–92.

Martín-Román, Ángel L. and Alfonso Moral. 2014. "Differences between Spanish and Foreign Workers in the Duration of Workplace Accident Leave: A Stochastic Frontier Analysis." in *Disadvantaged Workers: Empirical Evidence and Labour Policies*, edited by M. Á. Malo and D. Sciulli. Basel: AIEL Series in labor economics. Springer.

Martín-Román, Ángel L., Alfonso Moral, and Sara Pinillos-Franco. 2024. "Are Women Breaking the Glass Ceiling? A Gendered Analysis of the Duration of Sick Leave in Spain." *International Journal of Health Economics and Management* 24(1):107–34.

Martin-Roman, Angel Luis and Alfonso Moral. 2016. "Moral Hazard in Monday Claim Filing: Evidence from Spanish Sick Leave Insurance." *B.E. Journal of Economic Analysis and Policy* 16(1):437–56.

Martín-Román, Ángel and Alfonso Moral. 2017. "A Methodological Proposal to Evaluate the Cost of Duration Moral Hazard in Workplace Accident Insurance." *European Journal of Health Economics* 18(9):1181–98.

Mastekaasa, A. 2014. "The Gender Gap in Sickness Absence: Long-Term Trends in Eight European Countries." *The European Journal of Public Health* 24(4).

Mastekaasa, Arne and Anne May Melsom. 2014. "Occupational Segregation and Gender Differences in Sickness Absence: Evidence from 17 European Countries." *European Sociological Review* 30(5).

Meeusen, Wim and Julien van Den Broeck. 1977. "Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error." *International Economic Review* 18(2):444.

Modica, Maddalena, Maurizio Ferratini, Rosa Spezzaferri, Renata De Maria, Emanuele Previtali, and Paolo Castiglioni. 2014. "Gender Differences in Illness Behavior After Cardiac Surgery." *Journal of Cardiopulmonary Rehabilitation and Prevention* 34(2):123–29.

Mukuria, Clara, Donna Rowen, Mónica Hernández-Alava, Simon Dixon, and Roberta Ara. 2017. "Predicting Productivity Losses from Health-Related Quality of Life Using Patient Data." *Applied Health Economics and Health Policy* 15(5):597–614.

Nelson, Jon P. 2014. "Gender Differences in Alcohol Demand: A Systematic Review of the Role of Prices and Taxes." *Health Economics* 23(10):1260–80.

Nielsen, Helena Skyt. 1998. "Discrimination and Detailed Decomposition in a Logit Model." *Economics Letters* 61(1):115–20.

Oaxaca, Ronald. 1973. "Male-Female Wage Differentials in Urban Labor Markets."

International Economic Review 14(3):693.

Oksuzyan, Anna, Jordi Gumà, and Gabriele Doblhammer. 2018. "Sex Differences in Health and Survival." Pp. 65–100 in *A Demographic Perspective on Gender, Family and Health in Europe*. Cham: Springer International Publishing.

Østby, K. A., A. Mykletun, and W. Nilsen. 2018. "Explaining the Gender Gap in Sickness Absence." *Occupational Medicine* 68(5):320–26.

Schünemann, Johannes, Holger Strulik, and Timo Trimborn. 2017. "The Gender Gap in Mortality: How Much Is Explained by Behavior?" *Journal of Health Economics* 54:79–90.

Spierdijk, Laura, Gijsbert van Lomwel, and Wilko Peppelman. 2009. "The Determinants of Sick Leave Durations of Dutch Self-Employed." *Journal of Health Economics* 28(6):1185–96.

Stevenson, Rodney E. 1980. "Likelihood Functions for Generalized Stochastic Frontier Estimation." *Journal of Econometrics* 13(1):57–66.

Weisberg, Yanna J., Colin G. De Young, and Jacob B. Hirsh. 2011. "Gender Differences in Personality across the Ten Aspects of the Big Five." *Frontiers in Psychology* 2(AUG).

Wolf, Jennifer Moriatis, Lisa Cannada, Ann E. Van Heest, Mary I. O'Connor, and Amy L. Ladd. 2015. "Male and Female Differences in Musculoskeletal Disease." *Journal of the American Academy of Orthopaedic Surgeons* 23(6):339–47.

Yun, Myeong Su. 2004. "Decomposing Differences in the First Moment." *Economics Letters* 82(2):275–80.

Yun, Myeong Su. 2005. "A Simple Solution to the Identification Problem in Detailed Wage Decompositions." *Economic Inquiry* 43(4):766–72.

Figure 1. Standard Oaxaca-Blinder decomposition with all the components operating in the same direction.

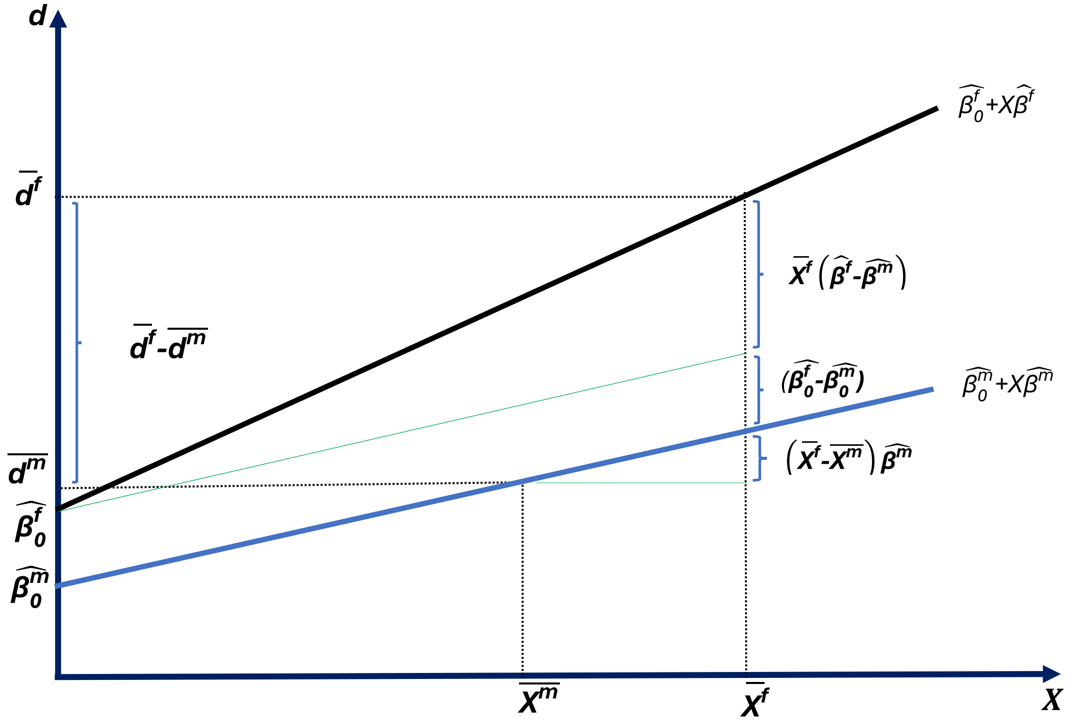


Figure 2. Standard Oaxaca-Blinder decomposition with the intercept differences operating in opposite direction.

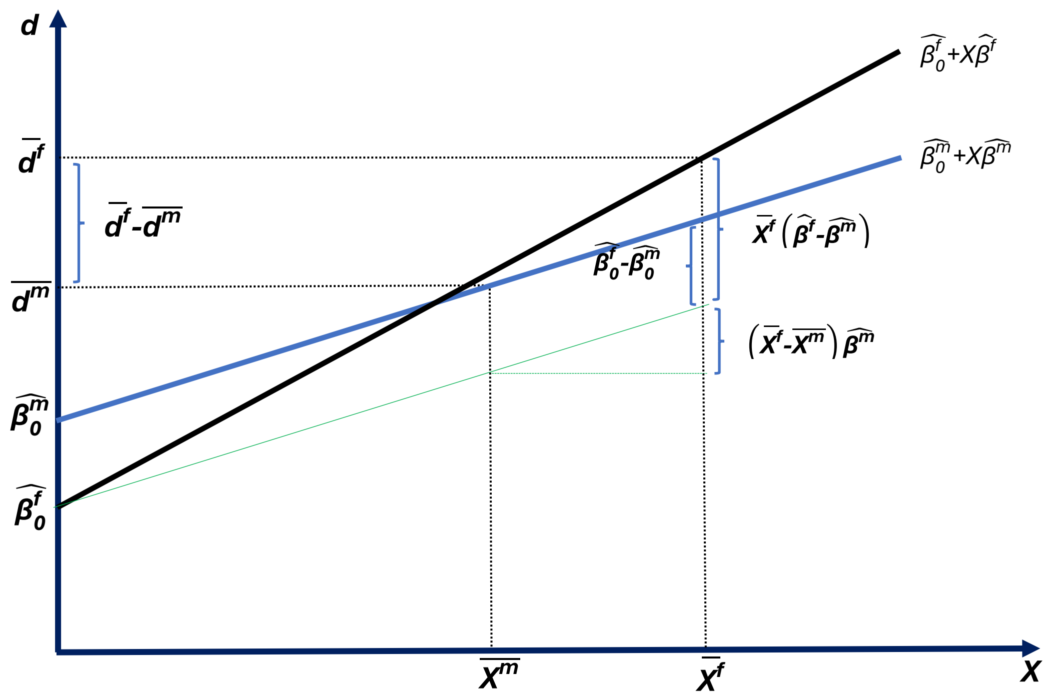


Figure 3. Nonlinear Oaxaca-Blinder decomposition with SFA estimation and the intercept differences acting in the same direction.

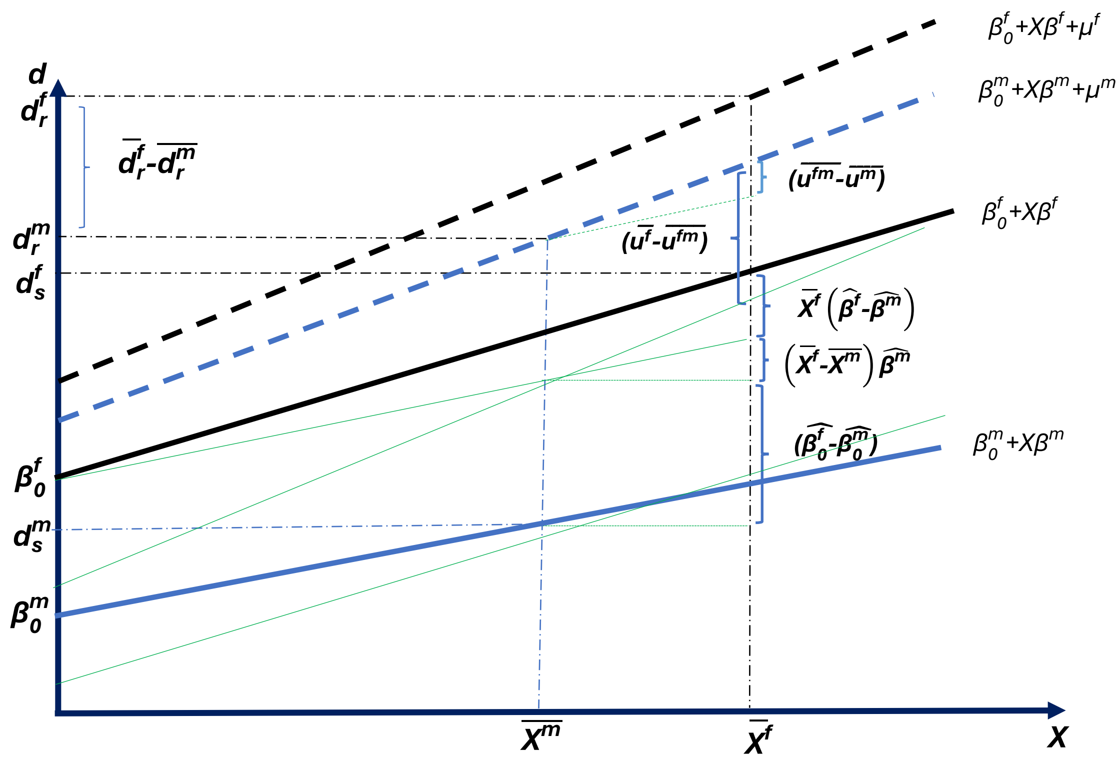


Table 1. Frontier estimations of the logarithm of the sick leave duration by gender (normalized).

	Female			Male		
	Coeff.	*P>z*	*Coeff. Norm*	*Coeff.*	*P>z*	*Coef . Norm*
			Type of injury			
Not specified			-0.096			-0.133
Superficial Injuries	-0.143	0.000	-0.239	-0.120	0.000	-0.253
Other injuries	-0.141	0.000	-0.237	-0.070	0.000	-0.203
Fractures	1.126	0.000	1.030	1.106	0.000	0.973
Strains	0.019	0.007	-0.077	0.047	0.000	-0.086
Dislocations	0.058	0.000	-0.038	0.118	0.000	-0.016
Sprain	0.033	0.000	-0.063	0.056	0.000	-0.077
Traumatic amputation	0.971	0.000	0.875	1.093	0.000	0.960
Concussion	0.015	0.048	-0.081	0.092	0.000	-0.041
Burns	-0.419	0.000	-0.515	-0.132	0.000	-0.266
Poisoning	-0.388	0.000	-0.483	-0.420	0.000	-0.553
Choking	-0.567	0.000	-0.663	-0.714	0.000	-0.847
Noise, heat	-0.184	0.000	-0.280	-0.155	0.000	-0.288
Psychological trauma	0.084	0.000	-0.012	-0.010	0.366	-0.144
Multiple injuries	0.153	0.000	0.057	0.248	0.000	0.115
Heart attack	0.917	0.000	0.821	0.992	0.000	0.859
			Part of the body			
Not specified			0.142			0.090
Head	-0.348	0.000	-0.207	-0.418	0.000	-0.328
Face	-0.567	0.000	-0.425	-0.466	0.000	-0.375
Eyes	-0.966	0.000	-0.824	-0.961	0.000	-0.871
Neck (spine)	0.129	0.000	0.271	0.070	0.000	0.160
Neck (rest)	0.048	0.002	0.189	-0.052	0.000	0.038
Back (spine)	-0.171	0.000	-0.029	-0.260	0.000	-0.170
Back (rest)	-0.184	0.000	-0.042	-0.278	0.000	-0.188
Trunk	-0.181	0.000	-0.039	-0.100	0.000	-0.010
Shoulder	0.235	0.000	0.376	0.270	0.000	0.361
Arm	0.101	0.000	0.243	0.128	0.000	0.219
Hand	-0.183	0.000	-0.041	-0.062	0.000	0.029
Finger (hand)	-0.279	0.000	-0.137	-0.068	0.000	0.023
Wrist	0.041	0.004	0.183	0.043	0.000	0.133
Upper limbs (not esp.)	0.073	0.000	0.215	0.120	0.000	0.210
Leg	0.072	0.000	0.214	0.238	0.000	0.329
Ankle	-0.099	0.000	0.043	0.006	0.577	0.096
Foot	-0.185	0.000	-0.044	-0.059	0.000	0.032
Finger (foot)	-0.550	0.000	-0.409	-0.310	0.000	-0.220
Lower limbs (not esp)	-0.046	0.002	0.096	0.069	0.000	0.159
Multiple parts	0.084	0.000	0.226	0.193	0.000	0.283
Ambulatory			-0.080			-0.098
Hospital care	0.160	0.000	0.080	0.196	0.000	0.098
No hospitalization			-0.245			-0.318
Hospitalization	0.490	0.000	0.245	0.636	0.000	0.318

Continue

Continuation

Minor			-0.455			-0.525
Serious	0.910	0.000	0.455	1.050	0.000	0.525
Accident			-0.195			-0.201
Relapse	0.390	0.000	0.195	0.402	0.000	0.201
			Age			
Less than 20			-0.254			-0.212
From 20 to 30	0.095	0.000	-0.159	0.053	0.000	-0.159
From 30 to 40	0.217	0.000	-0.037	0.151	0.000	-0.061
From 40 to 50	0.301	0.000	0.048	0.248	0.000	0.037
From 50 to 60	0.401	0.006	0.147	0.364	0.000	0.152
More than 60	0.508	0.000	0.255	0.454	0.000	0.242
			Occupation			
Company management			-0.057			0.008
Technical staff and scientists	0.050	0.001	-0.007	-0.034	0.003	-0.026
Professional support	0.026	0.077	-0.030	0.013	0.218	0.021
Administration employees	0.013	0.358	-0.043	-0.035	0.002	-0.027
Service workers	0.074	0.000	0.018	-0.021	0.045	-0.013
Skilled agriculture and fishing	0.098	0.000	0.042	0.053	0.000	0.061
Crafts and dealers	0.107	0.000	0.050	-0.029	0.005	-0.021
Machine operators	0.079	0.000	0.022	0.012	0.272	0.019
Unskilled	0.061	0.000	0.005	-0.028	0.008	-0.020
Constant	2.088	0.000	3.326	1.919	0.000	3.308
Observations		1,101,551			2,814,698	
/lnsig2v	-0.193	0.000		-0.373	0.000	
/lnsig2u	-1.774	0.000		-1.169	0.000	
sigma_v		0.908			0.830	
sigma_u		0.412			0.557	
sigma2		0.994			0.999	
Lambda		0.454			0.672	
LR test of sigma_u=0		chibar2(01) = 3.2e+03			chibar2(01) = 3.5e+04	

Source: Author's own based on SAW data

Table 2. Decomposition of the sick leave duration between female and male.

	Total difference	***Differences in standard duration***			***Differences in efficiency***	
	$\overline{d^{ff}}-\overline{d^{mm}}$	$\left(\widehat{\beta^f_o}-\widehat{\beta^m_o}\right)$	$\left(\bar{X}^f_k-\bar{X}^m_k\right)\beta^m_k$	$\bar{X}^f_k\left(\widehat{\beta^f_k}-\widehat{\beta^m_k}\right)$	$\left(\overline{u^f}-\overline{u^{fm}}\right)$	$\left(\overline{u^{fm}}-\overline{u^m}\right)$
Percentage	100%	20%	32%	208%	-179%	18%
Days	0.778	0.155	0.251	1.622	-1.390	0.140
			2.028		-1.250	

Source: Author's own based on SAW data

Table 3. Frontier estimations of the logarithm of the sick leave duration by specification.

	Female		Male	
Duration	***Coefficient***	***P>z***	***Coefficient***	***P>z***
		Ref.: Sprain		
Not specified	-0.033	0.000	-0.056	0.000
Superficial Injuries	-0.176	0.000	-0.176	0.000
Other injuries	-0.174	0.000	-0.126	0.000
Fractures	1.093	0.000	1.050	0.000
Strains	-0.014	0.000	-0.009	0.000
Dislocations	0.025	0.000	0.062	0.000
Traumatic amputation	0.938	0.000	1.037	0.000
Concussion	-0.018	0.000	0.036	0.000
Burns	-0.452	0.000	-0.188	0.000
Poisoning	-0.421	0.000	-0.476	0.000
Choking	-0.600	0.000	-0.770	0.000
Noise, heat	-0.217	0.000	-0.211	0.000
Psychological trauma	0.051	0.001	-0.066	0.000
Multiple injuries	0.120	0.000	0.192	0.000
Heart attack	0.884	0.000	0.937	0.000
		Ref.: Ankle		
Not specified	0.099	0.000	-0.006	0.577
Head	-0.249	0.000	-0.424	0.000
Face	-0.468	0.000	-0.471	0.000
Eyes	-0.867	0.000	-0.967	0.000
Neck (spine)	0.228	0.000	0.064	0.000
Neck (rest)	0.147	0.000	-0.058	0.000
Back (spine)	-0.072	0.000	-0.266	0.000
Back (rest)	-0.084	0.000	-0.284	0.000
Trunk	-0.082	0.000	-0.106	0.000
Shoulder	0.334	0.000	0.265	0.000
Arm	0.200	0.000	0.123	0.000
Hand	-0.084	0.000	-0.067	0.000
Finger (hand)	-0.179	0.000	-0.073	0.000
Wrist	0.140	0.000	0.037	0.000
Upper limbs (not esp.)	0.172	0.000	0.114	0.000
Leg	0.171	0.000	0.233	0.000
Foot	-0.086	0.000	-0.064	0.000
Finger (foot)	-0.451	0.000	-0.316	0.000
Lower limbs (not esp.)	0.053	0.000	0.063	0.000
Multiple parts	0.183	0.000	0.187	0.000

Continuation

Ref.: Unskilled				
Company management	-0.061	0.000	0.028	0.008
Technical staff and scientists	-0.011	0.002	-0.006	0.194
Professional support	-0.035	0.000	0.041	0.000
Administration employees	-0.048	0.000	-0.007	0.066
Service workers	0.013	0.000	0.007	0.001
Skilled agriculture and fishing	0.037	0.000	0.081	0.000
Crafts and dealers	0.045	0.000	-0.001	0.370
Machine operators	0.018	0.001	0.039	0.000
Hospital care	0.160	0.000	0.196	0.000
Hospitalization	0.490	0.000	0.636	0.000
Serious	0.910	0.000	1.050	0.000
Relapse	0.390	0.000	0.402	0.000
Ref.: From 30 to 40				
Less than 20	-0.217	0.000	-0.151	0.000
From 20 to 30	-0.122	0.000	-0.098	0.000
From 40 to 50	0.085	0.000	0.097	0.000
From 50 to 60	0.184	0.000	0.213	0.000
More than 60	0.292	0.000	0.303	0.000
Constant	2.299	0.000	2.104	0.000
Observations	1,101,551		2,814,698	
/lnsig2v	-0.193	0.000	-0.374	0.000
/lnsig2u	-1.766	0.000	-1.165	0.000
sigma_v	0.908		0.830	
sigma_u	0.414		0.559	
sigma2	0.995		1.000	
lambda	0.456		0.673	
LR test of sigma_u=0	chibar2(01) = 3.2e+03		chibar2(01) = 3.5e+04	

Source: Author's own based on SAW data

Table 4. Decomposition of the sick leave duration between female and male.

	Total difference	***Differences in standard duration***			***Differences in efficiency***	
	$\overline{d^{ff}}-\overline{d^{mm}}$	$\left(\widehat{\beta_o^f}-\widehat{\beta_o^m}\right)$	$\left(\overline{X_k^f}-\overline{X_k^m}\right)\beta_k^m$	$\overline{X_k^f}\left(\widehat{\beta_k^f}-\widehat{\beta_k^m}\right)$	$\left(\overline{u^f}-\overline{u^{fm}}\right)$	$\left(\overline{u^{fm}}-\overline{u^m}\right)$
Percentage	100%	216%	32%	12%	-179%	18%
Days	0.778	1.682	0.251	0.095	-1.390	0.140
		2.028			-1.250	

Source: Author's own based on SAW data

Appendix 1: Normalized regression to avoid identification problem.

The identification problem arises because when estimating groups of dummy variables, it is necessary to leave out one from the model. In this situation, the independent term not only changes based on the removed variable but also part of the decomposition related to that component. We thus calculated a normalized regression following Yun (2005) to solve this problem. According to Yun, if we start from an estimate of the duration expressed as:

$$d=b_0+\left(\sum_{i=2}^{I} s_i S_i \sum_{j=2}^{J} t_j T_j\right)+\sum_{k=1}^{K} b_k X_k +u$$

Where *S* and *T* are groups of *I* and *J* dummy variables, *X* includes *K* continuous variables, and *u* is the inefficiency term. From equation *(5),* we may obtain a normalized regression that does not omit reference groups, and we can calculate it as follows:

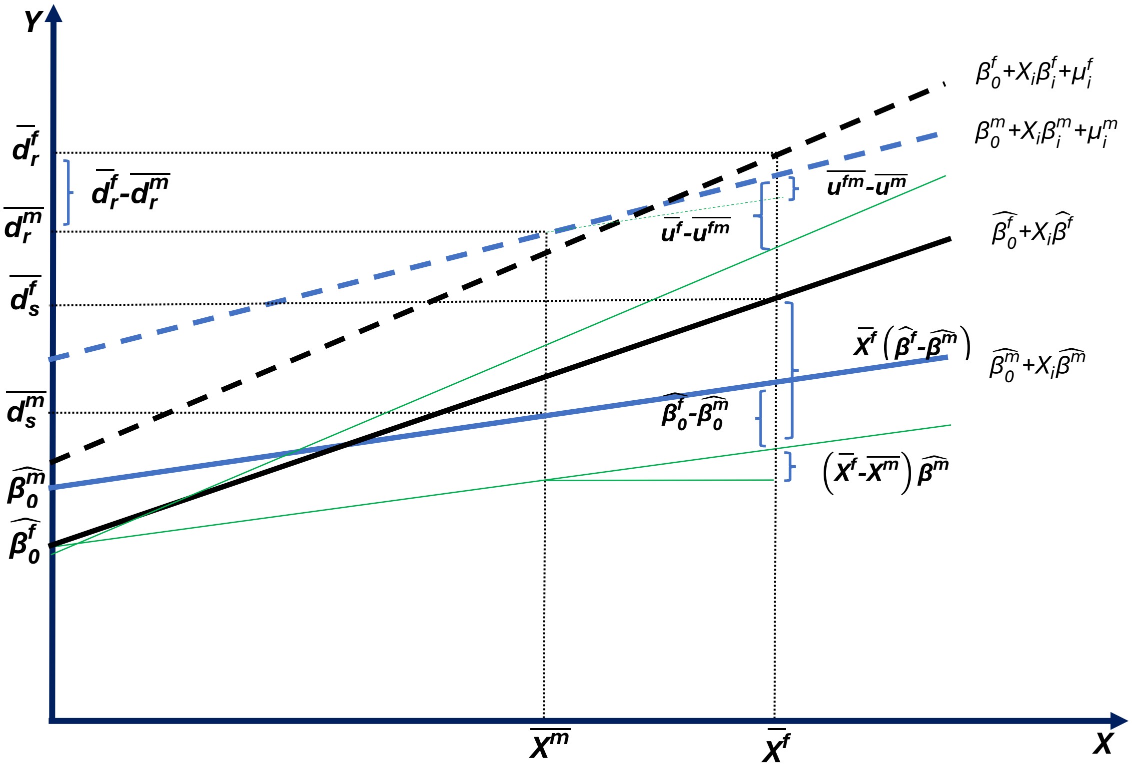
$$d=b_o^*+\left(\sum_{i=1}^{I} s_i^* S_i + \sum_{j=1}^{J} t_j^* T_j\right)+\sum_{k=1}^{K} b_k X_k +u$$

Where: $b_0^*=b_0+\bar{s}+\bar{t}$, $s_i^*=s_i-\bar{s}$ and $t_i^*=t_i-\bar{t}$

Being: $\bar{s}=\frac{\sum_{i=1}^{I} s_i}{I}$, $\bar{t}=\frac{\sum_{i=1}^{J} t_i}{J}$ and $s_1=t_1=0$

Appendix 2: Decomposition of the SFA estimation with intercept differences acting in the opposite direction to the rest of the components.

Figure A1. Nonlinear Oaxaca-Blinder decomposition with SFA estimation and the intercept differences acting in opposite direction.



Appendix 3: Nonlinear decomposition with age and age squared as continuous variables.

Table A1. Frontier estimations of the logarithm of the sick leave duration by gender (normalized).

	Female			*Male*		
	Coeff.	*P>z*	*Coeff. Norm*	*Coeff.*	*P>z*	*Coef . Norm*
			Type of injury			
Not specified			-0.096			-0.133
Superficial Injuries	-0.142	0.000	-0.239	-0.119	0.000	-0.252
Other injuries	-0.140	0.000	-0.237	-0.069	0.000	-0.202
Fractures	1.124	0.000	1.027	1.105	0.000	0.972
Strains	0.019	0.007	-0.077	0.048	0.000	-0.086
Dislocations	0.058	0.000	-0.038	0.118	0.000	-0.015
Sprain	0.033	0.000	-0.063	0.056	0.000	-0.077
Traumatic amputation	0.973	0.000	0.876	1.092	0.000	0.959
Concussion	0.016	0.048	-0.081	0.092	0.000	-0.041
Burns	-0.419	0.000	-0.515	-0.131	0.000	-0.265
Poisoning	-0.387	0.000	-0.483	-0.418	0.000	-0.552
Choking	-0.566	0.000	-0.662	-0.713	0.000	-0.846
Noise, heat	-0.181	0.000	-0.277	-0.155	0.000	-0.288
Psychological trauma	0.086	0.000	-0.010	-0.010	0.386	-0.143
Multiple injuries	0.154	0.000	0.058	0.249	0.000	0.116
Heart attack	0.914	0.000	0.818	0.988	0.000	0.854
			Part of the body			
Not specified			0.142			0.090
quHead	-0.348	0.000	-0.207	-0.418	0.000	-0.328
Face	-0.567	0.000	-0.425	-0.466	0.000	-0.375
Eyes	-0.966	0.000	-0.824	-0.961	0.000	-0.871
Neck (spine)	0.129	0.000	0.271	0.070	0.000	0.160
Neck (rest)	0.048	0.002	0.189	-0.052	0.000	0.038
Back (spine)	-0.171	0.000	-0.029	-0.260	0.000	-0.170
Back (rest)	-0.184	0.000	-0.042	-0.278	0.000	-0.188
Trunk	-0.181	0.000	-0.039	-0.100	0.000	-0.010
Shoulder	0.235	0.000	0.376	0.270	0.000	0.361
Arm	0.101	0.000	0.243	0.128	0.000	0.219
Hand	-0.183	0.000	-0.041	-0.062	0.000	0.029
Finger (hand)	-0.279	0.000	-0.137	-0.068	0.000	0.023
Wrist	0.041	0.004	0.183	0.043	0.000	0.133
Upper limbs (not esp.)	0.073	0.000	0.215	0.120	0.000	0.210
Leg	0.072	0.000	0.214	0.238	0.000	0.329
Ankle	-0.099	0.000	0.043	0.006	0.577	0.096
Foot	-0.185	0.000	-0.044	-0.059	0.000	0.032
Finger (foot)	-0.550	0.000	-0.409	-0.310	0.000	-0.220
Lower limbs (not esp)	-0.046	0.002	0.096	0.069	0.000	0.159
Multiple parts	0.084	0.000	0.226	0.193	0.000	0.283
Age	0.012	0.000	0.012	0.007	0.000	0.007
Age squared	0.000	0.006	0.000	0.000	0.000	0.000

Continue

Continuation

Ambulatory			-0.080			-0.098
Hospital care	0.160	0.000	0.080	0.196	0.000	0.098
No hospitalization			-0.245			-0.318
Hospitalization	0.490	0.000	0.245	0.635	0.000	0.318
Minor			-0.454			-0.525
Serious	0.909	0.000	0.454	1.049	0.000	0.525
Accident			-0.195			-0.201
Relapse	0.390	0.000	0.195	0.402	0.000	0.201
		Occupation				
Company management			-0.057			0.005
Technical staff and scientists	0.049	0.001	-0.008	-0.031	0.006	-0.027
Professional support	0.028	0.061	-0.029	0.018	0.105	0.022
Administration employees	0.013	0.384	-0.044	-0.032	0.003	-0.028
Service workers	0.075	0.000	0.018	-0.017	0.110	-0.012
Skilled agriculture and fishing	0.100	0.000	0.043	0.056	0.000	0.061
Crafts and dealers	0.107	0.000	0.050	-0.026	0.013	-0.021
Machine operators	0.080	0.000	0.023	0.014	0.175	0.019
Unskilled	0.060	0.000	0.003	-0.023	0.031	-0.018
Constant	1.878	0.000	2.864	1.746	0.000	2.926
Observations		1,101,551			2,814,698	
/lnsig2v	-0.193	0.000		-0.373	0.000	
/lnsig2u	-1.774	0.000		-1.169	0.000	
sigma_v		0.908			0.830	
sigma_u		0.412			0.557	
sigma2		0.994			0.999	
lambda		0.454			0.672	
LR test of sigma_u=0		chibar2(01) = 3.2e+03			chibar2(01) = 3.5e+04	

Source: Author's own based on SAW data

Table A2. Decomposition of the sick leave duration between females and males.

	Total difference	***Differences in standard duration***			***Differences in efficiency***	
	$\overline{d^{ff}}-\overline{d^{mm}}$	$\left(\widehat{\beta}_o^f-\widehat{\beta}_o^m\right)$	$\left(\overline{X}_k^f-\overline{X}_k^m\right)\beta_k^m$	$\overline{X}_k^f\left(\widehat{\beta}_k^f-\widehat{\beta}_k^m\right)$	$\left(\overline{u^f}-\overline{u^{fm}}\right)$	$\left(\overline{u^{fm}}-\overline{u^m}\right)$
Percentage	100%	-69%	34%	296%	-178%	17%
Days	0.778	-0.535	0.265	2.301	-1.389	0.135
			2.032		-1.253	

Source: Author's own based on SAW data