NATSEM
UNIVERSITY OF CANBERRA

APPENDICES A-E: THE IMPACT OF A SUSTAINED GENDER WAGE GAP ON THE ECONOMY

Report to the Office for Women, Department of Families, Community Services, Housing and Indigenous Affairs

PREPARED BY

Rebecca Cassells, Yogi Vidyattama, Riyana Miranti, and Justine McNamara

NOVEMBER 2009

Appendix A

Decomposition Approaches

Decomposing the wage gap adjusts for the overall level of differences between men and women in the labour market. (Walby and Olsen, 2002, p.104) Theoretically, it enables the wage gap to be compartmentalised into that portion which can be attributed to differences in characteristics of individuals, and a portion that is unexplained by these differences and could indicate discrimination[1]. Identification of components that are affecting the wage gap considerably can direct policy appropriately.

As noted in the main body of the report, there are several standard methodologies that have been developed in order to decompose the underlying components of wage gaps, each with benefits and limitations. As with any empirical analysis, the estimation technique used depends upon assumptions made about particular relationships. In this case, these assumptions are about how labour markets within Australia operate and how individual and organisational characteristics affect wages.

In this appendix, we discuss three widely used decomposition techniques – the Oaxaca-Blinder, Juhn-Murphy-Pierce and similar extensions of the Oaxaca-Blinder technique. We also discuss the relatively new simulation approach used by Walby and Olsen (2002, 2004) that is used in this report to gauge the underlying components of the gender wage gap.

1) Oaxaca-Blinder technique

Ronald Oaxaca and Alan Blinder produced two very similar estimation techniques to decompose the gender wage gap, almost simultaneously in 1973 (Oaxaca 1973; Blinder 1973). These two autonomous studies remain as the foundation of wage gap studies today, and they represent one of the most common decomposition techniques used in studies of wage inequality. The earnings regressions used for these decompositions are based on earlier work by Mincer (1974). Mincer developed a model of earnings differentials based on endowments of human capital, where individual earnings principally depended on endowments of formal education and work experience.

The purpose of Oaxaca's well-known 1973 study of the U.S. labour market was to separate or 'decompose' the differences in men and women's wages into those parts that can be explained by differences in human capital and those parts that are due to discrimination. There are several key assumptions that underlie Oaxaca's decomposition; the most notable

[1] Throughout the literature, a number of terms are used to describe differences in the characteristics of individuals and differences in the rewards for these characteristics. Individual characteristics are often referred to as 'endowments', 'levels', 'characteristics', 'human capital' or 'productivity differentials'. Rewards for these characteristics are also referred to as 'prices', 'remuneration', 'wage structure',' non-discriminatory wage structure', 'coefficients' or 'discrimination'.

is the assumption that individual characteristics are a measure of marginal productivity, and marginal productivity equates to wages (Oaxaca 1973, p.695).

Under the Oaxaca-Blinder methodology, earnings are estimated for separately for each gender using the following equation:

$$\ln W_i = \alpha + X_i \beta + \varepsilon_i \qquad \text{(A1)}$$

where W_i is the natural log of the wage for individual i; α is an intercept term or constant; X_i is a vector of regressors capturing the individual characteristics expected to impact on wages, such as education and experience; and ε_i is a residual term.

Following Oaxaca (1973), the gender wage differential can then be decomposed into that part attributable to differences in human capital endowments (the difference in the mean values or Xs) and that part attributable to differences in rewards for those endowments (the βs):

$$\bar{Y}_m - \bar{Y}_f = (\bar{X}_m - \bar{X}_f)\hat{\beta}_m + \bar{X}_f(\hat{\beta}_m - \hat{\beta}_f) \qquad \text{(A2)}$$

The first term on the right hand side of the equation measures the differences in the mean characteristics or endowments between men and women, and here is weighted by the estimated price for male endowments. These differences are often thought to be legitimate or 'explained' sources of the wage gap, however as expanded on below, these differences could also reflect pre-labour market choices or discrimination. The second part of the equation is the 'unexplained' or unjustifiable component of the wage gap, and is weighted by the mean of X. If, on average, the price of male labour exceeds the price of female labour, then there exists a male advantage.

The allocation of the gap into 'explained', and 'unexplained' components will depend on the choice of coefficients (or weights) used for the decomposition (βm or βf, where m and f indicate male and female respectively). Oaxaca assumes that rewards to endowments are interchangeable, that is, without discrimination, men could be paid as women are and women could be paid as men are, however he implies that the true wage structure lies somewhere between the female and male wage structures, and applies both in his calculations (Oaxaca 1973). The wage structure is defined as a set of prices for various labour market skills for both measured and unmeasured characteristics and the rents received for employment in particular sectors of the economy (Blau and Kahn 1997). The majority of applications of the Oaxaca-Blinder decomposition technique use the male coefficients as the weighting mechanism, with the assumption that the male wage structure is the current non-discriminatory price. Arguments for maintaining the male wage structure as the non-discriminatory price of labour are ones with tones of legality, where it is reasoned that in legal proceedings related to gender discrimination, it is the male wage structure that would be used as the non-discriminatory wage when determining compensation (Oaxaca and Ransom 1994; Ginther and Hayes 2003; Healy *et al.* 2009).

The use of either the female or male weight in determining the amount of the wage gap that is attributed to differences in endowments is one of the major criticisms of the Oaxaca-Blinder technique and there have been several extensions of this work in order to achieve a more appropriate non-discriminatory wage structure. These extensions are reviewed further below.

One of the benefits of the Oaxaca-Blinder technique is that it is relatively easy to understand, has its foundations based in a solid neo-classical human capital framework, and enables identification and quantification of the components of the wage gap that can be explained by differences in endowments and those that relate to differences in prices for these endowments. However, one of the key disadvantages of the Oaxaca-Blinder technique is that the unexplained component of the gender wage gap cannot be separated out into factors related to direct discrimination (that is, being a woman), and factors that are related to other attributes such as motivation, drive and initiative, that are difficult to quantify and include. Estimating separate wage equations for men and women is also a shortcoming, in that a direct comparison between men and women cannot be made. Also, as the Oaxaca-Blinder technique is based upon the neo-classical assumption that wages are equal to productivity, if all characteristics related to productivity (e.g. education, experience, motivation, initiative) cannot be included in the model due to data limitations, then these 'left-over' elements of productivity are included in the 'unexplained' component of the wage gap, along with a discrimination component, which may lead to an over-estimation of the unexplained component. Over the years, more and more explanatory variables have been included in wage gap analysis in order to try and overcome this shortcoming, and in the event that particular important variables, such as work experience, were not available, proxy indicators would be substituted in order to achieve a more fully specified model. Further, the model can also be limited by measurement error within the variables chosen to represent productivity. For example, not having the level of data needed to ascertain a true endowment effect may mean that the results are inaccurate.

Another critical assumptions that is inherent in the Oaxaca-Blinder methodology, and one that again stems from neo-classical economic theory, is the assumption that there exists individual choice (or freedoms), associated with gaining human capital factors, such as education, work experience and on the job training (Grimshaw *et al.* 2002 p. 4; Olsen and Walby 2002). Oaxaca notes the limitation of his model in that it 'does not take into account the effects of the feedback from labour market discrimination on the male-female differences in the selected individual characteristics', (Oaxaca 1973, p.708). The differences in human capital levels could be due to employer discriminatory practices, or the individual choice to invest less in human capital due to lower returns, which is itself influenced by current discrimination (Grimshaw *et al.* 2002, p.27). Grimshaw argues that these feedback effects result in an under-estimation of labour market discrimination (Grimshaw *et al*. 2002, p.27).

In addition, there is also an argument that because the wage gap is estimated around the mean value of the total population this is not giving a true estimation, and the Oaxaca-Blinder methodology has been extended to quantile wage regression analysis, where the wage distribution is separated into income groups. This has also been used in order to examine the 'glass ceiling' and 'sticky floor' effects that may exist within the labour market (see, for example, Kee 2006).

Overall, the Oaxaca-Blinder decomposition technique has four major disadvantages:

- Unsatisfactory choice of a true non-discriminatory wage structure.
- Feedback effects will mean that discrimination is under-estimated.
- Women and men cannot be compared directly due to separate wage estimations.
- Discrimination may be blurred with omitted variable bias, making it impossible to truly separate discrimination from other factors.

The advantages of the Oaxaca-Blinder methodology are that:

- Separate effects of endowments and prices can be calculated and quantified.
- Separate coefficients for men and women can be measured for each endowment.

EXTENSIONS OF OAXACA-BLINDER: CHOICE OF WEIGHTING MATRIX

There have been several variations of the Oaxaca-Blinder technique that seek to overcome the problem of gaining an appropriate non-discriminatory wage structure or weighting matrix (see for example Oaxaca and Ransom 1988, 1994; Neumark 1988; Cotton 1988; Reimers 1983). Each of these involves calculating a weighting matrix that is more reflective of the non-discriminatory price for wages.

Oaxaca and Ransom (1994) extend the original equation (A2), by introducing a third term which includes the estimation of the non-discriminatory wage structure for a pooled labour market. The Oaxaca-Ransom equation is as follows:

$$\ln \overline{W}_m - \ln \overline{W}_f = \overline{X}_m(\beta_m - \beta^*) + (\beta^* - \beta_f)\overline{X}_f + (\overline{X}_m - \overline{X}_f)\beta^* \qquad \text{(A3)}$$

Where β* is the estimated non-discriminatory wage structure, and the other notations are the same as the values given above in equation (A2). The first term on the right hand side of the equation estimates male wage advantage, the second term is female wage disadvantage, and the third term is an estimate of the productivity differential (Oaxaca and Ransom 1994, p.8). One of the underlying concepts of this equation is that, in wage discrimination, there not only exists a group that is disadvantaged (receiving below the non-discriminatory price), but there also exists a group that is advantaged (receiving above the non-discriminatory price). In the Oaxaca-Blinder technique, there is only ever one or the other, consequently, the result will either be under or over estimated.

The main difference between the Oaxaca-Ransom, Neumark, Cotton and Reimers approaches is in the way that they derive their weighting matrix. Oaxaca-Ransom and Neumark have both used the coefficients derived from the entire population in order to estimate their matrix, Cotton uses labour force weights (the share of the majority group), and Reimers chooses a weighting mechanism that lies directly between the two groups (see Silber and Weber (1999), Oaxaca and Ransom (1994) and Beblo *et al.* (2003) for comparisons).

Each of these weighting matrixes has its benefits and limitations. Cotton's approach gives a greater weight to the dominant labour group in the working population, with the assumption that the 'nondiscriminatory wage structure will be closer to the current white wage structure than to the current black wage structure' (Cotton 1988, p.239). Cotton's estimation of the non-discriminatory wage rate is given by:

$$\beta^* = P_m \beta^m + P_f \beta^f \qquad \text{(A4)}$$

Where P_m and P_f are the proportions of males and females (males and females have been substituted for white and blacks in this equation for relevance to the issue at hand).

Reimers estimates the non-discriminatory wage structure more simplistically, using a division of the two estimated male and female wage structures and is denoted by:

$$\beta^* = (\beta_m + \beta_f)/2 \qquad \text{(A5)}$$

Oaxaca and Ransom and Neumark use the pooled labour market in order to ascertain the non-discriminatory wage structure, and is shown by:

$$\beta^* = \Omega_\circ \beta_m + (I - \Omega_\circ)\beta_f \qquad \text{(A6)}$$

Where Ω = is a weighting matrix and I is a diagonal unit matrix (identity matrix as explained by Oaxaca and Ransom 1994, p. 12).

All three approaches have similarities; in fact Oaxaca and Ransom argue that Cotton's weighting technique can give the same result as the one they have proposed, and Reimer's is a special case of Cotton's methodology (Oaxaca and Ransom 1994). One of the disadvantages of both the Reimers and Cotton approaches is that the non-discriminatory wage is bounded or 'bracketed' by the separate male and female estimations, whereas the Oaxaca-Ransom methodology is not (Oaxaca and Ransom 1994, p.18). In the Oaxaca and Ransom (1994) comparison of the four alternative weighting matrices, they conclude that using a single gender wage structure is "too extreme", and that the pooled methodology gave the lowest standard errors. (Oaxaca and Ransom 1994, p.18). However, they do note that despite the empirical superiority of the decomposition technique, in a real world situation, it would be the male wage that would be taken as the non-discriminatory wage in wage discrimination litigation proceedings. However, within the Australian context, legislative changes in Queensland and New South Wales have moved away from using the male wage rate as the benchmark in industrial disputes related to wages, focusing more on the value of the work.

Silber and Weber (1999) compare five decomposition procedures (Oaxaca-Blinder, Neumark, Reimers, Cotton and Oaxaca-Ransom) using bootstrap techniques and find no significant differences between each technique. They do note, however, that the Neumark technique places a higher weight to the endowment effect and a lower weight to the remuneration effect when compared with the other four approaches (Silber and Weber 1999, p.365).

Overall, regardless of the weighting matrix used, combining the wage structures of men and women in an effort to achieve a true non-discriminatory wage structure offers a sensible middle ground in overcoming the shortcomings of using a single sex wage structure.

JUHN-MURPHY-PIERCE TECHNIQUE

An alternative technique to the Oaxaca-Blinder (1973) technique is the Juhn-Murphy-Pierce (1991) decomposition (for simplicity we refer to this as JMP), which is often used for comparisons over time or between countries. The main contribution of this technique is that it includes the whole wage distribution, rather than just the mean, in order to allow for changes in the overall gender wage gap. Juhn *et al.* (1991) analyse trends in the slowdown in black and white wage convergence to examine the effects of wage inequality on the black and white wage gap. The core advantage of the JMP decomposition over Oaxaca-Blinder (1973) is that it allows the decomposition of changes in the residual of the gender wage gap into “price” and “quantity or gap” effects (this is explained further below). We will start from the Oaxaca-Blinder (1973) decomposition as discussed in Yun (2009). Suppose we compare the wages of men and women and decompose their (mean) wage differentials, using this standard decomposition equation (Yun 2009, p.16):

$$\bar{Y}_m - \bar{Y}_f = \left(\bar{X}_m - \bar{X}_f\right)\beta_m + \bar{X}_f\left(\beta_m - \beta_f\right) + \left(\bar{\varepsilon}_m - \bar{\varepsilon}_f\right), \qquad \text{(A7)}$$

where m and f refer to males and females respectively, Y is the log of wages, X is the vector of explanatory variables, β are vectors of coefficients and ε is the residual. The ’over bar’ is the value of the sample average.

Similar to equation (A2), the first term on the right hand side of equation (A7) measures the “characteristic effect”, the second term measures the coefficient effect, and the third term represents the residual effect. The coefficient effect is often interpreted as the discrimination effect, whilst the residual effect measures differences in the unobserved characteristics distribution. When equation (7) is estimated using the OLS (Ordinary Least Squares) technique, the residual term collapses to 0 since both ε_m *and* ε_f are equal to 0 (Yun 2009). Juhn *et al.* (1991) argue that including the residual effect is important as it represents unobserved characteristics such as motivation and drive. However, if OLS estimation is used, these unobserved characteristics cannot be estimated. Therefore, Juhn *et al.* (1991) propose a decomposition equation which enables the examination of the effect of unobserved skills to be measured when using OLS to estimate wages. Following Blau and

Kahn (1997 pp. 6-7), who decompose the gender wage gap using the JMP methodology, the male wage equation for worker m in year t can be expressed as:

$$Y_{mt} = X_{mt}\beta_{mt} + \sigma_{mt}\theta_{mt} \qquad \text{(A8)}$$

where Y_{mt} is the natural log of wages for males, X_{mt} is the vector of explanatory variables for males, β_{mt} are vectors of coefficients for males, θ_{mt} is the standardised residual that captures unobserved skills of males (Kunze 2008), with a mean of zero and variance of one for each year, while σ_{mt} is the standard deviations of the residual for male wages for that year (to measure the inequality amongst the residual of the male wage).

Blau and Kahn (1997) assume that returns to observed characteristics are similar for males (m) and females (f) and thus the non-discriminatory wage structure is given by $\beta_m = \beta_f$.

So, under the assumption that prices from the male sample wage regression β_{mt} are equivalent to competitive prices and therefore discrimination to observed characteristics is ignored (Kunze 2008). Given this, the gender wage gap for year t (D_t) is written as follows:

$$D_t \equiv \overline{Y}_{mt} - \overline{Y}_{ft} = \Delta\overline{X}_t\beta_{mt} + \sigma_{mt}\Delta\overline{\theta}_t \qquad \text{(A9)}$$

Equation (A9) shows that the gender wage gap can be decomposed into a component due to gender differences in observable characteristics $(\Delta\overline{X}_t)$ weighted by the male coefficients (β_{mt}), and a component due to gender differences in the standardised residual of the male equation (unobserved variables such as motivation) $(\Delta\theta_t)$ multiplied by the money value per unit difference in the standardised residuals (σ_{mt}). The second term on the right hand side of equation A9 ($\sigma_{mt}\Delta\overline{\theta}_t$), is identical to the unexplained or discrimination effect in the standard Oaxaca-Blinder (1973) decomposition.

Blau and Kahn (1997) use the JMP decomposition to examine changes in the gender wage gap in the USA over time. Therefore, the difference in the gender wage gap between two years (between years t and t+1) can be decomposed using equation (A10) below. For simplicity, we drop the prefix m, therefore in the equations below, β_t refers to β_m, whilst σ_t refers to σ_m.

$$D_{t+1} - D_t \equiv (\Delta X_{t+1} - \Delta X_t)\beta_{t+1} + \Delta X_t(\beta_{t+1} - \beta_t) + (\Delta\theta_{t+1} - \Delta\theta_t)\sigma_{t+1} + \Delta\theta_t(\sigma_{t+1} - \sigma_t)$$
(A10)

The discussion about each component of the JMP decomposition below closely follows Blau and Kahn (1997, p.7).

The first term $(\Delta X_{t+1} - \Delta X_t)\beta_{t+1}$, is the endowment effect which measures the impact of the change in differences in endowments such as the level of educational attainment and years of work experience.

The second term $\Delta X_t(\beta_{t+1} - \beta_t)$ measures the effect of changing prices or rewards for the observed labour market characteristics of males. For example, if women have lower mean work experience levels in comparison to men, and there is an increasing return to male work experience, this will put some weight on women's lower experience levels and raise the gender wage gap.

The third term $(\Delta\theta_{t+1} - \Delta\theta_t)\sigma_{t+1}$ measures the gap/quantity effect or changes in the levels of the unobservable characteristics (Kunze 2008, p.71). This will reflect the contribution to the change in the gender wage gap, if the percentile rankings of the residuals of the female estimated wage have changed, while the distribution/inequality amongst the residuals of the male wage remain constant.

Finally, the fourth term $\Delta\theta_t(\sigma_{t+1} - \sigma_t)$ measures the unobserved price effect or changes in the returns to unobservable characteristics (Kunze 2008, p.71) and measures the impact of a change in inequality on the change in the male-female wage differential, assuming that women maintain the same position in the male residual wage distribution. This will measure the unobserved deficits in relative characteristics or discrimination that lower women's position in the male residual wage distribution. For example, Kidd and Shannon (2001) seek to calculate the unobserved price effect of the gender wage gap in Australia for two different periods of time: 1981-82 and 1989-90. They hold the percentile position of the mean female residual fixed at its 1981-82 position and the unobserved price effect is calculated as the difference in the value of the male residual at this percentile between 1981-82 and 1989-90.

Grimshaw and Rubery (2002, pp. 10) note that 'the gender-specific factors are reflected in the first and the third terms', which are the effect of gender differences in endowments and gender differences in wage rankings at a given observed characteristic. In contrast, the effects of a country's labour market structure are reflected in the second and fourth terms, which measure the effect of changes in returns to observed and unobserved characteristics (Blau and Kahn 1997).

Overall, in comparison to the Oaxaca-Blinder decomposition, the JMP decomposition has several advantages, however, there are also drawbacks related to the suitability of JMP as it relies on several strong assumptions which are considered difficult to verify (Yun 2009).

The advantages of the JMP decomposition methodology are as follows:

- Enables accurate estimates of the wage gap over-time or between countries.

- Includes and allows for the decomposition of changes in the residual into price and quantity effects, thus allowing for the relative importance of gender-specific-factors and the wage structure in the narrowing/widening of the gender wage gap (Blau and Kahn 1997).
- Minimises the problem of sample selection bias, which is strongly present in the Oaxaca-Blinder technique, by avoiding the need to make separate estimates of wage equations for women (Juhn *et al.* 1991). Sample selection bias usually occurs because female labour force participation changes over time and women who participate (or do not participate) in the labour market have specific characteristics. For example, women with lower levels of human capital often choose not to participate in the labour market (Grimshaw and Rubery 2002). In this case, women who participate in the labour market may not be representative of the female population (Grimshaw and Rubery 2002).

The drawbacks of the JMP decomposition are as follows:

- The assumption that there are changes in the distribution of male wage residuals may not necessarily be always the case. Changes in the distribution of male wage residuals also capture changes in sample composition and measurement error (Blau and Kahn 1997).
- The use of the male equation as the benchmark implies that the set of prices that influence women are similar to those that influence men, which then implicitly leads to the assumption that there are similar factors that raise wage inequality of both men and women (Blau and Kahn 1997).

In addition, the use of the JMP decomposition is not likely to be the best choice if the aim is to examine the impact of discrimination on the gender wage gap. Although, within a traditional framework of wage decomposition, the sum of the third and fourth terms represents changes in the unexplained differential which are commonly used as an estimate of discrimination Blau and Kahn (1997) acknowledge the complexity of the interpretation if the focus is on discrimination. In the Oaxaca-Blinder methodology discrimination is embedded in only one component of the equation, whilst in the JMP decomposition, discrimination is divided between the third and fourth components. Juhn *et al.* (1991) note that changes in the residual gap are a combination of changes in the distribution of unobserved characteristics and changes in discrimination. Although the second term of the Oaxaca-Blinder decomposition is also not perfect in explaining discrimination, Yun (2009) claims that researchers may need to choose between the use of the Oaxaca-Blinder decomposition and JMP decomposition depending on the needs of the researchers and whether the focus of the analysis is on discrimination or the price of unobserved skills. Yun (2009) suggests that the Oaxaca-Blinder decomposition is better to use for capturing discrimination in comparison to the JMP decomposition.

Further, the JMP decomposition would be more relevant if researchers are attempting to examine changes in the gender wage gap over time (for example, Juhn *et al.* 1991 and 1993; Blau and Kahn 1997; Kidd and Shannon 2001) or across countries (for example, Rice 1999;

Blau and Kahn 2001). Kidd and Shannon (2001) examine the gender wage gap in Australia during the 1980s and find that around 44 to 80 per cent of the gender wage gap (depending on the specification) is explained by observed characteristics (human capital endowment of experience and education). The gap/quantity effect shows the improvement of women's position within the male wage residual distribution which is related to a combination of women's unobservable characteristics or a decline in the level of discrimination due to policy changes. In contrast, Kidd and Shannon found that the unobserved price effect had little effect on the gender wage gap over this period.

SIMULATION TECHNIQUE OR 'GROSS DECOMPOSITION' APPROACH

Olsen and Walby (Olsen and Walby 2004; Walby and Olsen 2002) propose a different approach in order to gain a more indicative measure of the underlying components affecting the gender wage gap, which involves 'simulating the hypothetical changes needed to bring women's levels of these components into line with those of men' (Walby 2004, p.24). Walby and Olsen (2002, 2004) refer to their approach as a 'gross decomposition', as the slope coefficients are derived from an overall regression (both men and women are included), and label the Oaxaca-Blinder methodology as a 'net decomposition' as it uses 'net components', or male and female slope coefficients estimated separately (Walby and Olsen 2002, p.104).

This technique is relatively new and not widely used in wage gap decompositions. Recently, Watson (2009) has applied the Olsen-Walby technique in his decomposition of the gender wage gap for full-time managers in Australia.

From Walby and Olsen (2002), the decomposition is calculated using the formula:

$$\beta * (\Delta Xi)$$

Where, ΔXi is the hypothetical change in female mean values required to make them equal with the male mean values (or the difference in the means), and β^* is the slope coefficient estimated for men and women jointly[2]. In a hypothetical example, if the mean years of formal education for women is eight years and for men it is ten years, an increase of two years is required in order to bring women's years of education in line with men's. This extra two years of education is then multiplied by the corresponding coefficient (reward) for every extra year of education, which is, say 0.07 (seven per cent). This gives a simulated effect of 0.14 = 0.07*2. This means, that if women had the equivalent average level of formal education, their wage rate would increase by 0.14 (14 per cent). This formula is carried out for every component or variable that is used to predict wages, and the simulated effects are

2 Walby and Olsen (2002) argue that using the slope coefficient for the whole population is effectively using a weighted average somewhere between βm and βf, and therefore superior to choosing either men's or women's separately. (Walby and Olsen 2002, p.106)

then summed to find the overall wage gap, or, the proportion that would shift women's wages to be equal to men's wages. One of the key differences between the Olsen-Walby technique and the Oaxaca versions and the JMP approach is that it is one of an 'integration' rather than 'separation'. The simulation approach does not calculate separate 'reward' and 'endowment' components, but instead focuses on 'hypothetically moving the market in ways that equalise men's and women's experiences' (Olsen and Walby 2004, p. 69). Further, the gender component is illuminated with the Olsen-Walby approach, and is not lost, or spread across factors as in the other methodologies (Walby and Olsen 2002, p.104). This allows a direct discrimination component to be measured or proxied.

Olsen and Walby also choose to exclude from their decomposition the effects of factors which are 'female-advantaging' (those that help to decrease the wage gap), factors that do not change, and those that they consider to be controls only and not relevant to gender wage gaps, however relevant in estimating wages – for example, geographical variables. (Olsen and Walby 2004, p.63). Whilst removal of these components could be considered to produce biased estimates, or in fact an over-estimation of the wage gap, their justification for this approach is that the variables considered are ones with policy relevance to the gender wage gap.

In summation, the advantages of the Olsen-Walby simulation technique are as follows:

- The gender component is visible enabling the effect of direct discrimination or other aspects related to being a woman to be measured.
- There is the option to bring all of the 'policy relevant' variables into the forefront, and to treat all other variables as controls or irrelevant.
- Offsetting 'female advantaging' aspects are removed
- The tug-of-war about what component is due to 'rewards' and what is due to 'endowments' is removed.
- Feedback effects (pre-labour market discrimination) are to some extent addressed by giving women the 'best average situation among men' (Olsen and Walby 2004, p.8).

Some of the disadvantages of the Olsen-Walby technique include those discussed above in the Oaxaca-Blinder section, including measurement error associated with chosen variables, and omitted variable bias. Further, the removal of particular factors related to the wage gap if they are deemed 'controls' or policy irrelevant, could be considered as a shortcoming of this approach.

APPENDIX B

MICROECONOMIC MODEL SPECIFICATION, VARIABLE DESCRIPTIONS AND DETAILED DECOMPOSITION RESULTS

MODEL SPECIFICATION

We have chosen to use the Olsen-Walby decomposition technique in order to develop the most appropriate estimations of the key determinants of the gender wage gap. In doing so, we must first measure the effects of factors that are known to best predict wages, and which will include the main drivers of the gender wage gap. In doing this, we apply a standard statistical technique – regression analysis, which allows the impact of each factor that affects wages to be quantified. All variables are included in our regression analysis, however we then go on to decompose the wage gap using only variables that are relevant to wage differentials. Following Olsen and Walby (2004) we have included only those variables which meet the following criteria:

- they have a non-trivial, positive and statistically significant impact on the gender wage gap; and
- they are policy-relevant (although many factors would require substantial further research to inform the development of suitable policy responses).

Selecting these variables provides a practical way of quantifying the effects of the key determinants of the wage gap. It should be noted that this approach means that variables which are female-advantaging are not included in this figure, nor are a range of variables which have very small effects on the wage gap and which do not have clear policy relevance. Consider, for example, our set of hours of work variables. Only one of these variables – working 41 to 49 hours per week, has a positive effect on the wage gap, and this effect is very small. If we decompose the effect of this variable, we find that if as many women as men were to work 41-49 hours per week, two per cent of the gender wage gap would be explained or eliminated. However, this is not an area that is likely to be the subject of policy intervention. While some additional hours of work per week might have a modest effect on the ratio of women's wages to men's, the intensification of work in this way is widely regarded as being associated with a range of other possibly negative outcomes, for individuals and societies, and is unlikely to be the subject of policy intervention.

Data source and sample

We have used the most recent wave of the Household, Income and Labour Dynamics in Australia (HILDA) survey - 2007, in order to gain the best estimates of factors affecting the gender wage gap. The HILDA survey is a longitudinal survey, and has been conducted annually since 2001. The sample is designed in a way to best represent the entire Australian population, and survey weights are calculated in order to facilitate this. In 2007, there were 12,789 persons and 7,063 households in the sample.

For our sample of wage earners, we decided to exclude several groups of people in order to gain the most accurate measure of the wage gap – this is a common practice in gender wage gap studies, although decisions about which groups to include or exclude differ across studies. We excluded people still at school, self-employed people, people aged below 21 years, and those aged 65 years and above. We also excluded people with unusually low or high hourly wages[3]. Our final sample of wage earners is 6,137 people, of whom 3,073 were women and 3,064 men. When this sample is weighted, it represents 7,951,654 people, or 3,764,055 women and 4,187,599 men.

One of the advantages of using the HILDA data set is that it has a wide range of variables appropriate for estimating wages, including important work history variables such as the number of years in paid work, that are not available in other surveys such as the EEH or EEBTUM.

Selection Bias

In many empirical analyses it is important that samples which are used to examine relationships between variables are unbiased, meaning that those included in the sample should be there randomly, rather than due to any other influencing factor. In our situation, using a sample of wage earners only, it is likely that what is called a 'selection effect' is present. This means that sample membership is not random, but rather is influenced by a range of factors, some of which will also be related to our outcome variable (which here is wages). This is particularly the case for women wage earners – previous research has shown that women with lower human capital will be more likely to opt out of the labour force (and thus not be in our sample), thereby introducing a bias in our estimations (Heckman 1979). This situation has been rectified by the use of a selection model, which predicts the likelihood of persons being employed and 'corrects' for this within our estimation, thereby eliminating selectivity bias.

Variable Descriptions

As well as including variables capturing the most likely determinants of the wage gap, our model also includes a range of control variables, in line with those generally used in other Australian research. Additional control variables generally consist of variables that are likely to affect wages, but are not believed to be major drivers of the wage gap. These include marital status (found to have a positive effect on men's wages, but a negative effect on women's (Miller 2005)), and region (Kee 2006; Miller 2004; Miller 2005; Preston 2003; Wooden 1999). In Australian studies, regional control variables used include metropolitan versus non-metropolitan (Preston 2003; Wooden 1999), state and remoteness region (Kee

[3] A small amount of people appeared to be receiving unusual hourly wages of under $5 per hour and over $260 per hour. We have chosen to remove these people from our sample, as these hourly rates are believed to be inaccurate.

2006). We have also added control variables taking into account a person's health status and whether they have children aged four and under and/or children aged between five and 14.

Industry segregation

Industry segregation has been measured following Olsen and Walby (Olsen and Walby, 2004). We have used the Australian and New Zealand Standard Industrial Classification (ANZSIC), 1-digit industry classification (the broadest level) in order to measure the level of industry segregation a person is experiencing (see ABS Cat No. 1292.0 for further details of the classification). This has been measured by calculating the proportion of males within a major industry grouping and multiplying this proportion by ten in order to gain an index of industry segregation. Data from the ABS Labour Force survey, at a similar time period to when HILDA 2007 was conducted (August, 2007), has been used in order to gain more accurate results, as the labour force survey has a larger sample size and specialises in these variables (see ABS Cat No. 6291.0.55.001 for further details). The level of industry segregation an individual was experiencing, measured by the index, was then applied to the base dataset.

Occupational segregation

Occupational segregation has also been measured following Olsen and Walby (Olsen and Walby 2004). We have used the Australian and New Zealand Standard Classification of Occupations (ANZSCO), 1-digit occupation classification (the broadest level) in order to measure the level of occupational segregation a person is experiencing (see ABS Cat No. 122.0 for further details of the classification). This has been measured by calculating the proportion of males within a major occupation grouping and multiplying this proportion by ten in order to gain an index of occupational segregation. Data from the ABS Labour Force survey, at a similar time period to when HILDA 2007 was conducted (August, 2007), has been used in order to gain more accurate results, as the labour force survey has a larger sample size and specialises in these variables (see ABS Cat No. 6291.0.55.001 for further details). The level of occupational segregation an individual was experiencing, measured by the index, was then applied to the base dataset.

Firm size

The number of employees in a firm has been disaggregated into four categories – less than 20 employees, 20-100 employees, 100-500 employees and 500+ employees. We have used those employed in firms with less than 20 employees as the reference group in our regression estimation.

Children

Persons with children aged four or under have been included in our regression analysis as a dichotomous variable, as well as persons with children aged five to 14 years.

Partner

A person is considered to be partnered if they identify as being legally married or in a defacto relationship. This variable is dichotomous.

Educational Qualifications

We have divided the qualification field into three categories, those with a Bachelor or above (including graduate diploma, masters and doctorate), those with vocational training (including those with certificates, diplomas and associate diplomas) and those with year 12 or below. We have chosen those with year 12 or below as the reference group in our regression estimation.

Long term health condition

A person is considered to have a long term health condition if they identify as having a condition that has or is likely to last six months or more, restricts everyday activity and cannot be corrected by medication or medical aids (HILDA, Living in Australia Showcard set, Wave 7). Such conditions include sight problems not corrected by glasses or contact lenses, hearing problems, speech problems, blackouts, fits or loss of consciousness and any other long-term condition such as arthritis, asthma, heart disease, Alzheimer's disease and dementia (please refer to HILDA documentation for a full list).

Hours of work

We have chosen to create categorical variables from the continuous hours of work variable available in HILDA. There is discussion in the literature about the possible effects of variations in hours of work (especially unpaid overtime hours) on observed wages (Preston 2003), and some studies control for this with a set of dummy variables (Preston 2003; Wooden 1999). Through evaluation of the distribution of hours worked, and following standard definitions of part-time and full-time work in Australia, we have chosen four categories of comparison – working 1-34 hours per week, 35-40 hours per week, 41-49 hours per week and 50+ hours per week. We have chosen those that work 35-40 hours per week as the reference group.

Tenure in occupation and with current employer

Tenure in occupation and with current employer are continuous variables that integrate weeks and years into a year unit of measurement.

Union membership

Union membership is a dichotomous variable.

Sector of employment

Sector of employment has been divided into three groups – public, private and other. The other grouping consists of other government business enterprise or commercial statutory authority, other commercial, private sector not for profit organisation, and other non-commercial organisations.

Work schedule

The type of work schedule an employee is working has been divided into 'regular' and 'irregular', with the regular classification including those working a regular daytime schedule, evening shift or night shift. Irregular work schedules include a rotating shift (changes from days to evenings to nights), a split shift (two distinct periods each day) and being on call.

Geographic location

HILDA allocates the geographic residence of persons using the Australian Standard Geographic Classification (ASGC, ABS Cat. No. 1216.0). The classification we have chosen to use is that of 'Section of State', which divides locations into 'major urban', 'other urban', 'bounded locality' and 'rural balance'. We have combined those persons living in bounded localities and the rural balance to form a rural group.

Table B1 Olsen-Walby detailed simulated decomposition of the gender wage gap

	Men	Women	Change factor	Pooled coefficient	Simulation effect	Simulated change as a % of the wage gap	Cents/hour equivalent	$/35 hour week
	Mean	Mean	ΔX	B	BΔX			
Time in paid work (years)	19.13	17.09	2.04	0.0024242	0.004945368	0.0288024574	9.02	3.16
Tenure in occupation (years)	9.91	8.7	1.21	0.0055275	0.006688275	0.0389533713	12.19	4.27
Tenure in current employment (years)	7.25	6.41	0.84	0.000669	0.00056196	0.0032729271	1.02	0.36
Vocational qualification (%)	0.39	0.29	0.1	0.0823698	0.00823698	0.0479732278	15.02	5.26
Industry segregation (male proportion x 10)	6.01	4.39	1.62	0.0266268	0.043135416	0.2512261945	78.63	27.52
Firm size: 100-500 employed	0.21	0.18	0.03	0.1302226	0.003906678	0.0227529937	7.12	2.49
Firm size: 500+ employed	0.15	0.14	0.01	0.1935139	0.001935139	0.0112704977	3.53	1.23
Female	0	1	-1	-0.1022897	0.1022897	0.5957483305	186.47	65.26
Total					**0.171699516**	**1.00**	**313**	**109.55**

Note: Pooled coefficient refers to the estimated return for the entire group of wage earners for each additional unit of a particular variable, or in comparison to the reference group in the model. The reference group for firm size are those employed in a firm with less than 20 employees, and for educational qualifications, it is those with year 12 or below. All other variables are continuous or dichotomous.

Source: *Authors' calculations from HILDA, Wave 7unit record data.*

Table B2 OLS regression for log hourly earnings, 2007

	Coefficient	Std. Err.	t	P>t	[95% Confidence Interval]	
Time in paid work (years)	0.0024242	0.0004856	4.99	0	0.0014723	0.003376
1-34 hours per week (reference: 41-49 hours)	0.0282771	0.0136763	2.07	0.039	0.0014667	0.0550874
35-40 hours per week (reference: 41-49 hours)	0.0337767	0.0147377	2.29	0.022	0.0048856	0.0626678
41-49 hours per week (reference: 41-49 hours)	-0.0626823	0.0140833	-4.45	0	-0.0902907	-0.035074
Partnered	0.0781299	0.011779	6.63	0	0.0550389	0.101221
Children 0-4	0.0694577	0.015442	4.5	0	0.039186	0.0997295
Children 5-14	0.0075368	0.0122764	0.61	0.539	-0.0165292	0.0316028
Has long term health condition	-0.0478118	0.0140783	-3.4	0.001	-0.0754102	-0.0202134
Bachelor and above (reference: yr12 or below)	0.3092289	0.0133484	23.17	0	0.2830613	0.3353966
Vocational qualification (reference: yr12 or below)	0.0823698	0.0120043	6.86	0	0.0588371	0.1059025
Occupational segregation	-0.0084968	0.0025864	-3.29	0.001	-0.013567	-0.0034266
Industry segregation	0.0266268	0.0030945	8.6	0	0.0205605	0.032693
Tenure in current occupation	0.0055275	0.0006759	8.18	0	0.0042026	0.0068524
Tenure with current employer	0.000669	0.0008444	0.79	0.428	-0.0009863	0.0023243
Regular work schedule	-0.0116553	0.0135952	-0.86	0.391	-0.0383066	0.0149961
Firm size: 20-100 (reference: <20 employed)	0.0559333	0.0126326	4.43	0	0.031169	0.0806977
Firm size: 100-500 (reference: <20 employed)	0.1302226	0.0145037	8.98	0	0.1017903	0.158655
Firm size: 500+ (reference: <20 employed)	0.1935139	0.0166336	11.63	0	0.1609062	0.2261216
In a union	0.0288412	0.01253	2.3	0.021	0.0042779	0.0534045
Public sector employee (reference: private)	0.0510677	0.0154618	3.3	0.001	0.020757	0.0813783
Other organisation employee (reference: private)	0.003912	0.0162786	0.24	0.81	-0.0279998	0.0358237
Lives in other urban area (reference: urban)	-0.0613304	0.0130876	-4.69	0	-0.0869866	-0.0356741
Lives in rural area (reference: urban)	-0.061574	0.0164713	-3.74	0	-0.0938636	-0.0292844
Female	-0.1022897	0.0119084	-8.59	0	-0.1256344	-0.0789449
Constant	2.77725	0.029606	93.81	0	2.719211	2.835288
	Dependent var = lnhrlywage,	N=6116,	R-squared=0.2234,		Adjusted R-squared=0.2203	

Source: *Authors' calculations from HILDA, Wave 7 unit record data.*

Table B3 Olsen-Walby detailed simulated decomposition of the gender wage gap – all variables

	Men	Women	Change factor	Pooled coefficient	Simulation effect	Simulated change as a % of the wage gap	Cents/hour equivalent
	Mean	Mean	ΔX	B	BΔX		
Time in paid work – years	19.13	17.09	2.04	0.0024242	0.004945368	0.05	15.72
1-34 hours per week	0.1	0.41	-0.31	0.0282771	-0.008765901	-0.09	-27.87
41-49 hours per week	0.21	0.11	0.1	0.0337767	0.00337767	0.03	10.74
50+ hours per week	0.28	0.1	0.18	-0.0626823	-0.011282814	-0.11	-35.87
Partnered	0.68	0.67	0.01	0.0781299	0.000781299	0.01	2.48
Has children under 4	0.16	0.11	0.05	0.0694577	0.003472885	0.04	11.04
Has children 5-14	0.21	0.25	-0.04	0.0075368	-0.000301472	0.00	-0.96
Has long term health condition	0.15	0.15	0	-0.0478118	0	0.00	0.00
Bachelor qualification	0.27	0.33	-0.06	0.3092289	-0.018553734	-0.19	-58.98
Vocational qualification	0.39	0.29	0.1	0.0823698	0.00823698	0.08	26.18
Occupation segregation (male proportion x 10)	6.11	4.35	0.65	-0.0084968	-0.00552292	-0.06	-17.56
Industry segregation (male proportion x 10)	6.01	4.39	0.61	0.0266268	0.016242348	0.16	51.63
Tenure in occupation (years)	9.91	8.7	1.21	0.0055275	0.006688275	0.07	21.26
Tenure in current employment (years)	7.25	6.41	0.84	0.000669	0.00056196	0.01	1.79
Regular work schedule	0.83	0.83	0	-0.0116553	0	0.00	0.00
Firm size: 20-100 employed	0.28	0.33	-0.05	0.0559333	-0.002796665	-0.03	-8.89
Firm size: 100-500 employed	0.21	0.18	0.03	0.1302226	0.003906678	0.04	12.42
Firm size: 500+ employed	0.15	0.14	0.01	0.1935139	0.001935139	0.02	6.15
Union member	0.28	0.27	0.01	0.0288412	0.000288412	0.00	0.92
Public sector employee	0.14	0.25	-0.11	0.0510677	-0.005617447	-0.06	-17.86
Other sector employee	0.1	0.15	-0.05	0.003912	-0.0001956	0.00	-0.62
Other urban area	0.2	0.18	0.02	-0.0613304	-0.001226608	-0.01	-3.90
Rural area	0.11	0.11	0	-0.061574	0	0.00	0.00
Female	0	1	-1	-0.1022897	0.1022897	1.04	325.16
Constant	1	1	0	2.77725	0	0.00	0.00
				Sum of components	0.098463553	1.00	313.00

Source: *Authors' calculations from HILDA Wave 7 unit record data.*

APPENDIX C

GENDER INEQUALITY IN EDUCATION: IMPACT ON ECONOMIC PERFORMANCE

The World Bank (2001) undertook a comprehensive literature review on the relationship between gender and development, with a particular focus on gender inequality in education. The study claimed that conditions such as female illiteracy and low female education can inhibit productivity and earnings not only for women but also for whole economies. For example, in rural economies in Kenya, the lower yields of female farmers reflect lower levels of inputs or education relative to male farmers, thus increasing the education and input levels of female relative to male farmers can increase farming output as a whole (Quisumbing 1996 as discussed in World Bank 2001).

A number of other authors (Klasen 1999; 2002; Hill and King 1995; Dollar and Gatti 1999) have recognised the direct influences of gender inequality in education on economic performance (i.e. higher female-male ratios of total years of schooling are associated with higher economic growth).

Hill and King (1995) explored the impact of gender inequality in education (measured by female/male enrolment ratio) on Gross National Product (GNP). Their results showed that large gender disparities in educational attainment were associated with lower levels of GNP. These finding illustrated that in those countries where the ratio of female to male enrolments was less than 0.75, GNP was 25 per cent lower than those countries in which the ratio of female to male enrolments was close to one (greater than 0.95).

Dollar and Gatti (1999) applied a similar analysis when analysing the effect of education on growth of income per capita in developed countries, using female and male secondary educational achievement as the explanatory variable. They found that gender inequality in education lowered economic growth in developed countries: a one percentage point increase in the share of adult women with secondary school education increased income per capita by 0.3 percentage points in developed countries. However, their findings in relation to the effects of education for men and women on economic growth were mixed. When looking at secondary educational attainment, they found that female secondary education attainment is positively and significantly associated with growth in developed countries only, while male secondary education attainment is not significantly associated with growth in either developed or developing countries. Klasen (1999; 2002) however, suggests that these findings were possibly due to high multicollinearity between female and male educational attainment.

Klasen (2002) found that gender inequality in education has impacts on economic growth both in developing and industrialised countries. Differences between Klasen’s (2002) and Dollar and Gatti’s (1999) findings, Klasen argues, may be due to the shorter time period used by Dollar and Gatti (1999) (1975-90) in comparison to Klasen (2002) who used 1960-

1992 and also to their use of a different education variable, with Dollar and Gatti (1999) using a secondary educational attainment measure in contrast to Klasen (2002) who used total years of schooling, a more comprehensive measure of education.

In addition to the direct channels through which gender inequality in education may affect economic growth, Klasen (1999; 2002) measured other channels through which gender inequality might affect economic growth. He found that higher investment rates are related to higher productivity of the labour force and economic growth of income per capita. Reduced gender inequality in education is also related to higher investment rates. Klasen (2002) argued that a combination between lowering gender inequality and wage discrimination would increase investment as it would be profitable to invest in female-intensive industries.

Lagerlof (2003) further adds to our understanding of the channels through which gender inequality may affect economic growth by noting the possible impacts of human capital investment on the next generation. Using an overlapping generation framework applied to European data, Lagerlof (2003) found that a decrease in gender educational inequality induced a quantity-quality of children substitution which then created a fall in fertility, thus increasing human capital investment in children and increasing incomes per capita for the next generation. The results also showed that at the same time, increasing levels of human capital reduced mortality, inducing population growth to rise simultaneously with per capita income growth (GDP per capita). However, as the mortality level stabilises, and fertility continues to fall, population growth starts to fall, and per capita income also continues to rise and stabilise on a balanced growth path. Thus in the long run, Lagerlof finds that reducing gender inequality in education increases economic growth.

APPENDIX D

THE DEVELOPMENT OF GROWTH THEORY AND ITS APPLICATION TO STUDIES OF GENDER INEQUALITY

This appendix provides a history of the development of growth theory (originating from work by Solow and Swan in 1956), explaining the underlying concepts of economic growth which are embedded in this approach, and particular insights provided by endogenous growth theory. We discuss the complexities of these types of models and provide details of the ways in which previous studies of the macroeconomic impact of gender inequality have utilised growth theory.

The basic growth model, developed by Solow in 1956, is based upon how the flow of economic output, represented by GDP, is produced by two stocks of inputs - capital and labour in a production function. This basic model has been developed over the decades to include the impact of other aspects of the economy, such as human capital and financial development, on economic output. This impact upon economic output can be considered as the change in the two inputs (capital and labour), or the change in productivity, also known as total factor productivity (TFP).

The analytical framework of growth theory went through a boom after Solow and Swan launched their theories of growth in 1956. These neoclassical theories of growth assume one capital input, assuming a closed and single production function. Both the Solow and Swan models conclude that *in the long run* the growth of an economy will not be determined by aspects of the economy, with factors such as government policy having no long run impact. This type of system is referred to as an exogenous growth process. Starting with Romer (1986), Lucas (1988) and Barro (1990), the study of growth continued to expand, and frameworks were modified by including additional, previously un-measured types of capital input and developing several production functions for different types of output. This expansion and modification incorporated an understanding of economic growth as an endogenous process, and this understanding of the endogeneity of growth transformed the nature of the analytical framework. The endogeneity of growth reflects the theory that economic growth is determined by a range of variables that represent aspects of the economy. This increased interest in more complex economic relationships in predicting growth triggered a body of empirical work in search of the "real" sources of economic growth.

Initially, endogenous growth theory produced a very complicated growth model that could not be tested empirically (Sala-I-Martin 1997). Barro (1991) pioneered growth regression developments in order to produce an endogenous growth model that could be applied empirically, abandoning the structure of the theoretical model and using one linear model to ascertain which variables actually determined growth according to the data. Barro's (1991) work was followed by other empirical studies, examining a larger number of variables to seek the "real" determinants of growth (Mankiw *et al.* 1992; Levine and Renelt 1992; Sala-I-Martin 1997).

In the next section we describe and discuss the two most common frameworks that are used in growth analysis and how these are used to measure the impact of economic factors on economic growth. These two frameworks are (i) the Solow growth model [4] and (ii) growth accounting. We discuss the merits and limitations of each of these growth frameworks, provide additional detail about the development of methodologies based on growth theory and discuss the application of these techniques to studies of the impact of gender inequality on economic performance. As noted in the main report, we used a modified version of the Solow growth model in our calculations of the impact of the gender wage gap on economic growth.

SOLOW GROWTH MODEL

The Solow growth model uses a Cobb-Douglas production function that assumes constant returns to scale. The Cobb-Douglas production function can be described as

$$Y = K^{\alpha}(AL)^{1-\alpha} \qquad \text{(D1)}$$

where Y is output, K is physical capital, L is labour and A is an index of labour augmenting technology, while (AL) is defined as the number of effective units of labour (Mankiw *et al.* 1992)[5]. The assumption of constant returns to scale means if *all* production factors are increased in the same proportion, output will also increase by that proportion. This implies diminishing marginal returns for factors since the increase in output reduces if only one factor of production is increased.

In the capital accumulation process, labour and technology are assumed to grow at exogenous constant rates (that is, n and g in equation D2 below). Meanwhile, the accumulation of physical capital comes from the fraction of output invested in physical capital, which is constant and known as savings, s. In calculating the growth of physical capital, the model has also taken the rate of depreciation (δ) into account, to be subtracted from output. Therefore, the accumulation of capital per effective unit of labour can be described as:

$$\dot{k} = sy - (n + g + \delta)k \qquad \text{(D2)}$$

where k=K/AL, which is the stock of capital per effective unit of labour and y=Y/AL, which is output per effective unit of labour.

The capital accumulation process combined with the production function that assumes constant return to scale results in a steady state condition. This means that the return to output as a result of additional capital has been diminished and consequently, so has the ability to accumulate capital from the diminished output. This steady state condition has

[4] The Swan (1956) model was developed at the same time as the Solow (1956) model, but separately.

[5] α is known as the input share coefficient since it can represent the contribution of a particular input growth to the growth of output.

become the basis of the Solow model and predicts that increased savings and population growth are the main determinants of long run economic output per capita in the form of:

$$Ln\frac{Y}{L} = Ln(A) + gt + \frac{\alpha}{1-\alpha}Ln(s) - \frac{\alpha}{1-\alpha}Ln(n+g+\delta) \quad \text{(D3).}$$

The Solow neoclassical growth model has several critics. Most have focused on the inconsistency between the prediction in the model and the stylized facts. For example, the importance of the capital accumulation process as the only growth determinant (reflecting the exogenous growth theory that underlies this model) has attracted much criticism. Romer (2001) and Easterly and Levine (2001) point out that factors other than capital accumulation are responsible for income growth (i.e., endogeneity of growth), and point particularly to the role of technological progress.

GROWTH ACCOUNTING

An alternative theoretical approach to modelling economic growth is known as growth accounting, which originated in Solow's (1957) work and was further developed by Kendrick (1961) and Denison (1962).

Growth accounting is a technique used to decompose or 'break down' economic growth into components that are contributed by growth in population (labour), capital and other aspects of total factor productivity (Barro 1999). The main difference between the growth accounting framework and Solow's growth model is that growth accounting moves away from the assumptions in the Solow model related to the capital accumulation process and development of the production function.

Growth accounting's foundations are based upon a neoclassical production function as shown by:

$$Y = f(A, K, L) \quad \text{(D4)}$$

where Y is output, K is physical capital, L is labour and A is an index of technological stage or TFP. Growth accounting differentiates this production function into

$$\frac{\dot{Y}}{Y} = g + \left(\frac{F_K K}{K}\right)\frac{\dot{K}}{K} + \left(\frac{F_L L}{L}\right)\frac{\dot{L}}{L} \quad \text{(D5)}$$

or

$$LnY - \left(\left(\frac{F_K K}{K}\right)LnK + \left(\frac{F_L L}{L}\right)LnL\right) = g = f\left(x_1, x_2, \ldots\right) \quad \text{(D6)}$$

where g is technological improvement or TFP growth. As noted earlier, TFP can incorporate any aspect of the economy that can affect productivity, and thus it is possible to estimate g by replacing TFP with a set of determinant variables such as in equation D3 (Durlauf *et al.* 2004).

The main advantages of growth accounting are that the model is not based on the assumption that there will be convergence in the long run, and it allows the effect of many variables on economic growth to be estimated. On the other hand, the main problem with growth accounting is its assumption that marginal or incremental productivity is constant and represented by factor prices such as wages and rents (Barro 1999). This assumption has made the approach unable to cope with various developments in growth theory. For example, Romer (1986), and Lucas (1988) found that there may be increasing returns on some factors and spillover effects in the economy, which will be likely to increase the marginal productivity of the two inputs (i.e., not only change g but also $\left(\frac{F_K K}{K}\right)$ or $\left(\frac{F_L L}{L}\right)$). Another problem occurs when there are more inputs that cannot be distinguished clearly by the available data (Jorgenson and Griliches 1967). In this case, the marginal productivity of input cannot be calculated correctly and as consequence, TFP and its determinants cannot be estimated correctly either.

THE ENDOGENEITY OF GROWTH

Solow's growth model has been continuously developed to acknowledge more complex aspects and relationships within the economy, emerging out of theories that acknowledge the endogeneity of growth. One of the most well known developments is the addition by Mankiw *et al.* (1992) of the stock of human capital as another input in the production model. This new capital input is assumed to have a similar accumulation process to physical capital in the Solow-Swan model. The Mankiw *et al.* (1992) version of the model can be expressed as:

$$Ln\frac{Y}{L} = Ln(A) + gt + \frac{\alpha}{1-\alpha}Ln(s) - \frac{\alpha}{1-\alpha}Ln(n+g+\delta) + \frac{\beta}{1-\alpha}Ln(h) \quad \text{(D7)}$$

where h is the level of human capital and β is its input share to output growth. Empirically, Mankiw *et al.* (1992) claim the inclusion of additional capital in the growth regression reveals the real magnitude of the impact of the physical capital accumulation process on growth and argue that its impact is not as large as the Solow-Swan model predicted. Instead, the accumulation of human capital could account for a high percentage of the explanation for cross-country variations in income.

Another example of developments of the Solow-Swan model relates to the work done by Cass (1965) and Koopmans (1965) who investigated the endogeneity of the Solow-Swan physical capital saving based on Ramsey's (1928) notion of individual behaviour toward consumption and capital accumulation. They found that an individual's propensity to save depends on the return to capital, its time discount rate, and also the diminishing value of consumption utility. In other words, individuals will consider how much will be gained from saving, compared to the implicit loss of not being able to immediately consume. Becker *et al.* (1990) observe the endogeneity of population growth or, to be more exact, the fertility rate. They find that high fertility rates inhibit growth, especially because the production and rearing of children is very time-intensive. As a result, there will be an opportunity cost of having a child. It follows that the higher the individual income the higher this opportunity cost will be. This means higher incomes will have a negative impact on population growth.

In recent years we have seen a growing interest in theoretical and empirical determinants of economic growth in the context of understanding the endogeneity of growth. Durlauf *et al.* (2004) have argued that despite the complex structure of growth models, most factors, apart from capital and labour, can be classified as part of total factor productivity and consequently have a direct impact on economic performance. However, these factors can also affect growth indirectly through limiting or advancing the composition of capital and labour as input factors. Durlauf and Quah (1999) list 87 specific variables in 36 different categories that have been examined as growth engines or growth determinants, including gender issues. Within this framework, we can see that socio-economic factors including gender inequality can have an impact on economic growth through several channels, as well as directly.

In the next section, we discuss the approaches that have been taken to these issues by authors specifically examining the impact of gender inequality on economic growth.

METHODOLOGY USED IN PREVIOUS EMPIRICAL STUDIES ON THE IMPACT OF GENDER INEQUALITY ON ECONOMIC PERFORMANCE

Previous studies examining the impact of gender inequality (either inequality in education or wage inequality) on economic growth have used two main empirical methods.

(a) Econometric estimation

Simplistic exogenous gender inequality model

The empirical estimation used by Seguino (2000 p. 1216), is derived from the neoclassical production function as follows:

$$Y_{it} = A_{it} F(K, HKF, HKM)_{it} \qquad \text{(D8)}$$

where Y is output, A is technological change, K is capital stock, HKF_{it} *and* HKM_{it} are female and male human capital, i is the country index and t is time. Technological change is defined as $A_{it} = C_i(1+\phi t)e^{\sigma WGAP}$ where $WGAP_{it}$ is the gender wage gap and σ is the effect of gender wage differentials on growth.

Transforming the production function, Seguino (2000) then built several regressions using cross-country data from 1975-1995, taking into account a country's fixed effects to estimate the impact of the gender wage gap on economic growth, with this main equation exploring the direct channel and given by:

$$d\log Y_{it} = \phi + \sum \lambda_i + \alpha_1 WGAP_{it} + \alpha_2 d\log K_{it} + \alpha_3 HKF_{it} + \alpha_4 HKM_{it} + \varepsilon_{it} \qquad \text{(D9)}$$

where d is the difference operator (representing economic growth), ϕ is the growth rate of technological change when variables are measured at the mean, $\sum \lambda_i$ represents the fixed effects, $WGAP_{it}$ is the gender wage gap, HKF_{it} *and* HKM_{it} are female and male human capital, and ε_{it} is the error term.

Seguino (2000), in contrast to the methodology we use in our study, does not directly examine the ways in which the gender wage gap may be operating through indirect channels. Instead, she runs a separate estimation by regressing investment as a share of GDP on the gender wage gap and finds this relationship to be positive. Seguino (2000) also highlighted the importance of human capital variables as control variables.

In Seguino's regression, the results also show that while the coefficient of female human capital as a control variable is positive and significant on growth, the coefficient of male human capital is either negative and significant or completely insignificant. Klasen (1999; 2002) noted that these results were possibly due to a multicollinearity problem between female and male human capital. Klasen (1999; 2002) overcame this problem by instead including the ratio of female to male human capital along with a control variable which measured the overall state of human capital.

The next section will discuss an endogenous gender inequality model which acknowledges two-way causality between gender inequality and economic growth. That is, gender inequality influences growth, and growth influences gender inequality.

Endogenous gender inequality model

There is an additional body of econometric literature which acknowledges the endogeneity of gender inequality, and uses endogenous growth theory to estimate the relationship between gender inequality and macroeconomic performance. In relation to gender inequality studies, Klasen (2002) based his work on a modified version of the growth model, incorporating the Mankiw human capital refinements. He estimated the impact of gender inequality in human capital attainment on economic performance in a cross-country study. He did this by building up four equations as follows (Klasen 2002, p.355):

<u>Growth equation to estimate the direct channel of gender inequality on growth (D10)</u>

$$g = \alpha_1 + \beta_1 Inv + \beta_2 Popgro + \beta_3 LFG + \beta_4 ED_{60} + \beta_5 GED + \beta_6 RED60 + \beta_7 RGED + \beta_8 X + \varepsilon_1$$

<u>Indirect channels</u>

Investment equation (D11)

$$Inv = \alpha_2 + \beta_9 Inv + \beta_{10} Popgro + \beta_{11} LFG + \beta_{12} ED60 + \beta_{13} GED + \beta_{14} RED60 + \beta_{15} RGED + \beta_8 X + \varepsilon_2$$

Population growth (D12)

$$Popgro = \alpha_3 + B_{16} ED60 + \beta_{17} GED + \beta_{18} RED60 + \beta_{19} RGED + \beta_{20} X + \varepsilon_3$$

Labour force growth (D13)

$$LFG = \alpha_4 + B_{21} ED60 + \beta_{22} GED + \beta_{23} RED60 + \beta_{24} RGED + \beta_{25} X + \varepsilon_4$$

Where

g : average annual (compounded) rate of growth of GDP per capita 1960-1992

Inv :average investment rate 1960-1992

$Popgro$:average annual (compounded) rate of population growth 1960-1992

LFG :average annual (compounded) rate of labour force growth (15-64 years), 1960-1992

$ED60$:total years of schooling in 1960

$RED60$:female-male ratio of total schooling 1960

GED :annual (absolute) growth in total years of schooling 1960-1990

$RGED$:female-male ratio of growth in total years of schooling 1960-1990

X :other independent variables typically included in cross-country growth regressions (openness (average ratio of exports plus imports to GDP), initial income per capita in 1960 and regional dummy variables)

Using a path analysis method, Klasen (2002) was able to calculate the total effect of gender inequality upon growth, which is a summation of the direct and indirect effects.

Total effect = Direct effect + Indirect Effects

Examining the four main equations above, the total effect from gender inequality in education (i.e. $RED60$) can be calculated as follows:

$$\beta_6 + (\beta_{13} * \beta_1) + (\beta_{18} * \beta_2) + (\beta_{18} * \beta_9 * \beta_1) + (\beta_{23} * \beta_3) + (\beta_{23} * \beta_{10} * \beta_1) \text{ (D14)}$$

where β_6 is the direct effect of gender inequality in education,

$(\beta_{13} * \beta_1)$ is the indirect effect of gender inequality in education via investment,
$(\beta_{18} * \beta_2)$ is the indirect effect through population growth,

$(\beta_{18} * \beta_9 * \beta_1)$ is the indirect effect through population growth and investment,

$(\beta_{23} * \beta_3)$ is the indirect effect through labour force growth; and

$(\beta_{23} * \beta_{10} * \beta_1)$ is the indirect effect through labour force growth and investment

Klasen (2002) also then estimates a reduced form of the regression which omits the indirect channel to measure the total effect of gender bias directly on economic growth.

$$g = \alpha_5 + \beta_{26} ED60 + \beta_{27} GED + \beta_{28} RED60 + \beta_{29} RGED + \beta_{30} X + \varepsilon_5 \text{ (D15)}$$

Thus mathematically, β_{26} should replace equation (D14).

However, since equations (D11 to D13) do not contain all dependent variables from equation (D10), we may expect that β_{26} will not to be identical to above equation.

Since in equation (D10) above, economic growth depends not only on the total years of schooling and female-male ratio of total schooling in the initial period of 1960, but also on the growth of these variables (*GED* and *RGED*), there is a possibility that a reverse causality exists from economic growth to both *GED* and *RGED* . Therefore, *GED* and *RGED* are endogenous variables. Klasen (2002) trialled several specifications, including the use of instrumental variables, to address this simultaneity problem. Interestingly, Klasen (2002) found that the causality runs from gender inequality in education to economic growth and not vice versa.

Although Klasen (1999; 2002) concentrated on the impact of gender inequality in education, intrinsically, his "path analysis" methodology should be able to be applied to other types of gender inequality, including the gender wage gap.

Dollar and Gatti (1999, p. 6) also acknowledged the endogeneity issues of gender inequality in a much simpler framework, however they only measure the direct effect of gender inequality to growth, and do not examine indirect channels.

(b) Calibration or simulation models

While the econometric estimations discussed above focus more on the empirical aspects of the impact of gender inequality to growth, another stream of literature builds a more complex theoretical model, by setting up several functions in the economy, then carrying out a calibration or simulation on the baseline economy with relevant parameters.

Researchers who have applied calibration methodology to studying the impact of gender inequality on economic performance (Calvacanti and Tavares 2007 and Caro 2008) developed an overlapping generations model to study the cost of the gender wage gap, particularly following Galor and Weil (1996). Galor and Weil (1996) linked fertility and economic growth by showing that increasing women's relative wages reduces fertility and increases labour force participation, by raising the opportunity cost of having children.

In this "overlapping generations model", women and men are assumed to live for three periods. During the first period, women and men are children and consume a fraction of their parent's time endowment (parental care). During the second period of life, these children grow up to become adult men and women, organise as couples and make decisions together (divisions of household labour) about time at work, time raising children, and savings for the future. In the third period, each couple consumes what they saved in the second period. This overlapping generations model is based on very strong assumptions and is highly structured: for example, it is assumed that only men will be engaging in production involving physical work.

Calvacanti and Tavares (2007) and Caro (2008) used a simple growth model with endogenous female labour participation and fertility to estimate the cost of gender discrimination. They set up a production function in which gender wage discrimination is introduced where women only receive a fraction of men's salary, a utility function of the household and a budget constraint to set up equilibrium in the labour market, capital market and goods market.

After equilibrium was reached, calibration was conducted. Calvacanti and Tavares (2007) calibrated the performance across countries using the United States economy as the baseline economy to estimate whether wage discrimination explained the whole output difference in comparison to US output. On the other hand, Caro (2008) calibrated the performance of the Spanish economy by simulating different discrimination levels.

APPENDIX E

MACROECONOMIC ESTIMATION METHODS AND RESULTS

One of this project's key outputs is an estimation of the impact of the gender wage gap on economic performance or GDP as a proxy of economic output. In Appendix D, we discussed the range of approaches that have been taken in previous studies to conduct this type of estimation. We have opted to conduct the estimation using a regression technique based on the expansion of the Solow growth model, similar to that used by Klasen (1999; 2002) and Dollar and Gatti (1999). We consider this endogenous econometrics approach to be the most robust modelling technique for our purposes because it reflects the actual nature of the relationships that exist within the broad macroeconomic framework (based on a widely accepted model) without relying on a very strong assumption about the growth processes such as in an exogenous growth model or the strict structure of the calibration model.

As we are studying the impact of the gender wage gap in a particular country - Australia – we need to use time series data for our estimation (as opposed to cross-sectional data which can be used when a number of countries are being studied together). While both cross-sectional and time series data can pose problems for the type of regression analysis we are undertaking here, a single country, time series analysis has a number of advantages. First, the Solow model was originally built to examine the dynamics of one economy (Solow 2001) and although it can be adapted for cross-country analysis, was not explicitly intended for this purpose. Second, some technical problems common to cross-sectional, multi-country studies are less likely to be experienced in a one country, time series studies such as ours[6].

We use data from a variety of Australian Bureau of Statistics sources. The variables we include in the model (along with a measure of the gender wage gap) are those common to macroeconomic analysis of the kind being undertaken here. They are GDP, investment, human capital, fertility, labour participation and hours worked. Issues related to data availability (common in single country studies where data points may be limited) have affected both our modelling approach and our results. These issues are discussed in more detail below. In macroeconomic modelling, the longer the time period of data available, the more likely that such a time span smooths out the ups and downs of the business cycle. As one of our key variables – human capital – was only available from 1989 onwards, we were able to include only 20 data points in our analysis. However, we were able to test and adapt our model to address the issue of a relatively short span of data. Nevertheless, these data limitations should be kept in mind when interpreting our results.

The use of time series data raises the issue of the stability of the data over time – referred to as stationarity (Durlauf *et al.* 2004). In essence, unless we can be reasonably confident that

---

[6] These include fewer problems related to omitted variable bias (see Greiner *et al.* 2003).

the variables we use have a consistent average over the long run, we cannot be sure that the relationships we see from a regression analysis are accurate, or if in fact they simply (or partly) capture fluctuations over time in the data. We ran a series of tests related to stationarity (described below) and found that we could not be sufficiently confident of the stationarity of our data without making an additional adaptation to our model

In order to adapt the model to overcome possible problems with stationarity, we used first difference equations, rather than equations which use the data in what is known as a level form. This approach resolves problems of data stationarity by subtracting the equation at time t from the equation at time t-1. While this method means that we lose one observation point, the increased confidence that we have in our results due to 'first differencing' the model is very important.

Below we describe the database used in our modelling, including details of our rationale for choosing particular variables and data (including issues of data availability), and the characteristics of the data. We then describe our estimation technique, and present detailed results. Finally we discuss how we come up with an estimation of the total impact to the Australian economy of the gender wage gap, and provide details about the limitations of our modelling.

DATA AND VARIABLES

In this section we describe the ways in which we have operationalised variables for use within our growth model. We then summarise the chosen variables and the data used at the end of each variable description. Finally, the variables we have used and their data sources are provided in Table E1.

Economic Output

As discussed earlier in this report, economic performance is the main variable of interest in this study and it is often measured by economic output, in particular gross domestic product (GDP). GDP is a suitable measure for examining the performance of developed countries, while for developing countries other measures such as the Human Capital Index and the achievement of Millennium Development Goals have been used to measure economic performance. GDP measures the output produced by activities in the economy. However, when using GDP as a measure of economic performance, a set of decisions needs to be made in relation to definitions (and thus in turn about data).

There are three common approaches used to calculate GDP – value added, expenditure and income (Mankiw 1997). The value added approach measures GDP by summing up all goods and services in the economy subtracted by the intermediate products used to produce them. By doing so, the calculation will not "double count" the products that are part of other products. The second approach is based on income. The income approach to GDP is calculated by summing the income of capital, labour, and entrepreneurship having first deducted the cost of tools and materials. Finally, the expenditure approach uses an assumption that for every seller there is a buyer. Therefore, with the expenditure approach, GDP can be calculated by summing up all consumption, investment, government expenditure and net exports.

Figures of GDP which are based on these three approaches are available in the Australian Bureau of Statistics (ABS) publication Australian National Accounts: National Income, Expenditure and Product (Cat. no. 5206.0) in which quarterly data are available from June 1960 and in the publication Australian System of National Accounts (Cat. no. 5204.0) which provides annual data. Conceptually, the income, expenditure and production approaches to measuring GDP should achieve the same results; however this is dependent upon how the data are derived. In Australia, prior to 1994-95, these estimates were produced independently, and consequently there are statistical discrepancies between each estimate. However, over the period in which these estimates were calculated separately, the ABS also provided a compilation number, representing the average of these three approaches. Since 1994-95, there are no discrepancies between any of the estimates, with the exception of the quarterly estimates for the June 1994 quarter (ABS 2009).

Another issue in relation to the use of GDP as a proxy of economic output is whether to use GDP in constant or current prices. Current price GDP means that the currency amount of products in one economy is calculated based on the price of the product at that year or quarter. However, if our interest in GDP is to conduct economic growth analysis, and as this is based on the production (supply) side of the economy, we need the representation of quantity produced in the economy where the change in the value of output is not merely caused by the current prices of those outputs. The constant price GDP provides this representation since although it is reported in currency units it nevertheless is based on estimated fixed prices in the base year. Therefore, the changes in GDP value will represent the changes in the quantity of goods and services produced, avoiding the effects of price changes on the output calculation. The ABS now uses a chain volume measure of GDP in the Australian System of National Accounts as their measure of constant price GDP. The ABS uses the chain volume measure because it takes into account different growth rates of unit prices of goods and services as well as the growth rate of their quantities, which results in changes in price relativities. The chain index used in chain volume GDP is calculated based on annual changes so that price relativities remain the same (5216.0 - Australian National Accounts: Concepts, Sources and Methods, 2000).

Variable: Gross domestic product (GDP)

Data used: Chain volume measure of GDP, annual series, ABS Australian System of National Accounts (Cat. no. 5204.0)

Capital and Investment

Capital stock is one of the two major inputs that produce economic output (the other is labour). The addition to capital stock is known as 'investment'. Data measuring capital stock are available annually in the publication Australian System of National Accounts (ABS Cat. no. 5204.0). Investment can be calculated by an increase or decrease in capital stock. However, a more direct measure of investment is gross fixed capital formation, available in the publication Australian National Accounts: National Income, Expenditure and Product (ABS Cat. no. 5206.0) or total capital accumulation and net lending, available in the publication Australian System of National Accounts (ABS Cat. no. 5204.0). We opted to use total capital accumulation and net lending in the Australian System of National Accounts since this takes into account changes in financial capital as well in physical capital. We use this variable measured as a ratio of GDP (i.e. the rate of investment).

In most of the growth models discussed earlier, investment predicts economic performance. This theory is supported by empirical results that estimate the robustness of this variable compared with other growth determinants (Barro 1991, Levine and Renelt 1992, Mankiw *et al.* 1992, Caselli *et al.* 1996). Nevertheless, studies that compare the effects of investment and income on economic growth have shown that the superiority of investment as a predictor of economic growth is not as strong as earlier predictions would suggest (Mankiw *et al.* 1992; Easterly and Levine 2001). Moreover, the direction of causality could be from growth to investment rather than vice versa (Blomstrom *et al.* 1996). Barro (1996) shares this view by claiming that high growth prospects are one of the main drivers of high investment[7].

Variables: Gross fixed capital accumulation (Total investment)

Data used: Chain volume measure of gross fixed capital accumulation, Australian National Accounts: National Income, Expenditure and Product (Cat. No. 5206.0)

Labour and Fertility

In addition to capital, labour is the other major input to economic growth. There are several ways to measure labour input, including the number of people participating in the labour force, hours of work and their cost (i.e. wages). In the publication - Labour Force, Australia, Detailed, Quarterly (Cat. no. 6291.0.55.003), the ABS provides two of the three measures of labour input, which are the number of employed persons by full-time and part-time status and their hours of work. We use these measures rather than a wages measure, as the former are more commonly used in Australian macroeconomic modelling, and in order to better separate our labour variable from our gender wage gap variable (which is, of course, based on wages).

It is also important to include fertility in growth models such as the one we are using in this study. The Solow growth model (on which our estimation technique is based) has a strong assumption of diminishing marginal returns of labour. This means that although an increase in labour is likely to increase economic output, theoretically it will decrease labour productivity if there is no change in other inputs or aspects of the economy. In addition, the model also assumes that the labour force is growing in the rate of population growth. As a result, an increase in labour can also decrease the output or income per capita of an economy given that it is solely boosted by population growth. Most cross-country studies

[7] There is also a body of literature which suggests that different types of investment (for example, government investment versus private investment) may have different impacts on growth (see, for example, Barro 1990; Folster and Henrekson 1999).

have supported this hypothesis. The strongest result to confirm this negative impact of population growth on economic growth is from Mankiw *et al.* (1992) while Levine and Renelt (1992) produce a less significant relationship. However, if the fertility rate is controlled for in economic growth models, the impact of population growth on income per capita may also have a positive sign. Barro and Lee (1994) argue that this positive result is due to either positive net migration (reflecting an increase in skilled workers which has a positive impact on growth) or a low mortality rate that indicates a better health system.

Variables: Labour: Number of employed persons and hours of work; Fertility: Total Fertility Rate

Data used: Labour: Labour Force, Australia, Detailed, Quarterly (Cat. no. 6291.0.55.003)

Fertility: Fertility data in Australia is available annually in Births, Australia publication (ABS cat. no. 3301.0) or in Australian Historical Population Statistics (ABS cat. no. 3105.0.65.001).

Human Capital

Human capital is the most common additional variable used in growth estimations (Durlauf *et al.* 2004). Lucas (1988) defined human capital as a general skill level, indicating that human capital contributes to production by increasing worker productivity as well as directly increasing output through its contribution to technological improvement. Research has shown that the impact of human capital accumulation contributes to greater technical progress or Total Factor Productivity (TFP) growth by affecting knowledge accumulation (Romer 1990). The indirect impact of human capital accumulation mostly occurs through lower fertility or population growth (Becker *et al.* 1990; Galor and Weil 2000).

There are various types of human capital measurement, which are based on (i) health – mostly using life expectancy as a proxy; and (ii) education, mostly using school attainment as a proxy. Previous studies that have discussed gender inequality in human capital have mostly focused on education (Galor and Weil 2000; Klasen 1999; Seguino 2000), and that is the measure we use in this study. Annual data for this educational proxy of human capital, is available in ABS data on educational attainment in the publication Education and Work, Australia (ABS cat. no. 6227.0). Data begins in 1989, with details for the number of persons and workers by their highest educational attainment (in terms of non-school qualifications).

Variables: Proportion of persons with non school qualification

Data used: Education and Work, Australia (ABS cat. no. 6227.0)

Gender Wage Gap

The gender wage gap is the main input variable of interest in this study as our primary aim is to estimate the impact of the gender wage gap on economic growth. The connectivity between the gender wage gap and economic growth has been discussed in other parts of this report. The ABS conducts several surveys which have data and information about the gender wage gap. In this estimation we will use Average Weekly Earnings, Australia (Cat. no. 6302.0), which provides quarterly estimates of average gross weekly earnings of workers in Australia. Average Weekly Earnings data was chosen because it provides the longest time span of consistent data on wages.

Variable: Ratio of wage difference between male and female to male wage

Data used: Average Weekly Earnings, Australia (Cat. no. 6302.0) (Full time adults, ordinary time earnings)

Table E1 Summary of Variables and Data Used

Variable	Proxy	Sources	Availability
Economic Output	GDP	Australian System of National Accounts (Cat. No. 5204.0)	Annually
Economic Input			
Investment	Gross Fixed Capital Formation	Australian National Accounts: National Income, Expenditure and Product (Cat. No. 5206.0)	Annually and quarterly
Labour	Labour participation (Number of employed persons full time and part time)	Labour Force, Australia, Detailed, Quarterly (Cat. No. 6291.0.55.003)	Quarterly
	Hours of work	Labour Force, Australia, Detailed, Quarterly (Cat. No. 6291.0.55.003)	Quarterly
Fertility	Total fertility rate	Births, Australian Publication (Cat. No. 3301.0)	Annually
		Australian Historical Population Statistics (Cat. No. 3105.065.001)	Annually
Human capital (Education)	Proportion of persons with non school qualifications	Education and Work, Australia (Cat. No. 6227.0)	Annually
Gender wage gap	Ratio of wage difference between male and female to male wage	Average Weekly Earnings, Australia (Cat. No. 6302.0)	Annually and quarterly

STATISTICAL SUMMARY

During the past two decades Australia's GDP based on the chain volume measure grew by 3.4 per cent annually from $585 billion in 1989 to $1,084 billion in 2008. Meanwhile GDP per capita grew by 2.1 per cent annually from $35,000 in 1989 to $51,000 in 2008. Although both GDP and GDP per capita showed an increasing trend during this period, there was slight negative economic growth (especially in terms of GDP per capita) in the period 1991-1992. The overall increasing growth trend was also experienced by capital stocks, which increased from $1,750 billion in 1989 to $3,178 billion in 2008. This included an exponential trend in capital stocks between 2000 and 2008. However, there was a considerable drop in the flow of investment during 1989-1992, before experiencing a steady increase since then.

In the same period of 1989-2008, labour grew by 2.0 per cent annually, from 7.7 million people participating in the labour force in 1989, to 10.8 million people by 2008. In line with the fall in investment flow between 1990 and 1992, there was also a drop in labour force numbers over that period, but this was less marked than the drop in investment. The amount of hours of worked grew at a somewhat lower pace than overall labour force participation, at 1.8 per cent annually: from 267,000 hours a week in 1989 to 355,000 hours a week in 2008. Putting together the data on persons in the labour force with the information about hours worked, we find that the average hours of work of those in the labour force dropped by 0.2 per cent annually over the period, falling from 34.6 hours per week in 1989 to 32.9 hours per week in 2008. In terms of total hours of work, there was also negative growth in 2001 and 2006, as well as negative growth during the 1991-1992 downturns.

Overall population numbers grew slower than the number of labour force participants during the 1989-2008 period. At 1.3 per cent annually, the population grew from 16.7 million people in 1989 to 21.2 million people in 2008. The growth rate actually slumped from 1.8 per cent in 1989 to below 1.0 per cent in 1999. However, the growth rate then increased steadily to 1.2 per cent in 2005 and more gradually to 1.7 per cent in 2008. The latest increase in population growth is driven in part by the a jump in the Total Fertility Rate from 1.7 live births per 1000 women in the 2001 calendar year to 1.9 live births per 1000 women in 2007. This happened after fertility gradually decreased from 1.9 live births per 1000 women in 1990.

The percentage of persons aged 15-64 who attained a non-school qualification or tertiary education also increased substantially between 1989 and 2008. In 1989, 39.2 per cent of persons aged 15-64 held a non school qualification. This had increased to 53.9 per cent in 2008. The trend shows that the percentage of persons with non-school qualifications or tertiary education generally grew from year to year except for 1993, 1994 and 1997. In terms of gender inequality, the school attainment gap has closed considerably. In 1989, 45.4 per cent of men aged 15-64 had a non school qualification, compared to 39.2 per cent of women. These figures had increased to 55.3 per cent for men and by 52.6 per cent for women by 2008.

STATIONARITY

To be able to produce a reliable and consistent estimate when conducting econometric estimation, it is necessary that all variables have a stationary data stream – that is, that they are not being unduly influenced by other factors such as the normal peaks and troughs of the business cycle. Stationarity means that the variable is likely to converge to a single value or a trend path in the long run, and thus ensures that the mean, variance and relationships we are going to estimate will be achieved in the long run (Enders 1995). Without stationarity, we risk estimating a 'false' relationship that is based on insufficient serial autocorrelation within each variable. In addition, it is always important that the value that represents a variable at one moment, in a certain set of conditions, shows a consistent long run value, given everything else being the same, and not a value that will be likely to change in the short term.

Thus before undertaking the sort of estimation procedure we are using in this project, it was important to find out if the variables we are using have stationarity or not. Well-established tests exist to identify this, and we used these to explore the extent of non-stationarity in our data[8]. These tests, and the results from our preliminary modelling, when considered overall, led us to the conclusion that there was some degree of non-stationarity in our data. We addressed this through incorporating a first differencing approach into our modelling, as described in the following section.

THE ESTIMATION MODEL AND PROCEDURE

The Growth Model

Having completed our preliminary testing, and drawing lessons from our initial results, we then went on to develop our final estimation approach.

First, one of the most important findings from the stationarity tests of the potential variables was that our capital stock variable was not stationary in either the level (original) form or the first difference form. As a result, we could not use that variable to build the model. This in turns means that we cannot develop the model based on the growth accounting structure (i.e., equation D6 in Appendix D). Therefore, the model was built based on the Solow growth model, in particular, the structure that is offered by Mankiw *et al.* (1992) which is known as the augmented Solow growth model in the form of:

$$Ln\frac{Y_t}{L_t} = Ln(A_t) + \beta_1 Ln\left(\frac{I_t}{Y_t}\right) + \beta_2 Ln(h_t) + \beta_3 Ln(n_t) + \varepsilon_t \qquad \text{(E1)}$$

where A is Total Factor Productivity (TFP), Y is output (GDP), I is physical capital accumulation or investment, L is labour[9], h is human capital level (proxied by the proportion of population with non school qualifications) and n is population growth

---

[8] We used both the Augmented Dickey-Fuller (ADF) unit root test (Dickey and Fuller 1979) as well as an additional stationarity test based on Kwiatkowski *et al.* (1992). Results of this testing are available on request.

[9] We have chosen the ratio of labour (measured by total employed) to total population instead of the labour force participation rate to represent labour input for two reasons :

(i) mathematically, to be consistent with the dependent variable which is GDP per capita

(ii) It is more suitable in gaining an accurate measure of the impact of the gender wage gap on productivity, which covers the population who are employed only.

(proxied by the fertility rate). Following Seguino (2000), the technological or TFP change is formulated in the form of:

$$A_{it} = C_i(1+\phi t)e^{\sigma_0 Wgap_t} \qquad \text{(E2)}$$

where Ci is a country-specific variable which in our case should be constant because we only use one country, ϕ is the time effect and Wgap is the gender wage gap, calculated as $\left[\frac{w_m - w_f}{w_m}\right]$ where wm is average gross weekly earnings of male workers and wf is average gross weekly earnings of female workers.

Substituting this TFP formulation into the equation (1), we will have

$$Ln\frac{Y_t}{L_t} = C + \phi t + \sigma_0 Wgap_t + \beta_1 Ln\left(\frac{I_t}{Y_t}\right) + \beta_2 Ln(h_t) + \beta_3 Ln(n_t) + \varepsilon_t \qquad \text{(E3)}^{10}$$

The next step is to relax the assumption that the growth of labour will be equal to population growth, which allows us to better measure labour participation. This is done because we want to explore the impact of the wage gap on workers' productivity based on the total hours of work that they offer, rather than just whether or not they work at all. With the need to include the gender wage gap in the equation, we then adjust equation (E3) to be:

$$Ln\frac{Y_t}{P_t} = C + \phi t + \sigma_0 Wgap_t + \beta_1 Ln\left(\frac{I_t}{Y_t}\right) + \beta_2 Ln(h_t) + \beta_3 Ln(n_t) + \beta_4 Ln\left(\frac{L_t}{P_t}\right)$$
$$+ \beta_5 Ln\left(\frac{Hw_t}{L_t}\right) + \varepsilon_{0t} \qquad \text{(E4)}$$

[10] Given the fact that $Ln(1+\phi t) \approx \phi t$

where Hw is hours of work and P is the total population so the average hours of work is calculated as $\left[\dfrac{Total\ hours\ of\ work(Hw_t)}{Labour(L_t)}\right]$. With this equation, we can examine the direct effect of the gender wage gap on economic growth.

1.2 ESTIMATION METHODOLOGY

As discussed elsewhere, there are several possible modelling options within the overall framework described above, and we need to know more about the possible presence of indirect channels within the model in order to decide on a suitable estimation approach. Therefore, we first carried out a sensitivity analysis to examine the existence of indirect channels from the gender wage gap to economic growth, by taking in and out of the model variables that the literature suggests may be indirect channels of the gender wage gap's impact on economic growth (for example, fertility and investment). We found that when we added in and removed variables there were changes in the magnitude of the coefficients of the gender wage gap and that the coefficients of the gender wage gap in various combinations of specifications were not robust. These findings confirm the existence of indirect channels (if the only effect of the wage gap on economic growth was direct, then we would expect to see little change in the coefficients when variables are added into or removed from the model). These findings show that our best estimation option was to conduct simultaneous equations using three-stage least squares in which the independent variables are mostly endogenous, allowing us to examine the direct and indirect channels together.

The essential component of the three-stage least squares estimation is the existence of enough instrumental variables[11] to allow us to have combinations that have sufficient explanatory power in relation to the variable that is instrumented[12]. This should include adequate exogenous instrumental variable(s), otherwise the equations will end up in perfectly multicollinearity. In other words, we need to have variables in our model which explain much of the variation in the instrumented variable (here, rate of investment, average hours of work, fertility, human capital and labour participation), but are not determined by the outcome variable (which here is GDP per capita) which would retain the endogeneity problem. To achieve this, we use the lag of instrumented variables as the exogenous instrumental variables, given that the lag variables are likely to determine current conditions (that is, current GDP per capita) but theoretically cannot be determined by current GDP per capita. Using these lag variables means that we use the instrumented

---

[11] An instrumental variable (IV) is a variable which is uncorrelated with the error term in the main (growth) equation and is correlated with the endogenous explanatory variable.

[12] We need to make sure that the impact of the wage gap and other instrumental variables included in estimating each indirect channel will have a good representation of that channel. If not, this system will perform poorly.

variable at time t-1, t-2 and so on until we get a reasonable estimate of the instrumented variable at time t. Unfortunately, in doing so, we are losing the ability to estimate the impact of the gender wage gap through human capital accumulation, as data for this variable is not available before 1989. Introducing a human capital lag would cut many of the available observations in the model, and as the other instrumented variables are available from 1985 or earlier, we decided to take advantage of this longer period of data for the other variables, despite losing our ability to use human capital accumulation as an indirect channel. The human capital variable is still included in the model, but the extent to which the gender wage gap may operate through it to impact on GDP per capita cannot be estimated.

As our stationarity tests had given us slightly inconclusive results in regard to the stationarity of our variables (the tests suggested a possible problem with stationarity, but when analysed carefully did not provide sufficiently strong evidence to absolutely reject the possibility that our data was stationary), we needed to conduct some initial modelling in order to further inform our understanding of possible problems with stationarity. If overall we considered stationarity to be a problem, then we would need to address this by using first difference terms of our variables (subtracting the equation at time t from the equation at time t-1), rather than the original (or level) variables.

We conducted several estimations using both the level term and the first difference term of the variables, in order to see if our results could provide any further information about which of these approaches we should use. We found that, using the level terms, we came up with a very high R^2 value (99 per cent)[13]. Very high R^2 values like this are a cause for concern. In this case, such a high value may be an indication of model misspecification and also may point to the existence of spurious regression caused by variables that are not sufficiently stationary. In contrast, the estimation based on first difference variables produced an R^2 value of 77 per cent. Based on these findings, and in the light of the stationarity testing, we opted to estimate the impact of gender wage gap on economic growth using a first difference equation[14].

To estimate the equation using a first difference equation, we subtract equation (E4) with

[13] The R^2 value indicates the total amount of variation in the dependent variable (GDP per capita in the equation above) which is explained by the variables being used in the model.

[14] One possible disadvantage of the estimation using a first difference equation is, as discussed in Engle and Granger (1987) following Phillips (1957), that the coefficient estimated in this equation may not hold in the long run especially when there is a temporary relationship that causes a temporary disequilibrium in the model. However, we also conducted the estimation at the level form and in many respects, the two sets of results are similar, with the magnitude of the total effect of the gender wage gap on growth being closely comparable using both estimation methods (-0.51 against -0.57). An exception to this overall similarity between the two approaches relates to an analysis of the indirect channels of impact. The level estimations indicate that the negative effect of the wage gap on economic growth is much more strongly associated with a drop in investment than in average hours of work, while our first difference findings show that hours of work is a much more important channel. The latter result is much more commonly found in previous studies than the former, which provides further support for our choice of first differencing.

$$Ln\frac{Y_{t-1}}{P_{t-1}} = C + \phi(t-1) + \sigma_0 Wgap_{t-1} + \beta_1 Ln\left(\frac{I_{t-1}}{Y_{t-1}}\right) + \beta_2 Ln(h_{t-1}) + \beta_3 Ln(n_{t-1})$$
$$+ \beta_4 Ln\left(\frac{L_{t-1}}{P_{t-1}}\right) + \beta_5 Ln\left(\frac{Hw_{t-1}}{L_{t-1}}\right) + \varepsilon_{0t-1} \quad \text{(E5)}$$

and produce

$$\Delta Ln\left(\frac{Y_t}{P_t}\right) = \phi + \sigma_0 \Delta Wgap_t + \beta_1 \Delta Ln\left(\frac{I_t}{Y_t}\right) + \beta_2 \Delta Ln(h_t) + \beta_3 \Delta Ln(n_t) + \beta_4 \Delta Ln\left(\frac{L_t}{P_t}\right)$$
$$+ \beta_5 \Delta Ln\left(\frac{Hw_t}{L_t}\right) + \Delta\varepsilon_{0t} \quad \text{(E6)}$$

The next step is to build the system to recognise the endogeneity of the growth determinants and hence the indirect effect of the wage gap through other growth determinants. As with the main growth equation, we also estimate all equations in this system using the first difference form, especially given the variables we use in these subsequent equations are simply the first difference terms of those used in the main growth equation. This system can be written as:

$$\Delta Ln\left(\frac{I_t}{Y_t}\right) = \alpha_1 + \beta_6 \Delta Ln\left(\frac{I_{t-1}}{Y_{t-1}}\right) + \beta_7 \Delta Ln\left(\frac{I_{t-2}}{Y_{t-2}}\right) + \beta_8 \Delta Ln\left(\frac{I_{t-3}}{Y_{t-3}}\right) + \beta_9 \Delta Ln(h_t)$$
$$+ \beta_{10} \Delta Ln(n_t) + \beta_{11} \Delta Ln\left(\frac{L_t}{P_t}\right) + \beta_{12} \Delta Ln\left(\frac{Hw_t}{L_t}\right) + \beta_{13} \Delta Ln\left(\frac{Y_t}{P_t}\right) + \sigma_1 \Delta Wgap_t + \Delta\varepsilon_{1t} \quad \text{(E7)}$$

$$\Delta Ln(n_t) = \alpha_2 + \beta_{14} \Delta Ln(n_{t-1}) + \beta_{15} \Delta Ln(n_{t-2}) + \beta_{16} \Delta Ln(n_{t-3}) + \beta_{17} \Delta Ln(n_{t-4})$$
$$+ \beta_{18} \Delta Ln(h_t) + \beta_{19} \Delta Ln\left(\frac{I_t}{Y_t}\right) + \beta_{20} \Delta Ln\left(\frac{L_t}{P_t}\right) + \beta_{21} \Delta Ln\left(\frac{Hw_t}{L_t}\right) + \beta_{22} \Delta Ln\left(\frac{Y_t}{P_t}\right)$$
$$+ \sigma_2 \Delta Wgap_t + \Delta\varepsilon_{2t} \quad \text{(E8)}$$

$$\Delta Ln\left(\frac{L_t}{P_t}\right) = \alpha_3 + \beta_{23} \Delta Ln\left(\frac{L_{t-1}}{P_{t-1}}\right) + \beta_{24} \Delta Ln\left(\frac{L_{t-2}}{P_{t-2}}\right) + \beta_{25} \Delta Ln\left(\frac{L_{t-3}}{P_{t-3}}\right) + \beta_{26} \Delta Ln(h_t)$$
$$+ \beta_{27} \Delta Ln(n_t) + \beta_{28} \Delta Ln\left(\frac{I_t}{Y_t}\right) + \beta_{29} \Delta Ln\left(\frac{Hw_t}{L_t}\right) + \beta_{30} \Delta Ln\left(\frac{Y_t}{P_t}\right) + \sigma_3 \Delta Wgap_t + \Delta\varepsilon_{3t} \quad \text{(E9)}$$

$$\Delta Ln\left(\frac{Hw_t}{L_t}\right) = \alpha_4 + \beta_{31}\Delta Ln\left(\frac{Hw_{t-1}}{L_{t-1}}\right) + \beta_{32}\Delta Ln\left(\frac{Hw_{t-2}}{L_{t-2}}\right) + \beta_{33}\Delta Ln\left(\frac{Hw_{t-3}}{L_{t-3}}\right)$$
$$+ \beta_{34}\Delta Ln(h_t) + \beta_{35}\Delta Ln(n_t) + \beta_{36}\Delta Ln\left(\frac{L_t}{P_t}\right) + \beta_{37}\Delta Ln\left(\frac{I_t}{Y_t}\right) + \beta_{38}\Delta Ln\left(\frac{Y_t}{P_t}\right)$$
$$+ \sigma_4 \Delta Wgap_t + \Delta\varepsilon_{4t} \qquad \text{(E10)}$$

$$\Delta Wgap_t = \alpha_5 + \beta_{39}\Delta Wgap_{t-1} + \beta_{40}\Delta Wgap_{t-2} + \beta_{41}\Delta Wgap_{t-3} + \beta_{42}\Delta Ln(h_t)$$
$$+ \beta_{43}\Delta Ln(n_t) + \beta_{44}\Delta Ln\left(\frac{L_t}{P_t}\right) + \beta_{45}\Delta Ln\left(\frac{I_t}{Y_t}\right) + \beta_{46}\Delta Ln\left(\frac{Y_t}{P_t}\right) + \beta_{47}\Delta Ln\left(\frac{Hw_t}{L_t}\right) + \Delta\varepsilon_{5t}$$

(E11)

Estimation Results

Tables E2 to E7 show the estimation result of equations (E6)-(E11). The system is estimated using the data described above with t from 1990 to 2008. We utilise the coefficients from these tables to generate the impact coefficient that we use to estimate the impact of the gender wage gap on economic growth as shown in Table E8.

Table E2 Growth Equation

		Coefficient		Standard Error
$\Delta Ln\left(\frac{Y_t}{P_t}\right)$	R^2= 0.7693			
(GDP per capita)				
	$\Delta Ln\left(\frac{I_t}{Y_t}\right)$	0.081	**	0.032
	$\Delta Ln(h_t)$	0.060		0.065
	$\Delta Ln\left(\frac{Hw_t}{L_t}\right)$	0.222	***	0.077
	$\Delta Ln\left(\frac{L_t}{P_t}\right)$	0.695	***	0.117
	$\Delta Ln(n_t)$	-0.182		0.082
	$\Delta Wgap_t$	-0.250		0.313
	constant	0.017		0.002

Note: *, **, *** is the10%, 5% and 1% significance level, respectively
Source: The result of estimating equation (E6)

Table E3 Investment Equation

		Coefficient		Standard Error
$\Delta Ln\left(\frac{I_t}{Y_t}\right)$	R^2= 0.7139			
(Investment/GDP)				
	$\Delta Ln\left(\frac{I_{t-1}}{Y_{t-1}}\right)$	-0.097		0.205
	$\Delta Ln\left(\frac{I_{t-2}}{Y_{t-2}}\right)$	-0.263		0.18
	$\Delta Ln\left(\frac{I_{t-3}}{Y_{t-3}}\right)$	-0.024		0.149
	$\Delta Ln\left(\frac{Y_t}{P_t}\right)$	4.722	***	1.288
	$\Delta Ln(h_t)$	-0.485		0.434
	$\Delta Ln\left(\frac{Hw_t}{L_t}\right)$	-0.913		0.593
	$\Delta Ln\left(\frac{L_t}{P_t}\right)$	-1.124		1.373
	$\Delta Ln(n_t)$	0.751		0.542
	$\Delta Wgap_t$	-0.261		2.377
	constant	-0.088	***	0.029

Note: *, **, *** is the10%, 5% and 1% significance level, respectively
Source: The result of estimating equation (E7)

Table E4 Hours of Work Equation

		Coefficient		Standard Error
$\Delta Ln\left(\frac{Hw_t}{L_t}\right)$ (Average hours of work)	R^2= 0.7102			
	$\Delta Ln\left(\frac{Hw_{t-1}}{L_{t-1}}\right)$	0.101		0.168
	$\Delta Ln\left(\frac{Hw_{t-2}}{L_{t-2}}\right)$	-0.27	**	0.115
	$\Delta Ln\left(\frac{Hw_{t-3}}{L_{t-3}}\right)$	-0.33	***	0.122
	$\Delta Ln\left(\frac{Y_t}{P_t}\right)$	1.449	***	0.429
	$\Delta Ln(h_t)$	-0.156		0.134
	$\Delta Ln\left(\frac{I_t}{Y_t}\right)$	-0.064		0.073
	$\Delta Ln\left(\frac{L_t}{P_t}\right)$	-0.886	**	0.441
	$\Delta Ln(n_t)$	0.457	**	0.21
	$\Delta Wgap_t$	-1.432	**	0.563
	constant	-0.027	***	0.008

Note: *, **, *** is the10%, 5% and 1% significance level, respectively
Source: The result of estimating equation (E8)

Table E5 Labour Participation Equation

		Coefficient		Standard Error
$\Delta Ln\left(\frac{L_t}{P_t}\right)$	R^2= 0.7593			
(Labour Participation)				
	$\Delta Ln\left(\frac{L_{t-1}}{P_{t-1}}\right)$	0.013		0.115
	$\Delta Ln\left(\frac{L_{t-2}}{P_{t-2}}\right)$	-0.132		0.102
	$\Delta Ln\left(\frac{L_{t-3}}{P_{t-3}}\right)$	-0.033		0.087
	$\Delta Ln\left(\frac{Y_t}{P_t}\right)$	1.029	***	0.175
	$\Delta Ln(h_t)$	0.013		0.076
	$\Delta Ln\left(\frac{Hw_t}{L_t}\right)$	-0.241	**	0.103
	$\Delta Ln\left(\frac{I_t}{Y_t}\right)$	-0.033		0.041
	$\Delta Ln(n_t)$	0.294	***	0.093
	$\Delta Wgap_t$	0.378		0.355
	constant	-0.016	***	0.004

Note: *, **, *** is the10%, 5% and 1% significance level, respectively
Source: The result of estimating equation (E9)

Table E6 Fertility Equation

		Coefficient		Standard Error
$\Delta Ln(n_t)$	R^2= 0.5038			
(Fertility)				
	$\Delta Ln(n_{t-1})$	0.259		0.188
	$\Delta Ln(n_{t-2})$	0.921	**	0.372
	$\Delta Ln(n_{t-3})$	-0.2		0.339
	$\Delta Ln(n_{t-4})$	-1.038	***	0.365
	$\Delta Ln\left(\frac{Y_t}{P_t}\right)$	-0.567		0.709
	$\Delta Ln(h_t)$	0.019		0.17
	$\Delta Ln\left(\frac{Hw_t}{L_t}\right)$	0.187		0.226
	$\Delta Ln\left(\frac{L_t}{P_t}\right)$	0.42		0.492
	$\Delta Ln\left(\frac{I_t}{Y_t}\right)$	0.231	*	0.123
	$\Delta Wgap_t$	0.993		0.747
	constant	0.008		0.013

Note: *, **, *** is the10%, 5% and 1% significance level, respectively
Source: The result of estimating equation (E10)

Table E7 Wage Gap Equation

		Coefficient		Standard Error
$\Delta Wgap_t$	R^2= 0.5516			
(Gender wage gap)				
	$\Delta Wgap_{t-1}$	-0.463	**	0.182
	$\Delta Wgap_{t-2}$	-0.087		0.179
	$\Delta Wgap_{t-3}$	0.352	*	0.188
	$\Delta Ln\left(\frac{Y_t}{P_t}\right)$	0.032		0.14
	$\Delta Ln(h_t)$	-0.006		0.042
	$\Delta Ln\left(\frac{Hw_t}{L_t}\right)$	-0.118	**	0.05
	$\Delta Ln\left(\frac{L_t}{P_t}\right)$	-0.034		0.121
	$\Delta Ln\left(\frac{I_t}{Y_t}\right)$	0.037		0.027
	$\Delta Ln(n_t)$	0.141	**	0.056
	constant	-0.001		0.003

Note: *, **, *** is the10%, 5% and 1% significance level, respectively
Source: The result of estimating equation (E11)

ESTIMATING THE IMPACT

Estimation Coefficients

The coefficients that have been produced from the regression system above allow us to calculate the impact of the wage gap on economic growth. For example the coefficient σ_0 in the main growth equation (i.e., equation (E4)) represents the impact of the wage gap on economic growth while holding other variables constant (referred to as partial impact). This impact can be estimated as follows:

$$\partial Ln\frac{Y_t}{P_t} = \sigma_0 \partial Wgap_t \Rightarrow \frac{\partial y_t}{y_t} = \sigma_0 \partial Wgap_t \Rightarrow \partial y_t = \sigma_0 \partial Wgap_t y_t \quad \text{where } \frac{Y_t}{P_t} = y_t \qquad \text{(E12)}$$

In the same way, the partial impact of other growth determinants on economic growth can be estimated.

The impact of investment: $\frac{\partial y_t}{y_t} = \beta_1 \partial Ln\left(\frac{I_t}{Y_t}\right)$ (E13),

The impact of fertility: $\frac{\partial y_t}{y_t} = \beta_3 \partial Ln(n_t)$ (E14),

The impact of labour participation: $\frac{\partial y_t}{y_t} = \beta_4 \partial Ln\left(\frac{L_t}{P_t}\right)$ (E15),

The impact of average hours of work: $\frac{\partial y_t}{y_t} = \beta_5 \partial Ln\left(\frac{Hw_t}{L_t}\right)$ (E16),

The equations (E7)-(E10) in the system provide further information about how much these growth determinants will change if there is change in the wage gap. This can estimated as,

The impact on investment: $\partial Ln\left(\frac{I_t}{Y_t}\right) = \sigma_1 \partial Wgap_t$ (E17),

The impact on fertility: $\partial Ln(n_t) = \sigma_2 \partial Wgap_t$ (E18),

The impact on labour participation: $\partial Ln\left(\frac{L_t}{P_t}\right) = \sigma_3 \partial Wgap_t$ (E19), The impact on average hours of work: $\partial Ln\left(\frac{Hw_t}{L_t}\right) = \sigma_4 \partial Wgap_t$ (E20).

By substituting the equations in (E17)-(E20) into the equations in (E13)-(E16), we will have the impact estimate of the wage gap through other growth determinants as follows:

The impact through investment: $\frac{\partial y_t}{y_t} = \beta_1 \sigma_1 \partial Wgap_t$ (E21)

The impact through fertility: $\frac{\partial y_t}{y_t} = \beta_3 \sigma_2 \partial Wgap_t$ (E22)

The impact through labour participation: $\frac{\partial y_t}{y_t} = \beta_1 \sigma_1 \partial Wgap_t$ (E23)

The impact through average hours of work: $\frac{\partial y_t}{y_t} = \beta_3 \sigma_2 \partial Wgap_t$ (E24)

Therefore, the total differential of growth on the wage gap provides the total impact of the wage gap on growth as

$$\frac{\Delta y_t}{y_t} = (\sigma_0 + \beta_1 \sigma_1 + \beta_3 \sigma_2 + \beta_4 \sigma_3 + \beta_5 \sigma_4) \Delta Wgap_t \quad \text{(E25)}$$

Table E8 shows the impact estimate of the gender wage gap on economic growth

Table E8 The impact coefficients of the gender wage gap on economic growth

	σ		β		**Impact Coefficient**
Wage gap → economic growth	-0.25				-0.250
Wage gap→investment→economic growth	-0.261		0.081	**	-0.021
Wage gap→fertility→economic growth	0.993		-0.182		-0.181
Wage gap→average hours of work→economic growth	-1.432	**	0.222	***	-0.318
Wage gap→labour participation→economic growth	0.378		0.695	***	0.263
Total effects					**-0.507**

Note: *, **, *** is the10%, 5% and 1% significance level, respectively
Source: The coefficient from tables E2 to E7

Estimating the Total Impact on GDP

The previous section has shown how to get the total impact coefficient of the wage gap on GDP. The final step is to calculate the total volume of GDP change (in dollars) which would result from a decrease or increase of the wage gap. To do so, we should rearrange equation (25) as

$$\Delta y_t = (\sigma_0 + \beta_1\sigma_1 + \beta_3\sigma_2 + \beta_4\sigma_3 + \beta_5\sigma_4)\Delta Wgap_t \times y_t \quad \text{(E26)}$$

and multiply both the right hand side and the left hand side of equation (26) by P_t to get

$$\Rightarrow \Delta Y_t = (\sigma_0 + \beta_1\sigma_1 + \beta_3\sigma_2 + \beta_4\sigma_3 + \beta_5\sigma_4)\Delta Wgap_t \times Y_t \quad \text{(E27).}$$

Knowing that the wage gap is currently 0.17 and the chain volume GDP is at $1,084,146.00 million we can calculate the impact of a change of one percentage point in the gender wage gap, as well as the cost of the whole 17 per cent of the gender wage gap as seen in Table E9.

Table E9. Estimates of gender wage gap impact on economic output

	Current GDP ($ millions)	Change in GDP economic growth (%)	Change in GDP ($ millions)
Gender wage gap increases by one percentage point	1,084,146.00	-0.50	-5,496.65
Total cost of wage gap (wage gap is eliminated by 17 percentage points)	1,084,146.00	8.50	93,443.06

The confidence level

While the coefficient estimates of our main growth equation have relatively low standard errors (which meant that coefficients were statistically significant and thus supported the direction expected by the theoretical growth model), we did however find high standard errors when estimating the impact of the gender wage gap directly on growth and most of its components. This was a drawback of our model. The lack of observations available to us is the likely cause of this. The high standard errors mean that our impact estimate has wide confidence intervals. Nevertheless, the coefficient estimate of the main growth equation (equation (E6)) is significant and shows a relationship between the gender wage gap and GDP per capita that is supported by growth theory. Therefore, despite data availability issues and high standard errors, our result is one which makes sense theoretically.

The question is, to what extent we expect that our estimation reflects true relationships in the economy. As can be seen in table E8, the impact of the gender wage gap on economic growth through an increase in average hours of work is the only impact that we found to be statistically significant at the 5% significance level. That means that we can be confident that there is 95% probability that this will actually be the outcome of a change in the gender wage gap.

In this interpretation, we have taken into account in calculating the total gap the impact of all channels – both significant and non-significant. However, what would the gap look like if we only took into account the impact of the one significant channel (hours of work)? As shown in Table E10, the impact is still in the same direction, and still very substantial – if we take only the significant channel into account, we still see that for an increase in the gender wage gap of one percentage point, GDP falls by 0.318 per cent. Thus we find that just this single channel accounts for 62.7 per cent of the total expected impact.

Another way of considering these issues is to think about what would happen to the apparent impact of the gender wage gap on GDP if we lowered our confidence level to 80%. Would we still see only one significant channel, and would the apparent impact change? We did this (see Table E10), and found that this added another significant channel to the model. With a confidence level of 80% we found that the gender wage gap affected both fertility and hours of work significantly. If we then use the impact of these two channels only to estimate the effects of the gender wage gap on GDP, we find that this effect (at -0.499) is very similar in magnitude to the total effect when all channels are used.

These results provide further support for our overall findings – they suggest that even when we adjust various aspects of the assumptions built into our modelling, the overall story about the impact of the gender wage gap on GDP per capita remains much the same.

Table E10. The impact of wage gap in different confidence level

Gender wage gap increases by one percentage point	Change in GDP economic growth (%)	Compared to total impact (%)	Change in GDP ($ millions)
Total Impact with all channels	-0.507		-5,496.65
Total Impact with only channels significant at 95% confidence level	-0.318	62.7	-3,447.58
Total Impact with only those channels significant at 80% confidence level	-0.499	98.4	-5,409.89

Source: Authors' calculations

MODEL LIMITATIONS

The discussion in this appendix has shown that our model is placed strongly within current approaches to growth modelling, has been rigorously tested, and is capable of providing an estimate of the impact of the gender wage gap on economic growth, and the indirect channels through which such an impact may be occurring. However, this model is subject to some weaknesses and limitations, and these are to some extent reflected in our results. We discuss these limitations here.

The main source of weakness in our model is the fact that it relies on a limited (20 year) data span. These data limitations affected the choices we made about model specification and also the strength of our estimation.

Because not all variables which we needed to include in the model were available on a quarterly basis, we had to use annual data, with 1990 as the starting year. Because we needed to use a first difference equation in addition to using lagged versions of our variables in order to achieve a robust estimate of the coefficient of growth determinant, we set 1990 as our base year (i.e., at time t), which means that we needed to use data from 1985 to construct the lag variables. Our preliminary data testing showed that the combination of instrumental variables can only give a good estimation of the channels if we include three lags to explain the behaviour of each channel, except for fertility which needed four lags.

Therefore, we only have 19 years of observations for the estimation given that we use 2008 data as the end of the observation period. As noted earlier, the short data span meant that our variables were less likely to be significantly stationary in the level form, and this contributed to our decisions about model specification.

One of the key decisions we had to make was to use the first difference form rather than the level form of our equations. The very high coefficient of determination (R^2) of 99 per cent in the level form is likely a result of the short data span. While our tests of stationarity did not conclusively reject the level form data, this high R^2, as noted earlier, raises substantial concerns about the possibility of a spurious regression. On the other hand, the coefficient of determination (R^2) of 77 per cent in the first difference form is the better choice since this means that the model can still explain the variation of the dependent variable with a reasonable range of error. However, the use of first differences also means that the result may not give a picture of long term equilibrium. With only 19 years of data this is an expected result, since the average rate of convergence in growth models is around 35 years.

Another decision we had to make was not to estimate the impact of the wage gap through the indirect channel of human capital. While the short data span available to us is mostly due to the length of human capital data available from the Education and Work Survey, we nevertheless considered the inclusion of human capital in our model as being of significant theoretical importance. Our estimation result bears this out, showing that the inclusion of human capital is almost as important as the inclusion of physical capital investment (see Table E2). This can be seen from the magnitude of the human capital coefficient in the main growth equation at 0.06 compared to 0.08 for physical capital investment. The fact that human capital is so important in predicting growth, and our inability to model this indirect channel in this study, means that our model in fact may underestimate the overall effects of the gender wage gap on GDP – that is, had we been able to examine the human capital channel, this might have further contributed to the impact of the gender wage gap on economic growth. This is particularly so as some of our preliminary modelling showed that the wage gap may have a negative impact on human capital accumulation. However, given the high standard error of the human capital coefficient in our final model, we do not expect that such underestimation would be of a large magnitude.

As noted above, the impacts of the wage gap on growth components within our model have relatively high standard errors. These relatively high standard errors are another weakness of our model, making it difficult to detect statistically significant relationships between channels of impact and economic growth. The only channel for which we have statistically significant results is for the average hours of work variable, which shows that a decrease in the wage gap will increase the average hours of work and thus eventually increase economic output. However, as discussed above, our findings related to the total impact of the gender wage gap on GDP per capita appear fairly robust.

The final limitation of the model is that we cannot separate the impact of the gender wage gap on growth components based on gender. The short data span is partly the reason, but most importantly the high collinearity of the time series data between the two genders of labour force participation made this decision necessary – that is, male labour force participation and female labour force participation were too highly correlated with each other to be able to effectively separate out their explanatory power in the model. If variables which are highly collinear are used in a model such as the one developed here, the model will be both inefficient and liable to give misleading results.