

# **The London Difference in Gender Pay Gaps\***

**Mark B Stewart  
University of Warwick**

**July 2014**

## **Abstract**

The median gender pay gap is higher in London (and the south-eastern corner of Britain more generally) than in the rest of Britain, but this is a relatively recent phenomenon. The comparison between gender pay gaps in London and the rest of the country differs at different points in the wage distribution, as does the extent to which the difference is compositional. The gender pay gap for London is greater than that for the rest of Britain above about the 40<sup>th</sup> percentile of the wage distribution, but the phenomenon of higher gender pay gaps in London for other than compositional reasons is one that is confined to the top one-third of the wage distribution.

---

\* The author thanks the Office for National Statistics and the UK Data Archive for permission to use the Annual Survey of Hours and Earnings and New Earnings Survey data and the Secure Data Service for access. Responsibility for the analysis and interpretation of the data is solely that of the author. Address for correspondence: Economics Department, University of Warwick, Coventry, CV4 7AL, UK. E-mail: Mark.Stewart@warwick.ac.uk.

## 1. Introduction

The gender pay gap has fallen dramatically in Britain over the last 40 years. However this fall has not been equally paced across the regions of the country. In particular the fall has been far less in London and the south-eastern corner of Britain than in the rest of the country. The gender pay gap is now higher in London (and the south-eastern corner of Britain more generally) than in the rest of the country, which contrasts with the relative position even 10 years ago. This paper examines why this is the case and in particular analyses the extent to which the cause is compositional, i.e. due to differences in the characteristics of employees and jobs between London (or the south-eastern corner more generally) and the rest of the country.<sup>1</sup>

The headline measure favoured by the Office for National Statistics (ONS), the percentage difference in median hourly earnings (excluding overtime) of full-time employees, has fallen steadily over the last 40 years from nearly 30% in 1975 to slightly below 10% in 2012. Over the same period median pay in London and the south eastern corner of Britain has pulled away from that in the rest of Britain. In addition these increases in regional wage differentials have been greater for men than for women. In 1975 median pay in London, on the same measure as above, was 14% above that for Great Britain (GB) as a whole for men and 25% above for women. By 2012 it was 36% above for men and 33% above for women. This greater widening of the London premium for men than for women is also seen, although to a lesser extent, in the South East and East of England regions.

Correspondingly, the gender pay gap has fallen much less in the south eastern corner of Britain than in the rest of Britain and is larger in these regions in 2012 than in the other regions of Britain. In 1975 the gender pay gap was smaller in the South East, the East of England, and particularly London, than in GB as a whole, whereas all are larger than that for GB as a whole in 2012. The gender pay gap for GB as a whole has fallen by nearly 20 percentage points over this period, whereas that for London has only fallen by 10 percentage points.

---

<sup>1</sup> See Altonji and Blank (1999) and Bertrand (2011) for reviews of the extensive literature on the gender pay gap. See Leaker (2008), Manning and Swaffield (2008), Shackleton (2008) and DCMS (2014) for recent evidence on the gender pay gap in Britain.

This paper investigates this geographic variation in gender pay gaps. It examines the extent to which this variation is compositional, and in particular due to different mixes of occupations in the different regions. In an inter-temporal context the corresponding issue is why the gender pay gap in the south east corner of Britain has fallen so much more slowly than that in the remainder of Britain.

## **2. Gender pay gaps**

The ONS preferred headline measure of the gender pay gap is that in hourly earnings excluding overtime and for full-time employees. This is seen by ONS as providing the most appropriate comparison (see Hicks and Thomas, 2009). The choice is motivated by the fact that women are far more likely to work part-time than men and that men work more overtime hours than women. The hourly earnings measure is chosen over weekly earnings since, even among full-timers, women work fewer hours per week than men. The gap in the median is preferred for the headline figure by ONS because it is less susceptible to the influence of extreme values than the mean. This ONS headline measure will be used as the main measure of the gender pay gap in this paper too, but means (arithmetic and geometric) of the same hourly earnings measure will be looked at too since, in contrast to the headline measure, almost all the existing empirical literature on decomposing gender pay gaps focuses on the gap in means (often in the log of wages). In addition gender pay gaps at different percentiles in the wage distribution will also be examined.

The most accurate and comprehensive data source for data on earnings in Britain is the Annual Survey of Hours and Earnings (ASHE). The earnings statistics presented in this paper are based on this data source and its predecessor, the New Earnings Survey (NES). These provide the most accurate information on individual earnings available for Britain. The ASHE, developed from the earlier NES, is conducted in April of each year. It surveys all employees with a particular final two digits to their National Insurance numbers who are in employment and hence aims to provide a random sample of employees in employment. The ASHE is based on a sample of employees taken from HM Revenue and Customs PAYE records. Information on earnings and paid hours is obtained in confidence from employers, usually downloaded directly from their payroll computer records. It therefore provides very accurate information on earnings and paid hours. Indeed providing accurate information to

the survey is a statutory requirement on employers under the Statistics of Trade Act. Technical details of the ASHE are given in Bird (2004).

The median gross hourly earnings (excluding overtime) of male full-time employees in Great Britain in April 2012, according to the ASHE, was £13.33, while that for female full-time employees was £12.02.<sup>2</sup> Thus, considering full-time employees only, women earn £1.31 per hour or 9.8% less than men.<sup>3</sup> This percentage differential is commonly known as the gender wage differential, gender wage gap or gender pay gap. The figure is for a particular pay measure and a particularly defined employee group. The gender gap in means is larger than that in medians. The mean gross hourly earnings (excluding overtime) of male full-time employees for GB in April 2012 was £16.58, while that for female full-time employees was £14.08, giving a gender gap in the means of 15.1%

Gender pay gaps by region for selected years are given in Table 1. The left-hand half of the table gives these are in terms of medians as in the ONS-favoured measure. The 11 regions given are the current standard ONS regions (formerly known as “government office regions”) within Great Britain.<sup>4</sup> In 2012 the gender pay gap is largest in London, the East of England and the South East and all are larger than that for GB as a whole. In contrast, in 1975 all three of these regions had gender pay gaps smaller than the overall gap and the gap was the smallest in these three regions. By 1988 the gaps for the South East and the East of England were above the overall gap, but that for London was still below and was still the smallest of all the regions. The gap for London was still below that for GB in 2000, but by 2012 the overall GB gap had fallen below that for London.

Between 1975 and 2012 the gender pay gap for GB as a whole fell by 17.9 percentage points, while that for London fell by 9.1, that for the South East by 11.4 and that for the East of England by 13.6. These three regions exhibited the smallest falls (in percentage point terms) among the 11 regions.

---

<sup>2</sup> The earnings information collected by the ASHE relates to gross pay before tax, NI or other deductions and excludes payments in kind. Earnings figures given in this paper are based on employees on adult rates whose pay was unaffected by absence. This is the standard sample definition for ONS headline earnings figures. Full-time employees are defined as those who work more than 30 paid hours per week (or those in teaching professions who work more than 25 paid hours per week).

<sup>3</sup> This figure is the gender pay gap for GB. The ONS headline gap is for the UK and is slightly lower at 9.6%.

<sup>4</sup> Earlier regional classifications have been converted to the current ONS regions for this table and the rest of the analysis in this paper.

For comparison the right-hand half of Table 1 gives the gender gaps in mean earnings by region. The rankings of the regions are slightly different to those in terms of medians. In particular for 2012 the gender gaps in London and the South East are considerably bigger than that in the East of England. However it is still the case for means, as it was for medians, that these three regions have the largest gender gaps in 2012 and also that these three regions exhibited the smallest falls (in percentage point terms) in the gender gap.

Gender pay gaps (in medians) for London, the South East, the East of England and GB as a whole are shown in Figure 1 for each year from 1975 to 2012. Clearly the gender pay gap has fallen over this period in all cases. However it can also be seen that the rate of decline has been rather slower in the South East and the East and particularly London than that in the gap for GB as a whole. Indeed for London it would seem that there has been almost no fall in the gap since 1991. Some caution is required in such inter-temporal comparisons due to methodological changes in the surveys (see Appendix for details), but even allowing for these discontinuities the change in the London gender pay gap since 1991 seems to have been relatively small.

The time-series paths of the median pay of men and women in London, the South East and the East of England relative to the median for GB as a whole are shown in Figure 2. In all three regions male pay has risen faster than female pay (relative to the GB rise). For London, male pay was 14% above the GB median in 1975 and this had risen to 36% above in 2012, while for women it had risen from 25% to 33% above. For the South East male pay was 3% below the GB median in 1975 and had risen to 9% above it in 2012, while for women it had risen from 1% below the GB median to 3% above it. For the East region the picture is slightly different. Male pay was 2% below the GB median in 1975 and had fallen to 3% below it in 2012, while for women it had fallen from 1% below the GB median to 6% below it. However for all three regions the growth in relative pay (relative to GB as a whole) has been considerably greater for men than for women.

It is also instructive to look beyond these comparisons of medians.<sup>5</sup> It is well known that the shapes of the wage distributions for men and women differ considerably, resulting in gender

---

<sup>5</sup> This part of the analysis uses the ASHE microdata Secure Access files: Office for National Statistics (2013).

pay gaps being higher near the top of the wage distribution than near the bottom. As well as being true for GB as a whole, this is also true for each of the regions being considered here. Figure 3 shows the gender pay gaps for 2012 at each percentile of the wage distribution for the three regions focussed on above (London, South East and East) and for the rest of GB. In all four cases the gender pay gap rises as we move up the distribution. For the rest of GB first, the gender pay gap profile is fairly flat between about the 10th and 80th percentiles. It rises sharply up to the 10th percentile and particularly at the top of the distribution.

Relative to the rest of GB the rise in gender pay gap in the upper part of the distribution starts earlier in the East of England and even earlier in the South East. The rise starts earliest of all for London and indeed barely has a flat part to the profile. The London gender pay gap profile rises fast in the top half of the distribution. It goes from a lower gender pay gap than the other regions (between about the 5th and 40th percentiles) to a higher one than the other regions (above about the upper quartile). London's gender pay gap overtakes that for the rest of GB at about the 40th percentile, the East region at about the median, and the South East at about the upper quartile. In the following analysis it is therefore of interest to compare adjusted gender pay gaps at other points of the wage distribution as well as the median, and particularly at points above the median.

### **3. Estimation of adjusted wage gaps**

Wage gaps are often decomposed into a part that is “explained” by a set of factors and a part that is left “unexplained” by these factors. The latter can also be viewed as the “adjusted” wage gap, after removal of the effect of the specified set of factors. Suitable econometric decomposition methods are discussed in the survey by Fortin et al (2011). The classic regression-based decomposition proposed by Oaxaca (1973) and Blinder (1973) is now a standard technique in many areas of applied economics. It provides a method for decomposing the difference in the mean of an outcome variable between two groups or between two time periods and has been applied in a vast array of studies. Almost all studies that decompose gender wage gaps analyse the mean gap and apply the Oaxaca-Blinder (OB) decomposition.

This paper analyses both adjusted mean and adjusted median gender wage gaps: mean gaps because that is the main focus in the existing literature and of the standard decomposition technique, median gaps because this is now the favoured headline measure for the UK gender pay gap.

Consider first the decomposition or adjustment of the mean wage gap between two groups,  $A$  and  $B$ . The outcome variable of interest,  $y$ , is the log(wage). The estimated raw or overall mean wage gap is given by

$$\hat{\Delta}_o^\mu = \bar{y}_A - \bar{y}_B$$

If group  $A$  is women and group  $B$  men, then this is approximately the (negative) gender mean wage gap. Alternatively  $A$  and  $B$  might be London and the rest of GB or they could also be two time periods.

Linear regression models are specified for the log(wage) in each group

$$y_g = x' \beta_g + \varepsilon_g \quad g = A, B$$

The estimated raw mean wage gap can therefore also be written as

$$\hat{\Delta}_o^\mu \equiv \bar{y}_A - \bar{y}_B = \bar{x}_A' \hat{\beta}_A - \bar{x}_B' \hat{\beta}_B$$

One form of the OB decomposition can then be written as

$$\hat{\Delta}_o^\mu \equiv \bar{y}_A - \bar{y}_B = (\bar{x}_A - \bar{x}_B)' \hat{\beta}_A - \bar{x}_B' (\hat{\beta}_A - \hat{\beta}_B)$$

The first term,  $\hat{\Delta}_x^\mu = (\bar{x}_A - \bar{x}_B)' \hat{\beta}_A$ , is the “explained” part of the gap, also known as the composition effect. The second term,  $\hat{\Delta}_s^\mu = \bar{x}_B' (\hat{\beta}_A - \hat{\beta}_B)$ , is the “unexplained” part of the gap, also known as the wage structure effect. This component is the “adjusted wage gap” (after adjusting for the effect of the  $x$ -variables).

This “unexplained” component of an OB decomposition can also be interpreted as a “treatment effect” of the type that has been extensively studied in the program evaluation literature.<sup>6</sup> Fortin et al (2011) stress the usefulness of this link “to (i) clarify the assumptions underneath popular decomposition methods, (ii) propose estimators for some of the elements of the decomposition, and (iii) obtain formal results on the statistical properties of the various decomposition terms”.

There are several useful ways of viewing this decomposition and resulting adjusted wage gap. First, the adjusted wage gap combines the coefficient differences between the two groups using the group  $B$  means, i.e. the means of the “control” or “reference” group. Second, in the explained part, the composition effects of the  $x$ -variables are combined together using the group  $A$  coefficients in the adjustment for compositional differences, i.e. the coefficients of the “treatment” or “subject” group.

Third, the adjusted wage gap can also be written as

$$\hat{\Delta}_s^\mu \equiv \bar{x}'_B(\hat{\beta}_A - \hat{\beta}_B) = \bar{x}'_B\hat{\beta}_A - \bar{x}'_B\hat{\beta}_B = \hat{y}_A^C - \bar{y}_B$$

where  $\hat{y}_A^C = \bar{x}'_B\hat{\beta}_A$ , the counterfactual average log wage of group  $A$  if they had the average  $x$ -values of group  $B$ . So in the case of comparing women with men it estimates the counterfactual average if women were still paid as women (i.e. according to the female wage structure), but had the same average characteristics as the men. In the case of comparing London with the rest of GB it estimates the counterfactual average if London had the same average characteristics as the rest of GB, but was still paid according to the London wage structure.

Alternatively the counterfactual can be viewed as what men would be paid on average if they were paid according to the female wage structure, or what the average wage for the rest of GB would be if those working there were paid according to the London wage structure.

---

<sup>6</sup> Specifically it is the average treatment effect on the treated (ATT).

The counterfactual average log wage is then compared to the actual group  $B$  average log wage. So the adjusted mean wage gap is an estimate of what the mean gap would be if the two groups had the same average characteristics (i.e.  $x$ -variables).

Adjusted wage gaps can be constructed more generally using simple re-weighting estimators (also known as inverse probability weighting, inverse propensity weighting or propensity score weighting). For performing a decomposition for distributional statistics, such as here, Fortin et al (2011) in their survey argue that “reweighting is the method of choice” (page 74) because first it is simple to implement and second there are well established results in the program evaluation literature that show that the method is asymptotically efficient (Hirano et al., 2003; Firpo, 2007).

The OB estimator of the counterfactual mean can be shown to be a propensity score reweighting estimator based on a linear model for the conditional odds of being “treated”, i.e. of being a member of group  $A$  (Kline, 2011).

Using a reweighting estimator to estimate the counterfactual median and the adjusted median wage gap, the focus is on cumulative distributions, since the median can then be obtained by inversion. The approach constructs a counterfactual distribution which combines the characteristics of group  $B$  (e.g. men) with the wage structure of group  $A$  (e.g. women). It estimates the wage distribution that group  $A$  (e.g. women) would have if they had the same distribution of characteristics,  $x$ , as group  $B$  (e.g. men). Alternatively it can be viewed as an estimate of the wage distribution that group  $B$  (e.g. men) would have if they were paid like group  $A$  (e.g. women) workers. For the London vs. rest of GB case the counterfactual distribution estimates the wage distribution that London would have if those who worked there had the same distribution of characteristics as those in the rest of GB.

The method replaces the distribution of the vector of characteristics,  $x$ , of group  $A$  with the distribution of  $x$  of group  $B$  by using the re-weighting factor  $\Psi(x)$  defined as follows. Construct a dummy variable  $D_B$  that takes the value 1 in group  $B$  and the value 0 in group  $A$ . The re-weighting factor can then be written as

$$\Psi(x) = \frac{\Pr(x | D_B = 1)}{\Pr(x | D_B = 0)} = \frac{\Pr(D_B = 1 | x) / \Pr(D_B = 1)}{\Pr(D_B = 0 | x) / \Pr(D_B = 0)}$$

This can be easily computed by estimating a probability model for  $\Pr(D_B = 1 | x)$  and using the predicted probabilities to compute estimates of  $\Psi(x)$  for observations in group A. DiNardo et al (1996) suggest estimating a flexible logit model. Hirano et al (2003) propose using a non-parametric logit model, which is a series estimator applied to a logit model involving polynomial terms in the covariates of increasing order.

The steps involved in practical implementation are as follows (Fortin et al, 2011, p.65):

(1) Pool the data for groups A and B and run a logit or probit model for the probability of belonging to group B:

$$\Pr(D_B = 1 | x) = 1 - \Pr(D_B = 0 | x) = 1 - \Pr(\varepsilon > -h(x)\alpha) = \Lambda(-h(x)\alpha),$$

where  $\Lambda(\cdot)$  is a normal or logit link function and  $h(x)$  is a polynomial or other function of  $x$ .

(2) Estimate the reweighting factor  $\Psi(x)$  for observations in group A using the predicted probabilities of belonging to group B,  $\Pr(D_B = 1 | x)$ , and A,  $\Pr(D_B = 0 | x) = 1 - \Pr(D_B = 1 | x)$ , and the sample proportions in group B,  $\Pr(D_B = 1)$ , and A,  $\Pr(D_B = 0)$ , in the expression for  $\Psi(x)$  above.

(3) Compute the counterfactual statistic of interest (in this case the median) using observations from the group A sample reweighted using  $\Psi(x)$ .

The adjusted median wage gap, for example, is then the difference between this estimated counterfactual median and the actual median for group B.

#### 4. Empirical evidence

This section examines the extent to which regional differences in gender pay gaps are compositional, that is to say due to differences in job or employee characteristics. This paper focuses on the difference between London (and the south-eastern corner of Britain) on the one hand and the rest of GB on the other. It also focuses on raw gender pay gaps. Specifically it examines the counter-factual which asks: What would the London gender pay gap be if the characteristics of London female employees and their jobs were instead those of female employees in the rest of GB and the characteristics of London male employees and their jobs were instead those of male employees in the rest of GB?

Note that it is not addressing the counterfactual which asks: What would the London gender pay gap be if the characteristics of London female employees and their jobs were instead those of London male employees? This is a different question, also of interest of course, but not the focus here. This paper is concerned with the effect on regional gender pay gaps of the employee composition being different between London and the rest of GB, rather than of it being different between men and women.

The analysis here looks at gender pay gaps at various points in the distribution. Section 2 above presents both mean and median raw gender pay gaps and also looks at the profile of the gap across the distribution. Section 3 describes suitable techniques for the estimation of adjusted wage gaps at both the mean and the median and how these can be extended to other points in the distribution. The entire profile of the adjusted gap across the distribution is examined here in this way. The existing empirical economics literature on gender pay gaps most often uses the Oaxaca-Blinder decomposition to analyse the mean gap in the log of wages. It therefore typically considers the geometric mean of wages. This point in the distribution will additionally be analysed here. The analysis here will therefore examine gender pay gaps at percentile points across the distribution and additionally at the arithmetic and geometric means of the distribution.

The analysis here also considers alternative definitions of London or the south-eastern corner, the focal region of the analysis. The main comparison is for the standard ONS region definition of London. The second comparison then uses a broader concept, the south-eastern corner of Britain, which includes the South East and East of England standard ONS regions in addition to London. In contrast the third comparison takes a narrower area, just Inner London. To facilitate comparison between the three comparisons all three use the same comparator: “the rest of GB”, defined to be GB excluding the London, South East and East of England standard regions (i.e. excluding the whole of the south-eastern corner of Britain).

The following control variables are used in the analysis:

Controls for 2-digit occupations: 25 such occupations are identified and thus 24 dummy variables are included in the specification.

The age of the employee.

The log of the number of employees in the company.

Whether the employee's pay is set with reference to a collective agreement (a dummy variable).

Whether working in the public sector (central government, local authority, public corporation or nationalised industry) (a dummy variable).

The results of the analysis are presented in Table 2 and Figures 4 to 6. Looking at the standard definition of London first, Figure 4 shows the gender pay gaps at each percentile of the wage distribution for London and for the "Rest of GB" region. As stated in Section 2 the gender pay gap is higher in London than the Rest of GB above about the 40<sup>th</sup> percentile. The figure also shows the counterfactual gender gap at each percentile, calculated as described in the previous section using the control variables listed above. This is an estimate of what the gender pay gap profile would be if employees and jobs in London had the characteristics of those in the Rest of GB region. This is below the actual London gender pay gap above about the 20<sup>th</sup> percentile. It is also below the Rest of GB region gender pay gap between the 6<sup>th</sup> and 63<sup>rd</sup> percentiles. The actual London gender pay gap overtakes that for the Rest of GB region at about the 40<sup>th</sup> percentile. The counterfactual London one (i.e. the one with Rest of GB characteristics) does not do so until the 64<sup>th</sup> percentile.

Gender pay gaps for certain percentiles and the arithmetic and geometric means are given in Table 2, together with the bootstrap standard errors for the counterfactual estimates.<sup>7</sup> The gender pay gap between the medians is 12.0% in London and 9.2% in the Rest of GB region, a difference of just under three percentage points. The counterfactual London gender pay gap is 6.3%, i.e. just under three percentage points below that for the Rest of GB region. Thus if employees and jobs in London had the characteristics of those in the Rest of GB region, the gender pay gap between the medians would be three percentage points below that for the Rest of GB region rather than the three percentage points above that it actually is. At the median the entire difference between the gaps (and more) is due to differences in characteristics between London and the Rest of GB region.

As we move up the wage distribution the proportion of the difference in gender pay gaps between London and the Rest of GB that is removed by adjustment for the counterfactual declines. At the geometric mean the counterfactual reduces the London gender pay gap

---

<sup>7</sup> The bootstrap standard errors are based on 1000 replications.

almost all the way to that for the Rest of GB. At the 70<sup>th</sup> percentile it reduces the difference by about three-quarters; at the arithmetic mean by about two-thirds; and at the 90<sup>th</sup> percentile by about a half.

The broad findings just described for London are roughly the same when we either broaden the focal region from London to the whole of the south-eastern corner of Britain (Figure 5 and the second block of Table 2) or narrow it from the whole of London to Inner London only (Figure 6 and the third block of Table 2).

## **5. Concluding comments**

The comparison between gender pay gaps in London (or the south-eastern corner more generally) and the rest of the country differs at different points in the wage distribution, as does the extent to which the difference is compositional.

The higher gender pay gaps in the south-eastern corner of Britain, and particularly London, are primarily a phenomenon of the upper half of the wage distribution. In addition below about two-thirds of the way up the distribution the higher gender pay gap in London in the upper half of the distribution can be entirely accounted for by compositional differences. Indeed for much of that part of the distribution the counterfactual distribution is below that for the rest of GB. That is to say, if London had the same characteristics as the rest of GB its gender pay gap would be lower than the rest of GB rather than higher. It is only in the top one-third of the wage distribution that the London gender pay gap is above that for the rest of GB for non-compositional reasons.

## References

- Altonji, Joseph G. and Blank, Rebecca M. (1999), “Race and gender in the labor market”, in *Handbook of Labor Economics*, Volume 3, Edited by O. Ashenfelter and D. Card, Elsevier Science.
- Bertrand, Marianne (2011), “New perspectives on gender”, *Handbook of Labor Economics*, Volume 4b, Edited by O. Ashenfelter and D. Card, Elsevier Science
- Bird, D. (2004). “Methodology for the 2004 Annual Survey of Hours and Earnings”, *Labour Market Trends*, November, 457-464.
- Blinder, Alan (1973). “Wage discrimination: reduced form and structural estimates”, *Journal of Human Resources*, 8, 436–455.
- Department for Culture, Media and Sport (2014). “Secondary analysis of the gender pay gap”, March.
- DiNardo, John, Fortin, Nicole M. and Lemieux, Thomas (1996). “Labor market institutions and the distribution of wages, 1973-1992: a semiparametric approach”, *Econometrica*, 64, 1001–1044.
- Firpo, Sergio (2007). “Efficient semiparametric estimation of quantile treatment effects”, *Econometrica*, 75, 259–276.
- Fortin, Nicole, Lemieux, Thomas and Firpo, Sergio (2011). “Decomposition methods in economics”, in *Handbook of Labor Economics*, Volume 4a, edited by Ashenfelter and Card, Elsevier B.V.
- Hicks, Stephen and Thomas, Jenifer (2009). “Presentation of the gender pay gap: ONS position paper”, Office for National Statistics, November.
- Hirano, Kiesuke, Imbens, Guido W., and Ridder, Geert (2003). “Efficient estimation of average treatment effects using the estimated propensity score”, *Econometrica*, 71, 1161–1189.
- Kline, Patrick (2011). “Oaxaca-Blinder as a reweighting estimator”, *American Economic Review*, 101, 532–537.
- Leaker, D. (2008). “The gender pay gap in the UK”, *Economic and Labour Market Review*, 2(4), 19-24.
- Manning, Alan and Swaffield, Joanna (2008), “The gender gap in early-career wage growth”, *Economic Journal*, 118, 983-1024.

- Oaxaca, Ronald (1973). “Male-female wage differentials in urban labor markets”, *International Economic Review*, 14, 693–709.
- Office for National Statistics (2013). Annual Survey of Hours and Earnings, 1997-2012: Secure Access [computer file]. 4th Edition. Colchester, Essex: UK Data Archive [distributor], June 2013. SN: 6689, <http://dx.doi.org/10.5255/UKDA-SN-6689-3>.
- Shackleton, J.R. (2008), *Should We Mind the Gap? Gender Pay Differentials and Public Policy*, Institute of Economic Affairs.
- Stewart, Mark B. (2011). “The changing picture of earnings inequality in Britain and the role of regional and sectoral differences”, *National Institute Economic Review*, 218, 20-32.

**Table 1**

**Median and mean gender pay gaps by region for selected years**  
 Hourly earnings (excluding overtime) for full-time employees

%

	Median gender pay gaps				Mean gender pay gaps			
	1975	1988	2000	2012	1975	1988	2000	2012
<b>North East</b>	31.8	25.6	19.7	8.5	28.5	25.4	16.5	9.4
<b>North West</b>	29.3	22.9	18.8	8.7	28.8	24.6	21.6	11.6
<b>Yorkshire &amp; Humber</b>	28.9	23.4	17.4	9.1	26.9	23.3	17.4	11.7
<b>East Midlands</b>	30.3	26.0	17.8	11.8	30.8	26.7	19.4	15.3
<b>West Midlands</b>	31.4	25.8	17.4	10.8	28.7	25.3	20.4	13.8
<b>East of England</b>	26.0	24.2	16.7	12.4	23.6	26.9	20.4	15.6
<b>London</b>	21.1	18.6	16.0	12.0	27.0	27.7	24.8	20.0
<b>South East</b>	26.4	23.7	17.1	15.0	28.7	26.7	22.5	19.7
<b>South West</b>	24.6	23.1	18.8	10.5	24.6	24.1	20.4	13.6
<b>Wales</b>	30.0	21.9	15.9	9.5	28.6	20.9	14.4	7.9
<b>Scotland</b>	28.7	22.9	18.0	8.4	28.6	26.6	19.1	13.9
<b>Great Britain</b>	27.8	22.5	16.3	9.8	27.3	25.1	20.4	15.1

Notes:

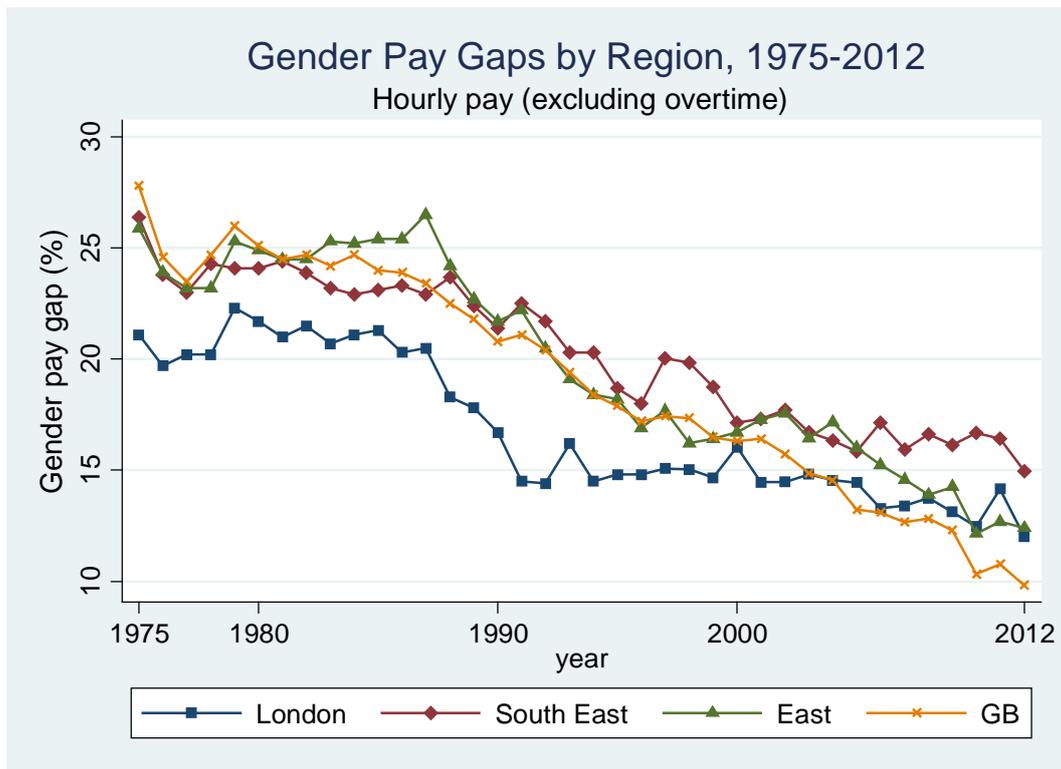
Sources: ASHE published tables and NES microdata.

**Table 2****Actual and counter-factual gender pay gaps, 2012**

	<b>Median</b>	<b>Geometric mean</b>	<b>70th percentile</b>	<b>Arithmetic mean</b>	<b>90th percentile</b>
<b>“Rest of GB”</b>	9.2	10.0	10.1	12.8	13.9
<b><u>London:</u></b>					
Actual	12.0	14.5	17.4	20.8	29.5
Counterfactual	6.3 (0.9)	10.1 (0.6)	11.9 (0.9)	15.5 (0.8)	22.0 (1.3)
<b><u>South-eastern corner</u></b>					
Actual	11.1	14.0	16.4	18.9	25.9
Counterfactual	7.2 (0.7)	10.5 (0.4)	11.5 (0.6)	15.0 (0.5)	20.3 (0.9)
<b><u>Inner London</u></b>					
Actual	16.5	17.0	21.1	23.5	32.7
Counterfactual	5.6 (1.3)	9.2 (0.9)	10.6 (1.2)	15.3 (1.0)	21.0 (1.7)

Notes:

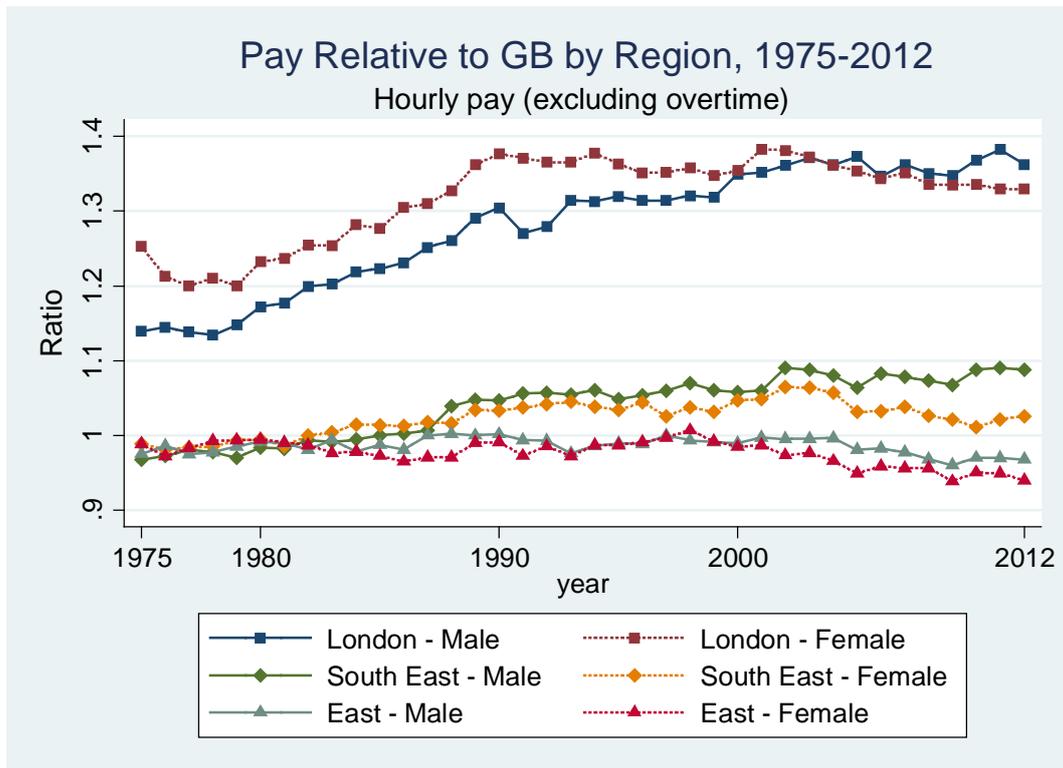
1. Source: Author’s calculations from Ashe 2012 microdata.
2. Control variables = occupation (2-digit), age, company size (log), collective agreement, public sector.
3. Bootstrap standard errors (based on 1000 replications) in parentheses below counterfactual estimates.



**Figure 1: Gender Pay Gaps by Region, 1975-2012**

Source: ASHE published tables and NES microdata.

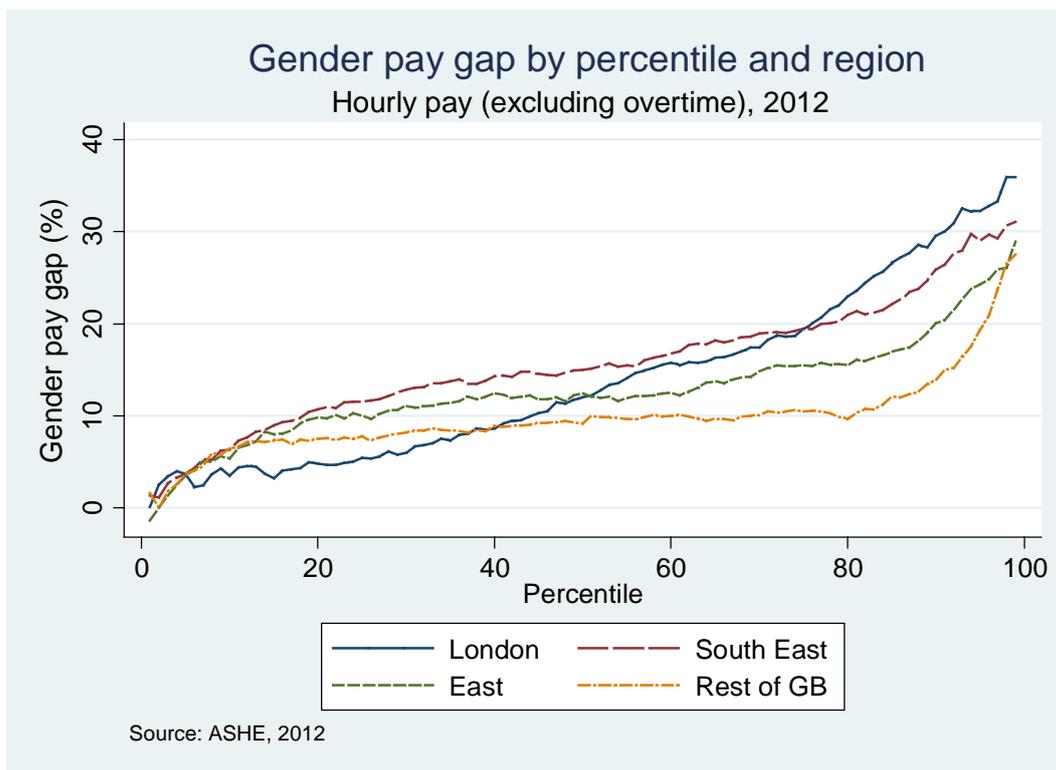
Note: See Appendix for series discontinuities due to methodological changes.



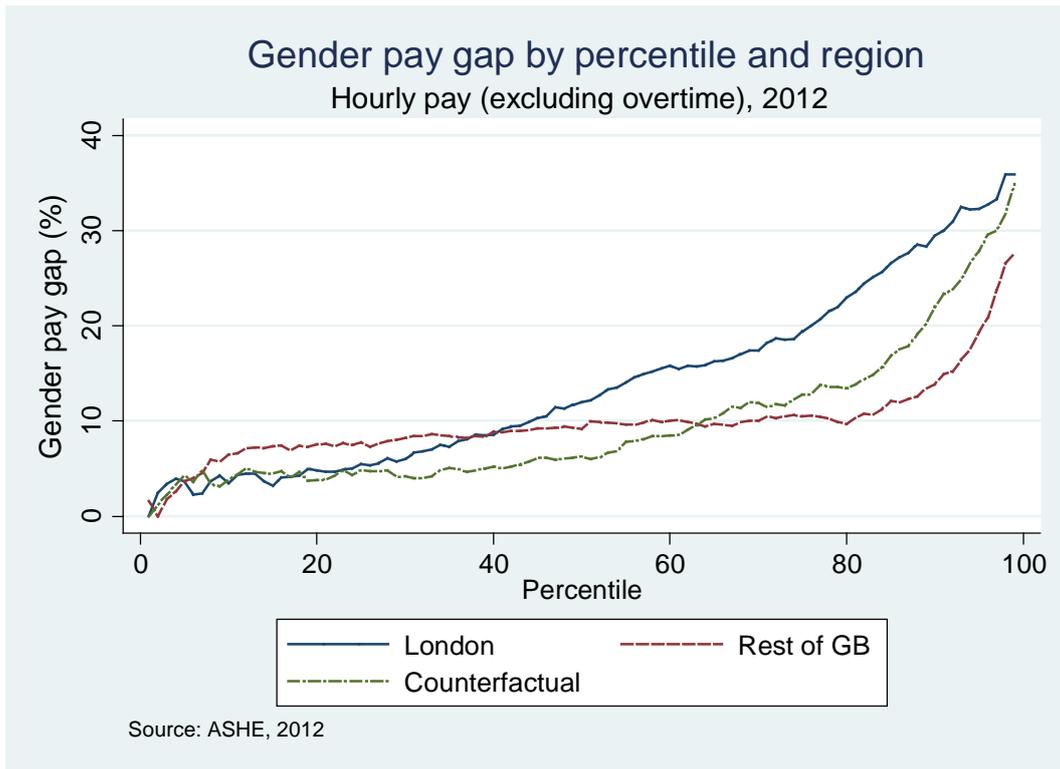
**Figure 2: Male and Female Pay Relative to GB by Region, 1975-2012**

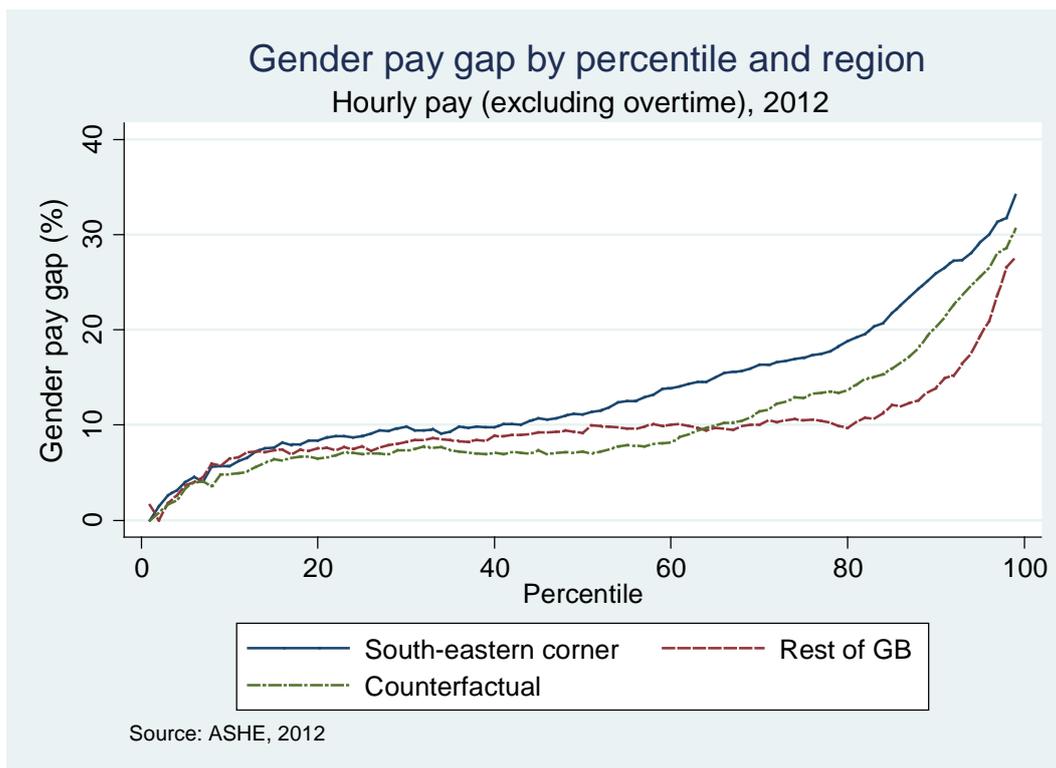
Source: ASHE published tables and NES microdata.

Note: See Appendix for series discontinuities due to methodological changes.

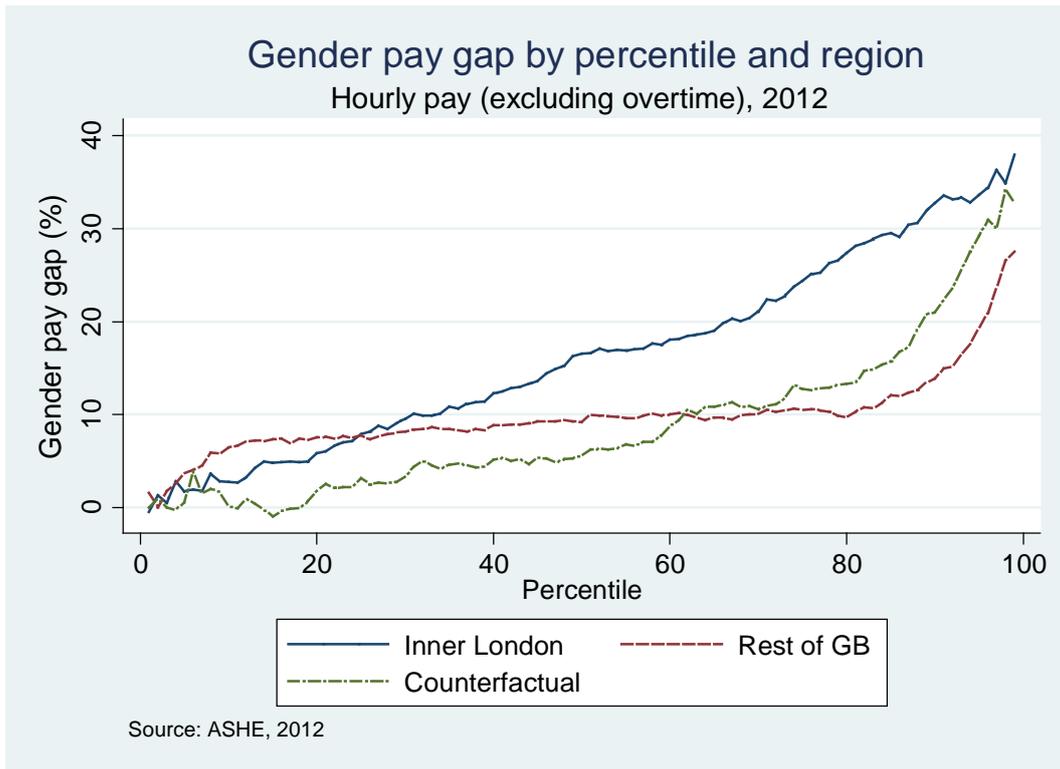


**Figure 3**

**Figure 4**



**Figure 5**



**Figure 6**

## Appendix: Discontinuities in the NES/ASHE data

Despite the accuracy of the NES/ASHE data, care has to be taken when comparing wage differentials or distributions at different points in time using them due to changes in survey methodology and coverage that have taken place. There are five main discontinuities to the series. The ASHE was introduced in 2004 and ONS have generated a semi-comparable back series for 1997-2003 by weighting and imputation of the NES data. The ASHE contains important improvements over the NES and there is an important discontinuity between the pre- and post-1997 figures as a result. There have also been changes to the ASHE methodology that cause discontinuities to the series in 2004, 2006 and 2011 and an earlier potentially important definitional change to the published NES figures in 1983.

For each of these main discontinuities separate figures are available that provide backward and forward compatibility and these also enable an evaluation to be made of the importance of the discontinuity for a particular statistic.<sup>8</sup> The table below gives figures for each of the discontinuity dates with and without the methodological and/or coverage changes made at that date. These are given for male and female median hourly earnings and the median gender pay gap. Figures are for gross hourly earnings excluding overtime (except 1983) of full-time adult employees whose pay in the reference week was not affected by absence. The 1983 figures are for gross hourly earnings including overtime pay and overtime hours.

Year	Without change			With change			Difference		
	Median wage (£)		Gender gap (%)	Median wage (£)		Gender gap (%)	Median wage (£)		Gender gap (%)
	Men	Women		Men	Women		Men	Women	
1983	3.56	2.60	27.0	3.50	2.62	25.1	-0.06	0.02	-1.82
1997	8.13	6.81	16.2	8.40	6.94	17.4	0.27	0.13	1.14
2004	11.09	9.53	14.1	10.96	9.37	14.5	-0.13	-0.16	0.44
2006	11.71	10.23	12.6	11.64	10.14	12.9	-0.07	-0.09	0.25
2011	13.11	11.91	9.2	13.11	11.74	10.5	0.00	-0.17	1.30

### 1983:

The 1983 break results from a change in the definition of an “adult”. The restriction to those paid on adult rates that is used with the ASHE was also used by the NES from 1983 onwards. However prior to 1983 the information on whether individuals were on adult rates was not collected. The equivalent published NES figures were instead for men aged 21 and over and women aged 18 and over. This change resulted in a 6p fall in the median wage for men and

<sup>8</sup> See Stewart (2011) for an examination of their impact on the measurement of wage inequality.

a 2p rise in that for women, giving roughly a 2 percentage point fall in the median gender pay gap.

1997:

The main differences between the NES and the ASHE (introduced in 2004) are: (i) that ASHE weights responses to the number of jobs in the Labour Force Survey to address unit non-response, (ii) that ASHE imputes for item non-response, and (iii) that the coverage of employees by ASHE is greater than that by NES. The back series for 1997-2003 was created by ONS by the weighting and imputation of NES data, but does not include the supplementary coverage. The 1997 break is therefore due to the introduction of weighting and imputation. This resulted in an increase in the median gender pay gap of a little over one percentage point.

2004:

A version of the 2004 data excluding the supplementary coverage was also created by ONS to be comparable with the back series generated by weighting and imputation of the 1997-2003 NES data. The 2004 break therefore results from the improved ASHE coverage of employees, particularly those changing or starting jobs between January/February and April. For 2004 – 2006 consistent ASHE data is available including the supplementary sample, in particular those changing or starting jobs between the drawing of the sample from the HMRC database in January / February and the ASHE reference period in April. This extra coverage from 2004 onwards resulted in only a small change in the median gender pay gap of less than one-half of a percentage point.

2006:

Data for 2006 – 2011 are available on a consistent basis using what is known as the “2007 methodology”. For the purposes of the current paper the main methodological change in 2006 is that employees of companies with a “special arrangement” (SA) with ONS are treated as a separate stratum for constructing weights. These SA companies provide their data electronically and have “internal systems set up to extract and return information on all relevant employees at the survey reference date”. As a result, as well as higher accuracy, they

have higher response rates and consequently need lower weights.<sup>9</sup> However this change has only a very minor impact on the gender pay gap.

2011:

The 2011 discontinuity is again due to a change in the construction of the weights used, this time due to a change in the occupational classification system. In 2011 there was a switch in the occupational coding system from the SOC2000 to the SOC2010 classification and this affected the constructed weights, since occupation is one of the determinants of these. This caused a fall in the median hourly earnings of women but no change for men, which resulted in an increase in the median gender pay gap of a little over one percentage point.

---

<sup>9</sup> An additional change for 2007 and 2008 was that the sample size was reduced by 20%. This was reversed and the full sample restored in 2009. Standard errors of the various statistics considered will be larger for 2007 and 2008 as a result.