

Sobre a discriminação sexual na formação de salários

Ana Rute Cardoso

Institute for Economic Analysis,
Barcelona GSE

Pedro Portugal

Banco de Portugal, Universidade
NOVA de Lisboa and IZA Bonn

Paulo Guimarães

Banco de Portugal

Pedro S. Raposo

UCP - Católica Lisbon School of
Business and Economics

Abril 2016

Resumo

Em Portugal, ao longo das últimas duas décadas, a proporção de mulheres entre os trabalhadores empregados aumentou de 35 para 45%. Esta evolução foi acompanhada por uma redução acentuada do hiato salarial de 32 para 20% resultante de uma melhoria dos salários das mulheres. Este progresso fica a dever-se quase inteiramente à melhoria das suas qualificações após duas décadas de investimentos em capital humano. Em 2013, as mulheres tinham características observáveis idênticas aos homens. Contudo, a discriminação de género permaneceu mais ou menos constante ao longo do período 1991-2013. Neste estudo, investigamos as origens da diferença salarial por género e concluímos que a distribuição de trabalhadores pelas empresas e categorias profissionais explicam cerca de dois quintos da diferença salarial por género. (JEL: J16, J24, J31, J71)

“Um dos aspectos da desigualdade é a singularidade - isto é, não o ser este homem mais, neste ou naquele característico, que outros homens, mas o ser tão-somente diferente dele.”

“Os espíritos altamente analíticos veem quase só defeitos: quanto mais forte a lente mais imperfeita se mostra a coisa observada.”

Fernando Pessoa

Introdução

Em 1991 os salários das mulheres portuguesas eram inferiores aos dos homens em cerca de um terço. Desde essa data um número crescente de mulheres, cada vez mais qualificadas, integrou o mercado de trabalho português. Em resultado desse fluxo, em 2013, a proporção de mulheres no “stock” de empregados em 2013 tinha aumentado de 35 para 45 por cento (Gráfico 1).

Agradecimentos: Os autores estão imensamente gratos pela ajuda providenciada por António Antunes, Hugo Reis, Lucena Vieira e Hugo Vilares e beneficiaram de discussões com os alunos de Economia de Trabalho da NOVA SBE.

E-mail: anarute.cardoso@iae.csic.es; pfguimaraes@bportugal.pt; pportugal@bportugal.pt; pedro.raposo@ucp.pt

Em simultâneo, de 1991 para 2013 o hiato salarial entre homens e mulheres reduziu-se em 12 pontos percentuais (Gráfico 2).

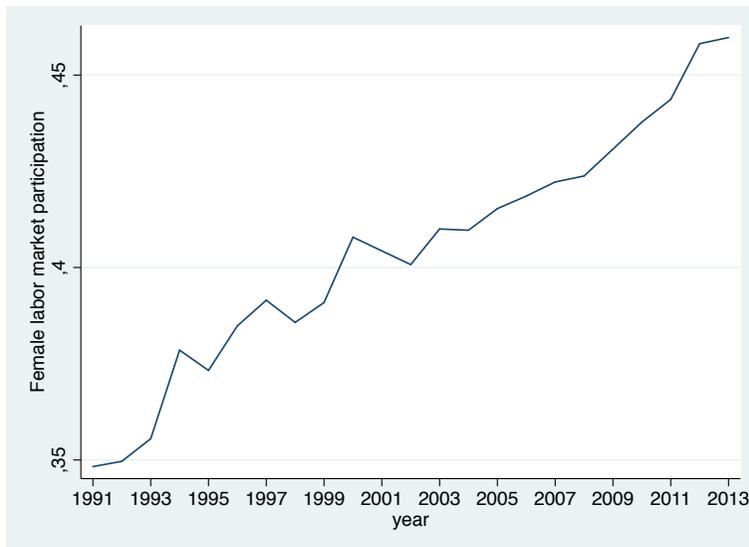


GRÁFICO 1: Taxa de participação feminina no mercado de trabalho

Esta evolução do hiato salarial reflete, essencialmente, uma evolução favorável das características das mulheres que favorecem o aumento de produtividade e, portanto, dos salários. De facto, quando o cálculo do hiato salarial é ajustado, através de uma regressão convencional, para as características observadas dos homens e das mulheres, a indicação de aproximação dos salários deixa de se verificar (Gráfico 2). Dito de outra forma, os progressos salariais do contingente feminino observados ao longo dos 22 anos são devidos quase exclusivamente à melhoria das suas qualificações (experiência profissional, antiguidade, etc.) e não à redução da componente não explicada do diferencial salarial, que é a componente convencionalmente associada à noção de discriminação sexual. Neste sentido, não há qualquer indicação de que a discriminação sexual na formação dos salários se tenha atenuado, pelo contrário, agravou-se ligeiramente.

Neste ensaio pretende-se fazer emergir os mecanismos responsáveis pela dimensão do hiato salarial através da execução de um conjunto de exercícios de decomposição da distribuição de salários. Em primeiro lugar, será explorada a metodologia de decomposição dos quantis de Machado e Mata (2005) com o objetivo de distinguir as alterações estruturais das composicionais na comparação das distribuições de salários por género, para 1991 e 2013. Em segundo lugar, a estimação de modelos de regressão com efeitos fixos de elevada dimensionalidade será combinada com a decomposição de Gelbach no sentido de identificar os desequilíbrios na afetação dos trabalhadores a empresas e categorias profissionais com regimes

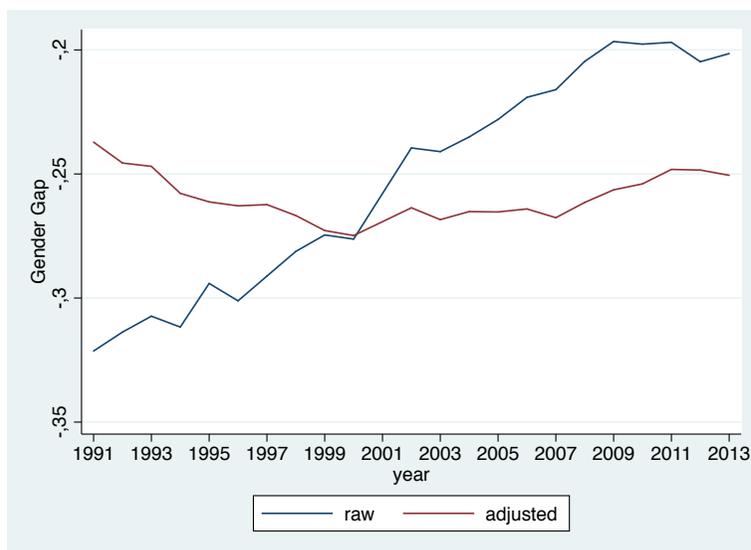


GRÁFICO 2: Discriminação salarial por género

de remuneração heterogêneos. Deste modo, será apresentada uma versão atualizada e aprofundada dos resultados do estudo de Cardoso, Guimarães e Portugal (2016). Por fim, a estimação de modelos de regressão de efeitos fixos de elevada dimensionalidade de Portugal e Guimarães (2010) será generalizada à estimação de coeficientes de regressão de elevada dimensionalidade com o propósito de medir hiatos salariais ao nível da empresa e da categoria profissional.

Os dados utilizados neste estudo correspondem aos registos individuais dos Quadros de Pessoal (1986-2009) e do Relatório Único (2010-2013) que, conjuntamente, são uma base de dados longitudinal em que a informação sobre empregadores, os trabalhadores, os acordos salariais e as categorias profissionais estão devidamente emparelhadas. A informação sobre salários é precisa e exaustiva e cobre os estabelecimentos do sector empresarial com pelo menos um trabalhador assalariado. A fiabilidade desta informação é potenciada pelo facto dos Quadros de Pessoal, até 2009, e do Relatório Único, após 2009, constituírem elementos essenciais para verificar se os empregadores obedecem aos acordos salariais estabelecidos através da negociação coletiva.

A próxima secção faz uma breve revisão da literatura. A secção 3 descreve os dados, enquanto os métodos são discutidos na secção 4. A secção 5 providencia os resultados principais dos determinantes do hiato salarial por género. A secção 6 conclui.

Revisão da Literatura

Estamos a assistir a um ressurgimento do interesse pelos determinantes do hiato salarial entre géneros, sob novas abordagens empíricas, dados mais ricos e renovadas perspectivas teóricas. A análise económica tradicionalmente tinha-se focado sobretudo na importância da participação feminina no mercado de trabalho e nas diferenças de atributos observáveis entre homens e mulheres. Qualquer um desses dois mecanismos pode ser entendido intuitivamente. Se a taxa de participação feminina é baixa, então há margem para que os atributos das trabalhadoras não sejam representativos das características da população feminina em geral. Esta selecção pode funcionar no sentido de elevar ou baixar os salários das mulheres, dependendo se as normas sociais, preferências, as condições económicas e políticas públicas atraem desproporcionalmente para o mercado de trabalho mais ou menos mulheres qualificadas (ao longo das dimensões que podem ser observáveis ou não observáveis). Em qualquer caso, com o aumento da taxa de participação das mulheres é esperado que se assista a uma diminuição da importância da selecção nos hiatos salariais (ver evidência entre países em Olivetti e Petrongolo (2008) ou a evidência ao longo do tempo para os EUA em Stanley e Jarrell (1998) e Jarrell e Stanley (2004)). Concomitantemente, as qualificações das mulheres e dos homens no mercado de trabalho irá influenciar o seu salário relativo (ver a ampla evidência de que a educação e a experiência contribuem para moldar as disparidades salariais, na revisão de Altonji e Blank (1999)). Sob esta vertente da literatura, a convergência no aproveitamento escolar entre homens e mulheres (se não a reversão da diferença, em favor das mulheres) e o aumento de forma persistente das mulheres na força de trabalho deveriam conduzir ao encerramento do hiato salarial. Surpreendentemente, uma pergunta perdura: Porque é que o hiato salarial é tão persistente, apesar da convergência acentuada das taxas de participação e das características observáveis de homens e mulheres, em especial nas economias desenvolvidas?

A literatura mais recente aborda esta questão. Blau e Kahn (2016) identificam uma redução do hiato salarial entre géneros nos EUA nas últimas décadas. Uma revisão da literatura recente para outros países aponta para um conjunto de factos estilizados e desafios a ultrapassar. Em primeiro lugar, a convergência das qualificações e da experiência profissional desempenhou um papel fundamental para reduzir o hiato salarial. Esses fatores têm actualmente um impacto silencioso sobre as diferenças salariais entre homens e mulheres. Pelo contrário, a indústria e a ocupação surgem como fatores geradores de diferenças salariais por género. Deste modo, é necessário compreender plenamente a alocação das pessoas de diferentes géneros pelos diferentes sectores de actividade e profissões (e a remuneração associada). Neste contexto, é imperativo melhorar o conhecimento das políticas de recrutamento e salariais das empresas. Um terceiro elemento digno de nota é que a disparidade salarial é persistentemente maior no

topo da distribuição das qualificações e dos salários. As fontes deste “efeito de tecto de vidro” também não são ainda totalmente conhecidas. Algumas das explicações plausíveis destacadas por Blau e Khan incluem diferenças de atributos psicológicos (por exemplo, o poder de negociação), que penalizam as mulheres no acesso ao topo da carreira profissional, diferenciais compensatórios para as características dos lugares de topo (por exemplo, jornadas de trabalho mais longas e com horários mais exigentes) e discriminação pura.

A recente facilidade de acesso a bases de dados que integram informação sobre trabalhadores e empregadores tem permitido aprofundar o estudo sobre algumas destas questões. Cardoso *et al.* (2016) (CGP) quantificam o impacto da distribuição dos trabalhadores pelas empresas e pelas categorias profissionais no hiato salarial entre homens e mulheres. No essencial, os autores consideram que a generosidade da política de remuneração das empresas (e das categorias profissionais) pode ser capturada pela presença de um efeito fixo da empresa (e da categoria profissional) numa regressão de salários. Ao estabelecer a comparação entre esses efeitos fixos entre homens e mulheres, CGP concluem que a afetação dos trabalhadores às empresas faz com que as mulheres estejam sub-representadas em empresas (e em categorias profissionais) que oferecem salários mais elevados. A desigualdade de género na afetação dos trabalhadores às empresas e categorias profissionais explica cerca de 40 por cento do hiato salarial. Esta quantificação, que generaliza a metodologia de Gelbach (2016), acomoda a presença de heterogeneidade das qualificações dos trabalhadores, sejam ou não observadas.

O estudo de Card *et al.* (2016) (CCK) explora uma outra dimensão do problema. Os autores propõem-se testar uma hipótese que tem sido objeto de intensas discussões noutras áreas do conhecimento: a hipótese de que as atitudes das mulheres relativamente à competição e à negociação as penalizam, em comparação com os homens. Sob esta premissa, as mulheres extrairiam menos rendas dos seus empregadores. Nesta linha de investigação, CCK consideraram a possibilidade de hiatos salariais ao nível da empresa e fizeram corresponder esses diferenciais a medidas de desempenho das empresas. Os autores quantificam a influência de dois canais de penalização salarial: a afetação dos trabalhadores às empresas e a negociação de salários. O exercício de decomposição do hiato salarial é, então, conduzido de forma a, alternativamente, fechar cada um dos canais. Atribuindo, artificialmente, o prémio salarial dos homens às mulheres é fechado o canal da negociação. Impondo uma distribuição igualitária dos homens e das mulheres pelas empresas, é fechado o canal da afetação. CCK concluem que o canal da negociação será responsável por 5 por cento do hiato salarial em Portugal. Por outro lado, o estudo confirma a importância da triagem dos trabalhadores pelas empresas, indicando que será responsável por 15 por cento do hiato salarial.

Outra vertente recente da literatura explora o papel dos diferenciais compensatórios para os lugares de topo, em particular jornadas de trabalho mais longas e com horários mais exigentes. Goldin (2014) e Bertrand e Katz (2010) estão entre os estudos que apresentam evidências convincentes sobre a importância deste canal.

O objetivo do presente trabalho é progredir ao longo da nova vertente da literatura que se baseia em dados de grande dimensão empregador-empregado para avaliar o papel da empresa moldar as disparidades salariais.

Dados

A base de dados dos *Quadros de Pessoal* (QP) é, por definição, uma base de dados longitudinal empregador-trabalhador-categoria profissional. QP é um inquérito anual obrigatório compilado pelo Ministério do Trabalho, Solidariedade e Segurança Social (MTSSS), e abrange praticamente todas as empresas que empregam trabalho pago em Portugal. Dada a natureza obrigatória do inquérito, problemas normalmente associados com dados de painel, como o atrito de painel, são consideravelmente atenuados.

A base de dados inclui informações específicas da empresa – localização, indústria (códigos SIC), o estatuto legal, a natureza da propriedade por nacionalidade, emprego, vendas – e informações específicas de cada um dos seus trabalhadores – rendimentos do trabalho, qualificações do trabalhador, género, idade, antiguidade, horas de trabalho, etc. A informação sobre os rendimentos é muito detalhada, precisa e completa. A base de dados inclui o salário base – salário bruto –, os benefícios regulares e o pagamento de horas extraordinárias. A informação sobre horas de trabalho normais e extraordinárias também está disponível. Uma vez que a informação sobre os rendimentos é reportada pela empresa, está sujeita a um menor erro de medição do que se fosse fornecida directamente pelo trabalhador. A lei obriga a que as informações contidas nos QP estejam disponíveis em espaço público dentro das instalações da empresa, o que reforça ainda mais a nossa confiança na sua informação.

Uma característica notável dos QP é que recolhem informações sobre a convenção colectiva que rege a relação entre o empregador e o empregado definida na sua dimensão salarial. Além disso, dentro de cada convenção colectiva, identifica a categoria profissional que o trabalhador detém. A importância da passagem de uma ampla classificação de ocupações tradicionalmente disponíveis para uma descrição mais rica das tarefas reais executadas pelos trabalhadores tem sido destacada na literatura (ver, por exemplo Autor (2013), Goos e Manning (2007), Autor *et al.* (2006) e Dustmann *et al.* (2009) na polarização do trabalho). Esta literatura recente mostra que, além da heterogeneidade da empresa e do trabalhador, os salários são

moldados pela heterogeneidade da tarefa que, portanto, deve ser considerada explicitamente na análise Torres *et al.* (2013).

Uma série de restrições foram impostas sobre o conjunto de dados brutos. Em primeiro lugar, limitamos a nossa análise aos trabalhadores a tempo completo em Portugal continental, entre 1986 e 2013.¹ Em segundo lugar, foram excluídos os trabalhadores dos sectores da agricultura e da pesca. Em terceiro lugar, os indivíduos com menos de 18 anos e maiores de 65 anos foram também retirados. Em quarto lugar, retirámos da análise os trabalhadores cujos salários mensais foram inferiores a 80 por cento do salário mínimo obrigatório, o que corresponde ao salário mais baixo admissível para os estagiários. Em quinto lugar, excluímos observações cuja combinação entre empresa e categoria profissional incluía apenas um trabalhador. Por fim, excluímos cerca de 1 por cento do número total de observações que não permitiam estabelecer ligações directas e indirectas entre trabalhadores e empresas. A nossa amostra final incluiu 27.921.002 observações (338.580 empresas; 5.126.998 trabalhadores; 95.196 categorias profissionais).

A variável dependente utilizada nas nossas equações é uma medida de salário real por hora dada pelo rácio entre a soma dos salários de base deflacionados, benefícios regulares (incluindo diuturnidades) e pagamento de horas extra e a soma das horas normais de trabalho e horas extra.

Efeitos fixos de elevada dimensionalidade e a decomposição de Gelbach

Nesta secção seguimos de perto a abordagem empírica de Cardoso *et al.* (2016). A ideia consiste em utilizar a decomposição de Gelbach para ajudar a entender as causas subjacentes ao hiato salarial por género observado. A novidade aqui consiste na aplicação da decomposição de Gelbach a um modelo salarial de regressão linear que tem em consideração todas as principais fontes de variação dos salários, incluindo as que são devidas a componentes não observáveis e que são capturadas pela inclusão de vários efeitos fixos de elevada dimensionalidade. O nosso ponto de partida é a tradicional equação salarial Minceriana:

$$\ln w_{ifjt} = \mathbf{x}_{ifjt}\boldsymbol{\beta} + \gamma g_i + \varepsilon_{ifjt} \quad . \quad (1)$$

Na equação acima, $\ln w_{ifjt}$ designa o logaritmo natural do salário horário em termos reais. Os diversos índices associados a w servem para realçar as fontes potenciais de variação. O índice i ($i = 1, \dots, N$) refere-se ao trabalhador, f ($f = 1, \dots, F$) à empresa, enquanto j é um índice que captura as diferenças nas categorias profissionais. O índice t é para o tempo ($t = 1, \dots, T$). O vector de

1. Utilizaram-se os anos entre 1986 e 1989 com o objectivo de obter as estimativas dos três efeitos fixos de elevada dimensionalidade na equação (3) com maior precisão.

variáveis explicativas, x , contém as características observadas do trabalhador e da empresa. Entre essas variáveis encontram-se o nível de educação do trabalhador, os anos de antiguidade na empresa e a dimensão da empresa. Intencionalmente, deixámos de fora do vector x a variável g , uma variável binária para o sexo do trabalhador, que captura a disparidade na remuneração entre homens e mulheres. O foco da nossa análise é o coeficiente associado a esta variável, pois fornece-nos a estimativa convencional do desvio salarial entre homens e mulheres. Assume-se que o termo de perturbação, ε_{ifjt} , segue as hipóteses habituais. Torna-se mais conveniente expressar a equação acima na forma matricial. Procedendo deste modo obtemos

$$Y = X\beta + \gamma G + \varepsilon \quad (2)$$

onde a simbologia utilizada carece de explicação. A especificação que aqui mostramos será designada por “modelo base” pois é a regressão linear tipicamente utilizada para quantificar o desvio salarial entre homens e mulheres. Basicamente, esta regressão permite obter uma estimativa da diferença percentual entre os salários dos homens e mulheres controlando para as características observáveis dos trabalhadores, como sejam a educação e a antiguidade, e as características das empresas, como por exemplo a sua dimensão. Contudo, e no seguimento do trabalho pioneiro de Abowd *et al.* (1999), reconhecemos a necessidade de tomar em consideração todos os fatores específicos aos trabalhos e empresas que contribuem para a variação salarial. Do ponto de vista empírico isto exige a utilização de uma base de dados com informação ao nível do trabalhador e da empresa. Como mostrou Abowd *et al.* (1999), com a introdução de efeitos fixos ao nível do trabalhador e da empresa torna-se possível controlar todas as características invariantes ao longo do tempo quer estas sejam ou não observadas. Fatores tais como a capacidade não observada do trabalhador, o seu historial familiar ou a aversão ao risco são todos devidamente levados em consideração. O mesmo se aplica às características não observadas das empresas, tais como a sua estrutura organizacional, capacidade de gestão, localização, etc. A riqueza da nossa base de dados permite-nos ir mais longe. Como já foi anteriormente explicado, a existência de informação detalhada sobre as categorias profissionais permite-nos a introdução de um efeito fixo que absorve toda as características fixas, observadas e não observadas, das categorias profissionais.

A adição de efeitos fixos para a empresa e para as categorias profissionais à equação base em (2) não deveria afectar a estimativa obtida para γ a não ser que homens e mulheres se distribuam de uma forma não balanceada pelas empresas e categorias profissionais. Posto de outro modo, se a estimativa de γ se alterar quando controlamos para os efeitos fixos da empresa e das categorias profissionais então isto significa que o processo de alocação de homens e mulheres pelas empresas e categorias profissionais é um factor que contribui para a existência de um hiato salarial por género. Contudo, o modelo completo que toma em consideração todas as fontes de variação

precisa também de incluir um efeito fixo para o trabalhador. A introdução do efeito fixo para o trabalhador absorve todas as características do trabalhador que não variam ao longo do tempo, incluindo a variável binária para o sexo (G). Como veremos adiante, esta limitação não nos impede de entendermos o que acontece a γ quando no modelo completo são incluídos simultaneamente controles para as três fontes de variação (trabalhador, empresa e categoria profissional). Para tal, torna-se necessário estimar o modelo completo que inclui os três efeitos fixos. Este modelo é dado por

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{D}\theta + \mathbf{F}\varphi + \mathbf{L}\lambda + \varepsilon \quad (3)$$

onde adicionamos três efeitos fixos de elevada dimensionalidade à equação (2). Na equação acima, \mathbf{D} é uma matriz que contém as variáveis binárias relativas aos efeitos fixos do trabalhador, \mathbf{F} é uma matriz similar mas para as empresas e \mathbf{L} é uma matriz para as categorias profissionais. Como habitual, mantemos como verdadeira a opção de exogeneidade estrita do termo de perturbação.

A elevada dimensão da nossa base de dados, com cerca de 28 milhões de observações, mais de 5 milhões de trabalhadores, 400.000 empresas e cerca de 95.000 categorias profissionais, levanta alguns desafios econométricos. Uma preocupação importante tem que ver com a elevada dimensionalidade dos efeitos fixos. A estimação de uma regressão linear com três efeitos fixos de elevada dimensionalidade é um problema não trivial devido ao tamanho das matrizes envolvidas. A transformação “*within*” é capaz de absorver um dos efeitos fixos mas a elevada dimensionalidade dos outros efeitos fixos impede a utilização da fórmula convencional do estimador dos mínimos quadrados. No entanto, a estimação deste modelo é possível se utilizarmos o algoritmo de Guimarães e Portugal (2010). Este algoritmo permite a obtenção da solução exacta dos mínimos quadrados sem que seja necessária a inversão de matrizes de elevada dimensionalidade.²

Dado que pretendemos utilizar as estimativas dos efeitos fixos para análise secundária é necessário garantir que as estimativas são identificadas. Para tal restringimos a análise a um subconjunto de dados conectados. A identificação deste subconjunto é feita recorrendo a um algoritmo proposto por Weeks e Williams (1964). A aplicação deste algoritmo aos nossos dados permitiu a identificação de um subconjunto de dados conectados que representam 99% dos dados originais. Dentro deste subconjunto as estimativas dos diversos efeitos fixos diferem apenas de um fator aditivo.

A decomposição de Gelbach (2016) pode ajudar-nos a perceber o que acontece à estimativa de γ quando partimos da equação base em (2) para o modelo completo com os três efeitos fixos adicionados em (3). A abordagem

2. Neste trabalho utilizámos o comando de Stata `reghdfe` que foi escrito por Sergio Correia e que implementa uma versão melhorada do algoritmo de Guimarães e Portugal (2010).

baseia-se na conhecida fórmula para variáveis omitidas do estimador de mínimos quadrados e tem a vantagem de permitir de uma forma inequívoca separar a contribuição de cada efeito fixo para o hiato salarial por género. Para percebermos a utilização desta decomposição no presente contexto convém recordar que pelo teorema de Frisch-Waugh-Lovell (FWL) é possível obter uma estimativa de γ no modelo base pela utilização de uma abordagem em dois passos. No primeiro passo, regredimos \mathbf{Y} em \mathbf{X} e calculamos os resíduos dessa regressão. Se definirmos $\mathbf{M} \equiv [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']$ então este passo consiste no cálculo de $\mathbf{M}\mathbf{Y}$. De forma similar, podemos calcular o resíduo da regressão de \mathbf{G} em \mathbf{X} , ou seja, $\mathbf{M}\mathbf{G}$. Com este procedimento expurgámos os efeitos das variáveis \mathbf{X} de \mathbf{Y} e \mathbf{G} . Assim, se agora correremos uma regressão linear simples entre $\mathbf{M}\mathbf{Y}$ e $\mathbf{M}\mathbf{G}$ podemos estar certos pelo teorema de FWL de que obteremos a estimativa de γ do nosso modelo básico. Ou seja,

$$\hat{\gamma} = (\mathbf{G}'\mathbf{M}\mathbf{G})^{-1}\mathbf{G}'\mathbf{M}\mathbf{Y} = \mathbf{M}_{\mathbf{G}}\mathbf{Y} \quad (4)$$

onde notamos de passagem que $\mathbf{M}_{\mathbf{G}} \equiv (\mathbf{G}'\mathbf{M}\mathbf{G})^{-1}\mathbf{G}'\mathbf{M}$ e que \mathbf{M} é uma matrix idempotente. Voltemos agora a nossa atenção para a equação do modelo completo (3). A correspondente equação ajustada deste modelo pode ser apresentada da seguinte forma

$$\mathbf{Y} = \mathbf{X}\hat{\beta} + \mathbf{D}\hat{\theta} + \mathbf{F}\hat{\varphi} + \mathbf{L}\hat{\lambda} + \hat{\varepsilon} \quad (5)$$

onde substituímos os coeficientes e o termo de perturbação pelas respectivas estimativas de mínimos quadrados. Note-se que $\mathbf{D}\hat{\theta}$, $\mathbf{F}\hat{\varphi}$ e $\mathbf{L}\hat{\lambda}$ são os vetores coluna contendo as estimativas dos efeitos fixos. Para implementar a decomposição de Gelbach apenas temos de pré-multiplicar a expressão acima por $\mathbf{M}_{\mathbf{G}}$. Fazendo isto obtemos do lado esquerdo da equação a fórmula do estimador de mínimos quadrados para γ enquanto que no lado direito os termos com \mathbf{X} e com $\hat{\varepsilon}$ desaparecem.³ Assim sendo ficamos apenas com três componentes do lado direito da equação, estando cada um deles associado a um efeito fixo. A soma dos três componentes é igual ao hiato salarial por género ($\hat{\gamma}$). Ou seja,

$$\hat{\gamma} = \hat{\delta}_{\theta} + \hat{\delta}_{\varphi} + \hat{\delta}_{\lambda} \quad (6)$$

Em termos práticos, cada $\hat{\delta}$ do lado esquerdo da equação é o coeficiente de uma regressão entre a estimativa do efeito fixo e a variável binária para o sexo, controlando também para as variáveis \mathbf{X} . Se, condicional na matrix de variáveis \mathbf{X} , a distribuição das mulheres pelas empresas fosse perfeitamente aleatória então esperaríamos que $\hat{\delta}_{\varphi}$ fosse próximo de zero. Isto significaria que a alocação das mulheres/homens pelas empresas não era um factor que contribuía para o hiato salarial. Lógica semelhante aplica-se à alocação dos homens/mulheres pelas categorias profissionais.

3. Por construção $\hat{\varepsilon}$ é ortogonal a \mathbf{X} e a \mathbf{D} daí resultando que é também ortogonal a \mathbf{G} . Em consequência $\mathbf{M}_{\mathbf{G}}\hat{\varepsilon} = \mathbf{0}$. Como $\mathbf{M}\mathbf{X} = \mathbf{0}$ é fácil verificar que $\mathbf{M}_{\mathbf{G}}\mathbf{X} = \mathbf{0}$.

Discussão dos resultados

A decomposição de Machado e Mata

Nesta secção explora-se a metodologia de decomposição dos quantis para analisar as diferenças na distribuição salarial entre mulheres e homens ao longo de um período de 22 anos. Para esse fim, usamos o método de decomposição Machado e Mata (2005), o que nos permite identificar as origens das diferenças na distribuição dos salários entre homens e mulheres. Repetimos o exercício em 1991 e em 2013 com o objectivo de comparar a forma como as possíveis explicações para as diferenças encontradas evoluíram entre o início do período (1991) e 22 anos mais tarde (2013).

As diferenças de género na distribuição dos salários podem resultar de diferenças na distribuição das variáveis explicativas (diferenças em termos das características da população, tais como educação e idade) ou de diferenças na distribuição condicional dos salários (o que pode ser visto como alterações do modo como os salários reagem às características da população, os “coeficientes”). O primeiro exercício captura um “efeito composição” e o segundo pode ser visto como um “efeito estrutural” (Autor *et al.* (2008)). Assim, construímos o exercício contrafactual estimando a distribuição marginal de salários que teria prevalecido para o sexo masculino se eles tivessem as características do sexo feminino (“efeito de composição”). Posteriormente, estimamos a distribuição marginal de salários que teria prevalecido para as mulheres se elas tivessem os mesmos retornos do que os homens (“efeito estrutural”).

Em 1991, os homens ganhavam mais que as mulheres e a sua vantagem salarial era maior para níveis de remuneração mais elevados (percentis maiores). Enquanto que os homens ganhavam mais 42 por cento (35,1 na escala logarítmica) do que as mulheres na mediana, a diferença foi de 51,7 por cento (41,7 na escala logarítmica) no oitavo decil (ver a terceira coluna do Quadro 1). É evidente que, a partir da visualização das colunas 4 e 5, as diferenças nos coeficientes eram mais influentes, dominando a diferença global na distribuição de salários em comparação com as diferenças na distribuição das variáveis. Na mediana, a diferença salarial foi de 11,5 por cento (10,9 na escala logarítmica) devido a diferenças nas variáveis, e foi de 27,4 por cento (24,2 na escala logarítmica) devido a diferenças nos coeficientes. Curiosamente, o efeito de composição é maior no primeiro decil, mas o efeito estrutural tornar-se mais influente à medida que os níveis salariais aumentam. O efeito estrutural gerou um hiato de género maior nos salários mais altos (percentis maiores).

Em 2013, o hiato ainda existe, é positivo e estatisticamente significativo, mas a sua magnitude foi reduzida. Embora os homens ganhem na mediana mais 22,8 por cento (20,5 na escala logarítmica) do que as mulheres, a diferença foi reduzida entre o maior e o menor percentil. Em 2013, o efeito estrutural

	Mulheres x0b0 (1)	Homens x1b1 (2)	(2)-(1) (3)	Efeito composição agregado (x1b1-x0b1) (4)	Efeito estrutural agregado (x0b1-x0b0) (5)
10 percentile	-0.433*** (0.000)	-0.268*** (0.001)	0.165*** (0.001)	0.090*** (0.001)	0.074*** (0.000)
20 percentile	-0.351*** (0.000)	-0.116*** (0.000)	0.235*** (0.001)	0.096*** (0.001)	0.139*** (0.000)
30 percentile	-0.274*** (0.000)	0.010*** (0.001)	0.284*** (0.001)	0.101*** (0.001)	0.183*** (0.000)
40 percentile	-0.191*** (0.000)	0.130*** (0.001)	0.322*** (0.001)	0.105*** (0.001)	0.216*** (0.000)
50 percentile	-0.099*** (0.001)	0.251*** (0.001)	0.351*** (0.001)	0.109*** (0.001)	0.242*** (0.000)
60 percentile	0.008*** (0.001)	0.384*** (0.001)	0.375*** (0.001)	0.111*** (0.001)	0.264*** (0.000)
70 percentile	0.142*** (0.001)	0.539*** (0.001)	0.397*** (0.001)	0.114*** (0.001)	0.282*** (0.000)
80 percentile	0.321*** (0.001)	0.737*** (0.001)	0.417*** (0.001)	0.117*** (0.001)	0.300*** (0.000)
90 percentile	0.602*** (0.001)	1.036*** (0.001)	0.433*** (0.001)	0.123*** (0.002)	0.310*** (0.001)

QUADRO 1. Decomposição da discriminação salarial por género (1991)

	Mulheres x0b0 (1)	Homens x1b1 (2)	(2)-(1) (3)	Efeito composição agregado (x1b1-x0b1) (4)	Efeito estrutural agregado (x0b1-x0b0) (5)
10 percentile	-0.163*** (0.000)	-0.064*** (0.000)	0.099*** (0.001)	0.021*** (0.001)	0.078*** (0.000)
20 percentile	-0.063*** (0.000)	0.067*** (0.000)	0.130*** (0.001)	0.014*** (0.001)	0.117*** (0.000)
30 percentile	0.028*** (0.000)	0.186*** (0.001)	0.158*** (0.001)	0.008*** (0.001)	0.150*** (0.000)
40 percentile	0.124*** (0.000)	0.306*** (0.001)	0.183*** (0.001)	0.002** (0.001)	0.181*** (0.000)
50 percentile	0.229*** (0.001)	0.434*** (0.001)	0.205*** (0.001)	-0.004*** (0.001)	0.209*** (0.000)
60 percentile	0.349*** (0.001)	0.575*** (0.001)	0.226*** (0.001)	-0.009*** (0.001)	0.235*** (0.000)
70 percentile	0.495*** (0.001)	0.739*** (0.001)	0.244*** (0.001)	-0.015*** (0.001)	0.259*** (0.000)
80 percentile	0.684*** (0.001)	0.947*** (0.001)	0.262*** (0.001)	-0.016*** (0.001)	0.279*** (0.001)
90 percentile	0.968*** (0.001)	1.256*** (0.001)	0.288*** (0.001)	-0.013*** (0.002)	0.301*** (0.001)

QUADRO 2. Decomposição da discriminação salarial por género (2013)

domina em valor absoluto o efeito de composição em toda a distribuição (Quadro 2).

As mulheres, em 2013, não só são mais semelhantes aos homens, mas também apresentam melhores características (Quadro 3). As mulheres, em 2013, são trabalhadoras mais velhas e mais experientes refletindo o aumento da sua taxa de participação no mercado de trabalho. As qualificações da força de trabalho aumentou consideravelmente durante este período refletindo o envelhecimento da população. As mulheres, em 2013, estão a trabalhar em empresas maiores e são claramente mais educadas do que os seus colegas

	1991		2013	
	Mulheres	Homens	Mulheres	Homens
Idade	33,98	38,27	40,25	40,73
Antiguidade	8,87	10,17	9,47	9,55
Tamanho da empresa	5,09	5,50	4,82	4,77
Educação	6,36	6,27	9,86	9,29

QUADRO 3. Discriminação salarial por gênero: estatísticas descritivas (Composição)

	1991		2013	
	Mulheres	Homens	Mulheres	Homens
Idade	0,0190	0,0470	0,0267	0,0452
<i>Idade</i> ²	-0,0001	-0,0005	-0,0002	-0,0004
Antiguidade	0,0095	0,0104	0,0171	0,0216
<i>Antiguidade</i> ²	-0,0002	-0,0001	-0,0002	-0,0002
Tamanho da empresa	0,0422	0,0629	0,0258	0,0426
Educação	0,0739	0,0810	0,0723	0,0783
Proporção de mulheres	-0,2527	-0,1031	-0,2537	-0,1023
Constante	-1,0997	-1,6832	-1,2180	-1,7495

QUADRO 4. Discriminação salarial por gênero: regressões de quantis (β)

do sexo masculino. Existem diferenças significativas no retorno da educação tanto em 1991 como em 2013. Apesar de terem características semelhantes, o retorno do capital humano geral e específico é muito menor para as mulheres em comparação com os seus homólogos do sexo masculino (Quadro 4). As políticas salariais utilizadas pelas empresas maiores tendem a beneficiar os trabalhadores do sexo masculino numa extensão muito maior do que as mulheres. Por último, as empresas cuja força de trabalho é mais densamente povoada por mulheres (mais segregado) tende a gerar uma penalização salarial, principalmente, para os trabalhadores do sexo feminino.

A decomposição de Gelbach

O hiato salarial que identificamos constitui um diferencial médio entre os salários de dois trabalhadores de sexo diferente que de outra forma seriam totalmente idênticos. Uma questão fundamental diz respeito ao efeito da heterogeneidade não observada por trás dessas diferenças (ver Gráfico 3). De seguida, considera-se como a afectação entre as empresas com diferentes políticas de remuneração, a atribuição a categorias profissionais distintas e a alocação de trabalhadores com competências (não observadas) diferentes influenciam o hiato salarial entre os gêneros. O foco na decomposição da diferença salarial por gênero é, portanto, sobre as contribuições de cada uma dessas três fontes de heterogeneidade não observada.

Antes de prosseguir, vale a pena discutir a interpretação dos três efeitos fixos de elevada dimensionalidade adicionados na equação (3). O efeito fixo da empresa, no essencial, captura a política salarial (constante) da empresa. As empresas com políticas de remuneração generosas irão apresentar um efeito fixo de empresa positivo, empresas com políticas salariais de baixos salários irão gerar um efeito fixo negativo. No Gráfico 4 é comparada a distribuição dos efeitos fixos das empresas para os trabalhadores de ambos os géneros.⁴ É evidente a partir do gráfico que os homens preenchem de forma desproporcional as empresas com políticas salariais mais generosas.

No Gráfico 5 apresenta-se a distribuição empírica dos efeitos fixos dos trabalhadores. O efeito fixo do trabalhador condensa a influência das características constantes (observadas e não observadas) dos indivíduos sobre os seus salários. Estes efeitos podem ser uma medida para o capital humano (ou produtividade) do trabalhador ou podem simplesmente refletir a discriminação de género que não está associado com a afectação dos trabalhadores entre empresas e categorias profissionais. O gráfico mostra que a diferença salarial entre homens e mulheres está firmemente enraizada na componente individual de salários, mais notoriamente na aba superior da distribuição. Este resultado pode ser o resultado de características observadas ou não observadas. Mais à frente identifica-se o papel específico destas competências individuais não observadas.

Por fim, surge a distribuição empírica dos efeitos fixos das categorias profissionais. O efeito fixo da categoria profissional reflete em grande parte o nível da remuneração das tarefas a um nível muito desagregado. De certa forma, a inclusão de efeitos fixos da categoria profissional baseia-se na equação de salários Minceriana, que na sua primeira geração incluiu uma ampla definição de ocupações. Na configuração atual, é fornecida uma detalhada contabilização das tarefas necessárias para preencher uma categoria profissional. As distribuições dos efeitos fixos da categoria profissional apresentados no Gráfico 6 exibem uma diferença perceptível em termos de género, o que sugere que a alocação de trabalhadores nas diferentes categorias profissionais desfavorece significativamente as mulheres.

Os resultados da decomposição Gelbach são apresentados no Quadro 5. Observa-se que a penalização salarial de 22,6 pontos percentuais (25,6 na escala logarítmica) obtida a partir da estimativa da equação (1) pode ser decomposta na contribuição de três componentes distintas: trabalhador, empresa e categoria profissional. Uma parte significativa da diferença salarial é explicada pela heterogeneidade das políticas salariais das empresas. A alocação de trabalhadores pelas diferentes empresas é responsável por 5,6 pontos percentuais (5,8 dos 25,6 na escala logarítmica) do hiato salarial.

4. É de salientar, no entanto, que nesta comparação a influência de variáveis tais como a indústria ou o tamanho da empresa estão incorporadas no efeito fixo da empresa.

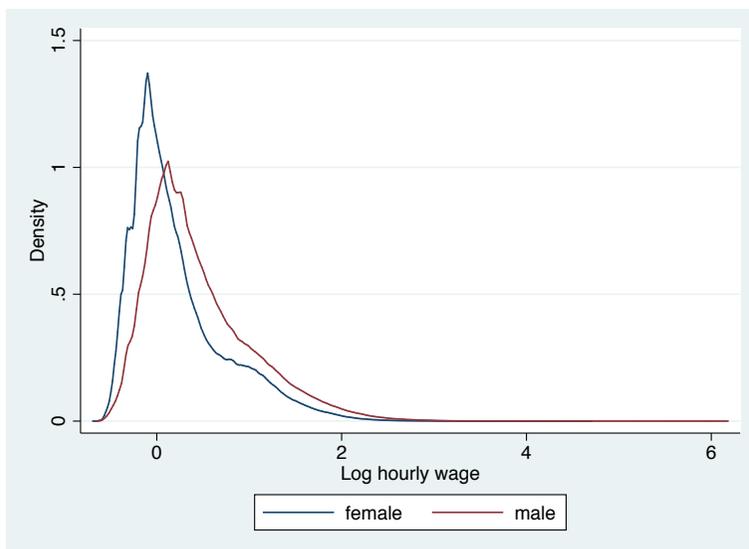


GRÁFICO 3: Discriminação salarial por género: (log) salário real por hora

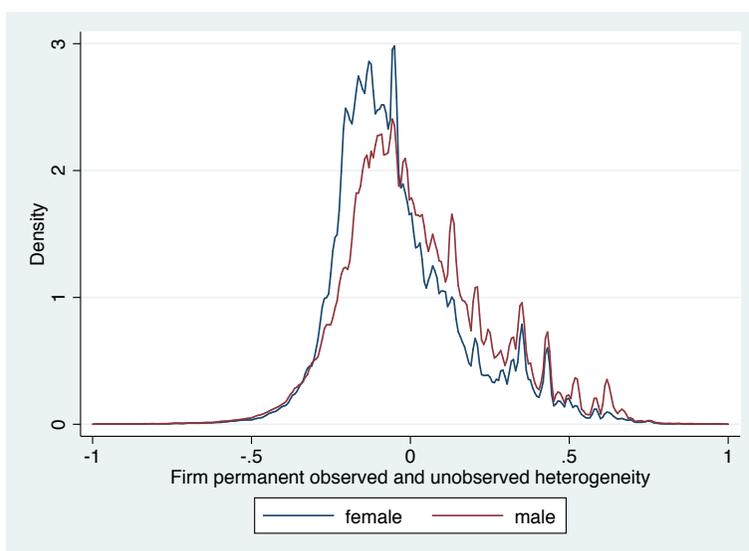


GRÁFICO 4: Discriminação salarial por género: Empresa

Isto significa que as mulheres estão empregadas desproporcionalmente em empresas com políticas salariais menos generosas. Dito de outro modo, se os trabalhadores fossem aleatoriamente atribuídos às empresas, o hiato salarial seria reduzido em cerca de um quinto. Na mesma tabela encontra-se evidência de que a afectação à categoria profissional, seja através de políticas de promoção ou por meio de atribuições iniciais, é significativamente

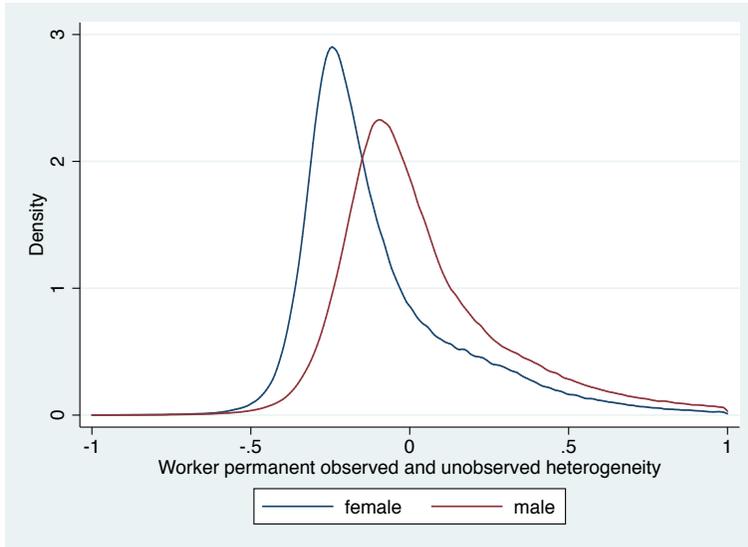


GRÁFICO 5: Discriminação salarial por gênero: Trabalhador

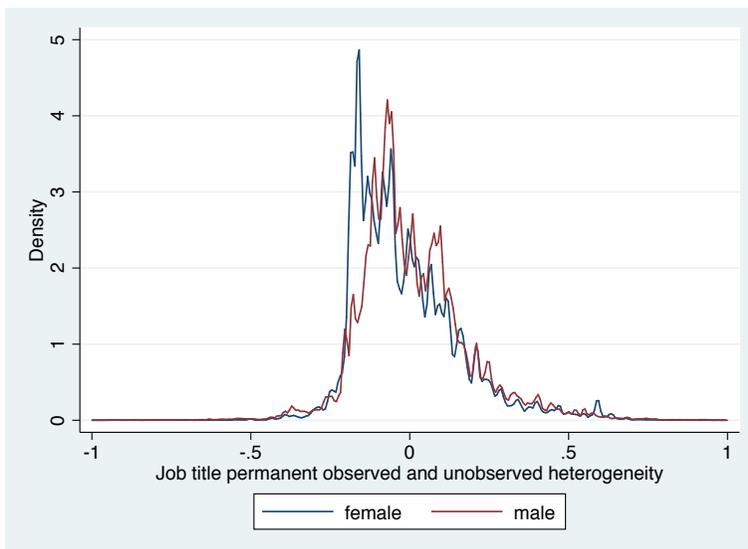


GRÁFICO 6: Discriminação salarial por gênero: Profissão

influenciada pelo gênero, contribuindo 4,3 pontos percentuais para aumentar o hiato. Em conjunto, o processo de triagem para as empresas e para as categorias profissionais são responsáveis por cerca de 40 por cento do hiato salarial. As características não observadas (permanentes) dos indivíduos são responsáveis pelos restantes 60 por cento. Estas características não observadas do trabalhador (pelo menos aos olhos do investigador) podem

ser equiparadas a competências não observadas ou, simplesmente, a alguma forma de discriminação de género.

gap	worker fe	firm fe	job fe
-0.2560	-0.1547	-0.0580	-0.0433

Note: Decomposições baseadas em Gelbach (2016).

QUADRO 5. Decomposição condicional do hiato salarial por género

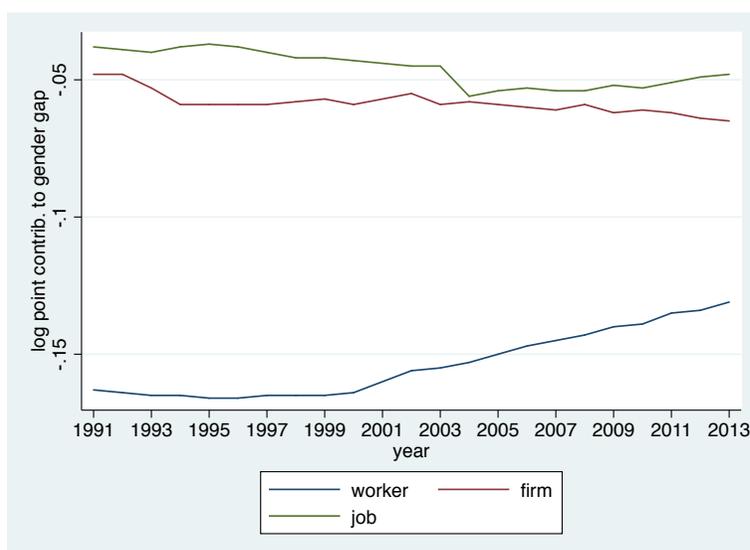


GRÁFICO 7: Decomposição condicional do hiato salarial por género (por ano)

O Gráfico 7 exibe a decomposição do hiato salarial ao longo do tempo. A alocação de trabalhadores do sexo feminino pelas empresas e categorias profissionais não sofreu grandes alterações ao longo das últimas duas décadas. No entanto, é detetável que a afectação às empresas e às categorias profissionais está agora ligeiramente menos favorável para as mulheres (-1,7 e -1,0 pontos percentuais cento para as empresas e para as categorias profissionais, respectivamente, ao longo do período 1991-2013). Em compensação, a penalização salarial resultante do papel da heterogeneidade individual não observada foi visivelmente atenuado (3,2 pontos percentuais), em especial desde o início do século. Em geral, a combinação da evolução dos três fontes de heterogeneidade resultou numa pequena (0,5 pontos percentuais) diminuição do hiato salarial entre os géneros ao longo destes 22 anos.

A decomposição do hiato salarial

A abordagem com base nos efeitos fixos de elevada dimensionalidade permite controlar totalmente para a distribuição de trabalhadores pelas empresas e categorias profissionais. No entanto, esta abordagem é silenciosa quanto à heterogeneidade das diferenças dentro da empresa e da categoria profissional. Os hiatos salariais determinados ao nível da empresa têm sido interpretados como evidência de discriminação de género que emerge da insuficiência de poder de negociação relativo das mulheres (Card *et al.* (2016)). Neste sentido, a abordagem anterior é generalizada para acomodar a estimativa dos hiatos salariais dentro da empresa e da categoria profissional. No essencial, para o caso da empresa, estima-se a seguinte equação de regressão:

$$\ln w_{ifjt} = \mathbf{x}_{ifjt}\beta + \varphi_f + \gamma_f g_i + \varepsilon_{ifjt} \quad . \quad (7)$$

onde a equação (1) é aumentada para incluir um efeito fixo por empresa (φ_f) e um efeito de género específico da empresa (γ_f). É importante salientar que não se está a incluir um efeito fixo do trabalhador e por isso a diferença de género ao nível da empresa não é filtrada da heterogeneidade do indivíduo (observada e não observada). A identificação do parâmetro do género específico da empresa em conjunto com o efeito do trabalhador fixo exigiria restrições adicionais de normalização, a fim de reter uma escala comum.

Os resultados da estimação da equação anterior estão expostos no Gráfico 8, em que a distribuição empírica dos hiatos salariais por empresa para 1991 são contrastados com os de 2013. O histograma pode ser interpretado como a distribuição da discriminação dos empregadores (ou, no sentido de Card *et al.* (2016) , como o reflexo do poder de negociação das mulheres). O gráfico indica que a maioria dos empregadores têm diferenças salariais de género negativas e que a distribuição dos hiatos salariais se deslocou de forma limitada entre 1991 e 2013. É interessante notar que uma fração não desprezível de empregadores praticam hiatos salariais positivos, ou seja, pagam salários mais elevados às suas trabalhadoras.

O facto dos hiatos salariais determinados ao nível da empresa serem altamente correlacionados com a segregação ao nível da empresa é uma indicação de que não é simplesmente consequência da variação amostral (o coeficiente de correlação linear é igual a -0,476). A indicação de que uma maior proporção de mulheres leva a uma diferença salarial mais negativa é consistente com a ideia de que as mulheres têm um défice de poder de negociação.

Em comparação com a distribuição por empresa, a distribuição das estimativas do hiato salarial por categoria profissional exhibe uma menor dispersão, em especial no ano de 2013. Em contraste com os hiatos salariais determinados ao nível da empresa, os hiatos salariais determinados ao nível da categoria profissional não são influenciados pelo grau de segregação

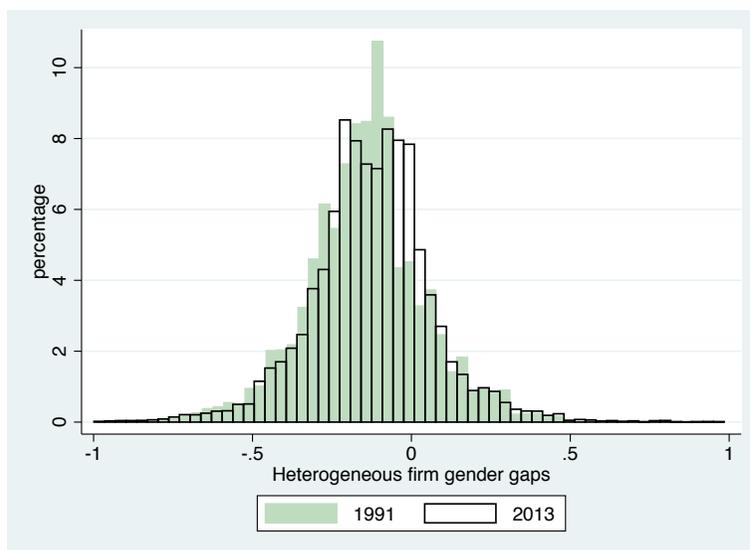


GRÁFICO 8: Hiatos salariais por empresa

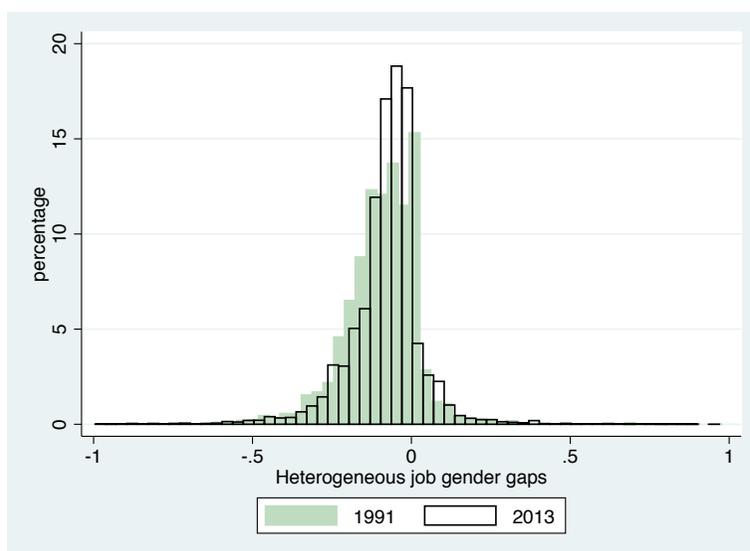


GRÁFICO 9: Hiatos salariais por categoria profissional

sexual (o coeficiente de correlação linear é igual a 0,006). Isto é, enquanto empresas mais segregadas tendem a exibir maiores desigualdades salariais entre homens e mulheres, as categorias profissionais levam a salários mais baixos mas não a mais elevados hiatos salariais. Uma explicação possível para este resultado é a sobreocupação feminina de algumas profissões (Groshen (1991)).

Conclusões

Ao longo do período compreendido entre 1991 e 2013, o mercado de trabalho português observou um aumento notável da participação de mulheres, que se traduziu num aumento de 10 pontos percentuais na taxa de feminização entre os trabalhadores empregados no sector empresarial. Este movimento foi acompanhado por uma significativa aproximação entre os salários médios das mulheres e dos homens. Neste estudo, argumenta-se, porém, que a atenuação do hiato salarial se deveu, no essencial, a alterações na composição das qualificações dos trabalhadores em favor do contingente feminino e não a mudanças estruturais associadas à noção de discriminação sexual.

De facto, neste período, a medida do hiato salarial ajustada para as características observadas dos trabalhadores manteve-se relativamente constante em torno de 25 por cento. No primeiro exercício de decomposição mostrou-se que a afetação dos trabalhadores a empresas e categorias profissionais é decisivamente influenciada pelo género. Se assim não fosse, isto é, se os trabalhadores fossem alocados aleatoriamente às empresas, o hiato salarial seria reduzido em 5,8 pontos percentuais. Identicamente, se a afetação dos trabalhadores às categorias profissionais fosse independente do género, o diferencial de salários entre homens e mulheres seria atenuado em 4,3 pontos percentuais. Conjuntamente, a distribuição seletiva dos trabalhadores por empresas e categorias profissionais com diferentes políticas salariais é responsável por cerca de dois quintos do hiato salarial.

A afetação das mulheres às empresas e categorias profissionais não melhorou ao longo das últimas duas décadas. Pelo contrário, deteriorou-se um pouco, uma vez que em 2013 as mulheres estão menos presentes em empresas e categorias profissionais com políticas de remuneração mais generosas. Em compensação, porém, o papel da parte não observada das qualificações das mulheres favoreceu uma diminuição contida do hiato salarial. Esta indicação poderá ser interpretada como o efeito duma diminuição (limitada) da discriminação sexual.

Por fim, o exercício de decomposição dos hiatos salariais por empresa revelou que a magnitude desses hiatos salariais é poderosamente influenciada pela taxa de feminização que se observa para a empresa, o que parece sugerir que as mulheres detêm menos poder negocial do que os homens. Em contraposição, o nível de segregação sexual por categoria profissional, que está decisivamente associado a salários mais baixos, não parece ter impacto significativo na medida de hiato salarial observado dentro de cada categoria profissional.

Referências

- Abowd, John M., Francis Kramarz, e David N. Margolis (1999). "High Wage Workers and High Wage Firms." *Econometrica*, 67(2), 251–334.
- Altonji, Joseph G. e Rebecca M. Blank (1999). "Race and gender in the labor market." In *Handbook of Labor Economics*, vol. 3, edited by Orley Ashenfelter e David Card, pp. 3143–3259. Elsevier.
- Autor, David H. (2013). "The "task approach" to labor markets : an overview." *Journal for Labour Market Research*, 46(3), 185–199.
- Autor, David H., Lawrence F. Katz, e Melissa S. Kearney (2006). "The Polarization of the U.S. Labor Market." *American Economic Review*, 96(2), 189–194.
- Autor, David H., Lawrence F. Katz, e Melissa S. Kearney (2008). "Trends in U.S. Wage Inequality: Revising the Revisionists." *The Review of Economics and Statistics*, 90(2), 300–323.
- Bertrand, C. Goldin, M. e L. F. Katz (2010). "Dynamics of the gender gap for young professionals in the financial and corporate sectors." *American Economic Journal: Applied Economics*, 2(3), 228–255.
- Blau, F. e L. Kahn (2016). "The gender wage gap: Extent, trends and explanations." *Journal of Economic Literature*, forthcoming.
- Card, David, Ana Rute Cardoso, e Patrick Kline (2016). "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women." *The Quarterly Journal of Economics*, forthcoming.
- Cardoso, Ana Rute, Paulo Guimarães, e Pedro Portugal (2016). "What drives the gender wage gap? A look at the role of firm and job-title heterogeneity." *Oxford Economic Papers*, 68(2), 506–524.
- Dustmann, Christian, Johannes Ludsteck, e Uta Schönberg (2009). "Revisiting the German Wage Structure." *The Quarterly Journal of Economics*, 124(2), 843–881.
- Gelbach, Jonah (2016). "When do covariates matter? and which ones, and how much?" *Journal of Labor Economics*, forthcoming.
- Goldin, C. (2014). "A grand convergence: Its last chapter." *American Economic Review*, 104(4), 1091–1119.
- Goos, Maarten e Alan Manning (2007). "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain." *The Review of Economics and Statistics*, 89(1), 118–133.
- Groshen, Erica L. (1991). "The structure of the female/male wage differential: Is it who you are, what you do, or where you work?" *Journal of Human Resources*, 26(3), 457–472.
- Guimarães, Paulo e Pedro Portugal (2010). "A simple feasible procedure to fit models with high-dimensional fixed effects." *Stata Journal*, 10(4), 628–649.
- Jarrell, S. B. e T. D. Stanley (2004). "Declining bias and gender wage discrimination? a meta-regression analysis." *Journal of Human Resources*, 39(3), 828–838.

- Machado, José A. F. e José Mata (2005). "Counterfactual decomposition of changes in wage distributions using quantile regression." *Journal of Applied Econometrics*, 20(4), 445–465.
- Olivetti, C. e B. Petrongolo (2008). "Unequal pay or unequal employment? A cross-country analysis of gender gaps." *Journal of Labor Economics*, 26(4), 621–654.
- Stanley, T. e S. B. Jarrell (1998). "Declining bias and gender wage discrimination? a meta-regression analysis." *Journal of Human Resources*, 33(4), 947–973.
- Torres, Sónia, Pedro Portugal, John T. Addison, e Paulo Guimarães (2013). "The sources of wage variation: a three-way high-dimensional fixed effects regression model." Working Papers w201309, Banco de Portugal, Economics and Research Department, URL <http://ideas.repec.org/p/ptu/wpaper/w201309.html>.
- Weeks, David L. e Donald R. Williams (1964). "A Note on the Determination of Connectedness in an N-Way Cross Classification." *Technometrics*, 6(3), 319–324.