UCD CENTRE FOR ECONOMIC RESEARCH

WORKING PAPER SERIES

2012

The Relationship Between Low Birthweight and Socioeconomic Status in Ireland

David Madden, University College Dublin

WP12/14

April 2012

UCD SCHOOL OF ECONOMICS
UNIVERSITY COLLEGE DUBLIN
BELFIELD DUBLIN 4

The Relationship Between Low Birthweight and Socioeconomic Status in Ireland

David Madden

(University College Dublin)

March 2012

Abstract: There is now fairly substantial evidence of a socioeconomic gradient in low birthweight for developed countries. The standard summary statistic for this gradient is the concentration index. Using data from the recently published *Growing Up in Ireland* survey, this paper calculates this index for low birthweight arising from preterm and intra-uterine-growth-retardation. It also carries out a decomposition of this index for the different sources of low birthweight and finds that income inequality appears to be less important for the case of preterm births, while fathers education and local environmental conditions appear to be more relevant for IUGR. The application of the standard Blinder-Oaxaca decomposition also indicates that the socioeconomic gradient for all sources of birthweight appear to arise owing to different characteristics of rich and poor, and not because the return to characteristics differ between rich and poor.

Keywords: Low birthweight, Concentration Index, decomposition.

JEL Codes: I10, I14.

Corresponding Author: David Madden,
School of Economics,
University College Dublin,
Belfield,
Dublin 4,
Ireland.

Phone: 353-1-7168396
Fax: 353-1-2830068
E-Mail: david.madden@ucd.ie

I thank the Irish Social Science Data Archive for making the Growing Up in Ireland data available. I would also like to thank Brian Nolan for helpful discussion. The usual disclaimer applies.

The Relationship Between Low Birthweight and Socioeconomic Status in Ireland

1. Introduction

There is now fairly substantial evidence of a socioeconomic gradient in low birthweight (LBW) for developed countries (see Kramer et al, 2000 and for evidence for Ireland see McAvoy et al, 2006and Niedhammer et al, 2011). The incidence of LBW (weighing less than 2500 grams) tends to fall as socioeconomic status increases and the phenomenon is observed for a variety of measures of socioeconomic status (such as income, education and employment status).

Low birthweight is of concern for a number of reasons. It is associated with fetal and infant mortality, as well as with short and long-term morbidity. In addition, there is fairly extensive evidence that LBW is also regarded as a risk factor for a number of health and non-health outcomes in later life. See for example, Almond and Currie (2011a, 2011b), Black et al (2007) and Currie (2009, 2011). For evidence for Ireland see Delaney et al (2011) and McGovern (2011).

One of the principal measures for summarising the link between socioeconomic status and a given health outcome is the concentration index. Curiously, despite the relatively abundant literature detailing the link between LBW and socioeconomic status, there is virtually no calculation of the concentration index for this key outcome. This paper attempts to fill this gap by calculating the concentration index for LBW for Ireland for a representative sample of infants. In addition the decomposition of the concentration index can provide valuable insights into the factors lying behind the socioeconomic gradient.

In the next section of this paper we outline how to calculate the concentration index and also how it may be decomposed. In section 3 we describe the data and variables used, while in section 4 we present results for the concentration index for LBW in Ireland. Section 5 offers concluding comments.

2. The Concentration Index

Suppose we have a health variable, h, where h_i is the value of that variable for individual i. Then if r_i is the fractional rank of individual i in the income distribution (or whatever measure of household resources is being used), the concentration index is

$$C = \frac{2 * \mathrm{cov}(h_i, r_i)}{\mu_h}$$

where μ_h is the mean value of the health variable (Kakwani et al, 1997). C can take on a value from -1 to +1, where a negative (positive) value indicates that the health variable is concentrated among the relatively poor (rich). Since LBW can be regarded as both a measure and predictor of ill-health, a negative value of C will indicate a situation favouring the better-off and so could be regarded as pro-rich inequality.

One attractive property of the concentration index is that it is possible to decompose C into inequalities and elasticities of health determinants. If the vector $\boldsymbol{X}$ refers to those variables influencing h, then if we assume that the health variable can be described by a linear regression of the form

$$h_i = \alpha + \beta_k X_{ki} + \varepsilon_i$$

then C can be written as

$$C = \sum_k \left(\frac{\beta_k \bar{x}_k}{\mu_h} \right) C_k + \frac{GC_\varepsilon}{\mu_h}$$

where the index k refers to the regressors in the equation, C_k is the concentration index for each of the individual regressors, β_k is the coefficient for each health determinant and $\bar{x}_k$ is the mean value of each individual regressor. GC_ε is the generalised C for the residual from the regression.

The analysis above refers to the situation where the health variable is continuous. In the case of the incidence of LBW h_i is a binary variable which takes on values of 0 or 1. In this case a normalisation must be applied to the concentration index (since the bounds would not be -1 and +1). Wagstaff (2005) suggested a normalisation of $C_n = C/(1-\mu_h)$. In a recent contribution Erregeyers (2009a) suggested that the appropriate normalisation be $C_E = 4\mu_h C = 4\mu_h(1-\mu_h)C_n$. The subsequent debate (see Wagstaff, 2009 and Erregeyers, 2009b) indicates that the issue is not quite resolved yet. In our analysis here we will apply the Erregeyers normalisation to the concentration index and its decomposition.

It could be asked, why not simply calculate the concentration index for the total distribution of birthweight as opposed to just focussing on LBW? The reason we do not do this is because, from a public policy point of view, we are not concerned with how the distribution of birthweight varies with household resources above the critical threshold of 2500 grams. While the extent to which birthweight above 2500 grams varies with household resources may be of interest in its own right, we argue that it is not of relevance in the context of the socioeconomic gradient of LBW, presuming we accept the conventional threshold of 2500 grams.

3. Data and Variables Used

Our data comes from the *Growing Up in Ireland (GUI)* survey, 9 month old infant cohort (for a summary guide to this survey see Quail et al, 2011). The 9 month cohort comprised 11134 children born between 1st December 2007 and 30th June 2008. The sampling frame was drawn from the Child Benefit Register. Child Benefit is a payment made with respect to all children aged 16 years or under, and has many features which render it an ideal sampling frame for this exercise (see Quail et al, 2011, for details). We also use the sampling weights provided to further ensure that the sample is representative.

For the vast majority of subjects in the sample (over 99%), the responses were provided by the biological mother. In this study we drop those subjects where the answer was not provided by the biological mother and we also drop non-singleton births. The principal dependent variable used in this study is the birthweight of the child. Birthweight is recorded in the survey in intervals of 100 grams and there is data censoring at both the top and bottom of the distribution. All birthweights in excess of 4600 grams are listed as 4600. Meanwhile all birthweights below 1499 grams are listed as 1499. In addition birthweights in the 1500-2499 interval are simply listed as 2499. Given this censoring of the data it seems preferable to analyse LBW as a binary rather than continuous variable.

As pointed out by Kramer et al (2000), LBW is in some respects an unsatisfactory outcome for epidemiological analysis, since birthweight may be determined by both duration of gestation and by the rate of fetal growth. Thus LBW may occur either because an infant is born too early (a preterm birth) or because it is small for his/her gestational age (this can be regarded as a case of intra uterine growth restriction, IUGR). This distinction is of particular importance since evidence suggests that the determinants of gestational duration (and hence the issue of preterm) may be quite different from those of IUGR (see Kramer et al, 2000 and the references therein). This in turn may imply a different socioeconomic gradient for preterm compared to IUGR. It may also have consequences for the decomposition of the concentration index outlined in section 2, as the relative contribution of each factor to the concentration index may differ between overall LBW, preterm and IUGR. A detailed description of all covariates is provided in the appendix.

We distinguish between preterm and cases of IUGR in the following way. In addition to answering questions about birthweight, subjects are also asked after how many weeks of pregnancy the baby was born. We define all those answering less than 37 weeks as preterm. IUGR cases are then defined as those who are LBW but *not* preterm. For the sake of comparison we carry out all analysis for overall LBW, preterm and IUGR (we also carried out analysis using a lower cut-off of 32 weeks, results available on request).

The particular measure of household resources we use to calculate the concentration index is equivalised net income. Net income is the response to a question asking for the net income of all household members. It is defined as total gross household income less statutory deductions of income tax and social insurance contributions. It is then equivalised by dividing by the square root of household size.

Of course, we would expect LBW to also be correlated with other measures of socioeconomic status such as education or class. However as these are not cardinal variables it is not possible to calculate concentration indices with respect to these variables. We also include education of both parents as explanatory variables in the decomposition of the concentration curve.

4. Results

Before providing results regarding the concentration curve and the contribution of various covariates, we first present some summary statistics for overall birthweight (bearing in mind that the data is censored at either end), low birthweight, preterm and IUGR. In tables 1-4 we present the data for all observations, as well as by education, income quintile and social class. We divide education into five categories: lower secondary (i.e. left school before age 17/18), Leaving Certificate (completed secondary school education), Diploma/Cert (obtained qualification after secondary school but did not go on to third level education), third level education and postgraduate education. We divide social class into four categories, according to those provided in the survey. These are Professional/Managerial, Non-Manual and

Skilled Manual, Semi-Skilled and Unskilled Manual and then a group listed as “never worked at all, no class”.

The socioeconomic gradient is evident in pretty much all cases, but it is more pronounced in some cases compared to others. In particular the prevalence of LBW and IUGR is concentrated amongst the “lowest” of the groups. For example, in the case of IUGR by social class there is little evidence of a gradient for three of the classes (with incidence in a narrow range between 1.8% and 2%) and then a jump to 4% for the never worked/no class group. This differential gradient by socioeconomic class and by measure of LBW suggests a significant role for confounding factors and we examine these when decomposing the concentration index.

In the results which follow we calculate the concentration index for LBW, preterm and IUGR in table 5 and we also provide regression-based decompositions of these indices in tables 6-8. The choice of right-hand side variables is influenced by the review by Kramer et al (2000). The variables chosen are: age of mother and age squared (to allow for a non-linear effect), gender of child (male children are typically heavier at birthweight), education of both parents, working status of mother, mother’s smoking and drinking, body mass index (BMI) of mother and BMI squared, log of disposable household income, two measures reflecting local environmental conditions, general health of mother plus information on some specific conditions (mental and physical) and ethnicity of mother (Irish or non-Irish). Full details of all explanatory variables are provided in the appendix.

Before discussing these there is a specific data issue which must be addressed. In calculating the concentration indices and in providing the decomposition, we must bear in mind that some observations are missing and in particular it is possible that such observations may not be missing at random. Compared to our base sample size of 10969 there are over 800 observations where income data is missing. When we take into account the right-hand side variables used for the decomposition we lose approximately another 670 observations. There are a variety of approaches one can take when faced with missing data (see Horton and Kleinman, 2007). One possibility is to adopt the "complete-case" approach, whereby we only use those observations with no missing values for any variable. This gives a sample size of 9469. The concentration indices for LBW, preterm and IUGR calculated using this sample are -0.434, -0.410 and -0.516 respectively.

However, the complete-case approach can be regarded as overly *ad hoc* as a means to address missing data. Since the variable with the greatest number of missing observations is income, and since income is critical in terms of calculating concentration indices it seems worthwhile to try to deal explicitly with the missing values for this variable at least. One possible way of doing this is to estimate an income function, by regressing equivalised income on age and education. We then substitute the fitted values from this regression for those observations where income is missing. This provides a sample size of 10196. The concentration indices calculated for this sample are -0.461, -0.411 and -0.511. This is known as the regression prediction or conditional mean imputation approach (where we have applied it to income only). While there are some differences from the indices calculated using the complete case approach, they are of a similar order of magnitude suggesting that the missing observations do not unduly bias the results. Note that missing observations for variables other than income are only relevant when carrying out the decomposition. We could calculate concentration indices only, using the conditional mean approach which would give a full sample of 10969 (which gives values of -0.476, -0.470 and -0.461). However in this case, we would not be able to carry out the decomposition. Thus the approach we adopt could be characterised as a combination of the complete case approach with conditional mean imputation applied to the income variable.

Turning now to the actual results, we must bear in mind that this measure summarises the gradient with respect to equivalised disposable income only (since education and class are non-cardinal variables it is not possible to calculate concentration indices with respect to these variables). Table 5 provides concentration indices for LBW, preterm and IUGR. All indices are negative and in the -0.4 to -0.5 region, indicating that incidence of the phenomenon in question is substantially concentrated amongst the less well-off. In all cases the p-values are less than 0.05. It is also comforting to note that the values of the concentration index do not appear to be unduly sensitive to the way in which we treat the missing observations.

Before analysing the decomposition of this index, it is useful to try to get an intuitive sense of what these figures actually mean. The sign of the concentration index indicates the direction of any relationship between the health variable and rank in the distribution of whatever measure of household resources is being used. The magnitude reflects both the strength of the relationship and the degree of variability in the health variable. In addition, Koolman and van Doorslaer (2004) have shown that multiplying the value of the index by 75 gives the percentage of the ill-health variable which, in the case of a negative index, would need to be redistributed from the poorer half to the richer half of the population to arrive at a distribution with an index of zero. Thus if 30-35% of LBW could be transferred from the poorer 50% of the population to the richer 50% of the population, the concentration index would be zero and there would be no socio-economic gradient in LBW. For the sake of comparison it is also worth noting that the concentration indices reported here are higher than those reported by the World Bank for under-five mortality in Vietnam (Wagstaff et al, 2007).

Turning now to the decomposition results in tables 6-8, perhaps the first issue to note is that the right-hand side variables explain a considerable portion of the income-related inequality, ranging from over 100% in the case of preterm to about 75% in the case of IUGR. Recall that in order for any variable to contribute to income-related inequality it must (a) influence the measure of LBW (which can be examined via the elasticities column) and also (b) itself be related to the distribution of income (which can be examined via the value of its own concentration index). Note also that variables can contribute both positively and negatively to the overall concentration index. In the discussion which follows it is important to bear in mind that the overall concentration index for all our measures of LBW is negative i.e. it is more concentrated amongst the less well-off. The column labelled "contribution %" shows the percentage contribution of each factor to overall income related inequality in LBW. A positive value indicates that this factor operated to bring about the concentration of LBW amongst the less well-off. A negative value indicates that the factor operated in the opposite direction i.e. on its own, this factor would have led to LBW being more concentrated amongst the *better-off*.

Bearing this in mind we can see that in table 6, the most important factors with respect to overall LBW, in absolute terms, were age, working status, smoking, drinking, income and overall health. We combine the contributions of age and age squared to arrive at a contribution of -22%. The regression results show that incidence of LBW first of all declines with age, bottoms out at about 27 and then rises with age. The concentration curve for age is positive, though relatively low (i.e. older mothers are better off on average than younger mothers, but not by an awful lot). Combining all these factors together, the impact of age, on its own, is to increase the incidence of LBW amongst the better-off.

Working full-time and part-time also make substantial negative contributions to the overall concentration index, to the tune of about -27%. Bonzini et al in a review of the evidence concerning working and a number of adverse birth outcomes (including preterm delivery and LBW) suggest that for preterm delivery there was extensive evidence of an effect of certain occupational exposures viz. long working hours, shift-work, lifting, standing and heavy workload. No evidence was found for LBW and pre-eclampsia. The findings in tables 6-8 partially confirm this in that greater effects are observed for overall LBW and preterm than for IUGR. The *GUI* data has information on whether or not a mother worked during pregnancy, how many hours per week she worked, and how long before birth she gave up work. Of these three variables, the effect of working *per se* is the most dominant. Once working/non-working is included then there is no separate effect for the number of hours worked nor the number of weeks before birth when work stopped. Unfortunately we do not have information on whether the work involved heavy lifting or standing for long periods. Given that working is associated with a greater likelihood of LBW, and given that it is also associated with higher incomes, the combined effect is that working tends to increase the concentration of LBW amongst the better-off and hence makes a negative contribution to the concentration index.

Smoking makes one of the largest *positive* contributions to the concentration index. In this case the chain of causality is clear. Smoking is associated with LBW and is heavily concentrated amongst lower income mothers (see Kramer et al, 2000) and the combination of these factors imply that it makes a contribution of about 25% to the concentration index.

What is perhaps slightly more surprising is that drinking alcohol also contributes positively (just over 15%) to the concentration index. Drinking is concentrated amongst better-off mothers and LBW has an elasticity of -0.173 with respect to drinking. As outlined in the appendix the drinking variable used is ordinal with seven categories, ranging from never drinks alcohol to drinking alcohol every day. Unfortunately, our data does not provide information on drinking behaviour *during pregnancy*, but rather drinking behaviour at time of interview. This is unfortunate since finding a negative effect of alcohol consumption on the probability of LBW is unusual (see Kramer et al, 2000) and it would be comforting to know that our measure of alcohol actually referred to the pregnancy period. However it is worth noting that the negative relationship between alcohol consumption and LBW is also found in the

9-year cohort of the *GUI* survey and in that instance alcohol refers to alcohol consumption during pregnancy.

The biggest single contributor to the negative concentration index for LBW is income itself (to be more precise the log of equivalised income). While it may seem strange that income should be a right-hand variable in this decomposition, perhaps the easiest way to view this is to consider what the gradient would be if everyone had the same income. In this case there clearly could be no gradient, in the sense of a relationship between LBW and income, since everyone would have the same income. Correspondingly, if LBW is negatively related to income, then any factor which leads to a widening of income inequality will increase the (negative) value of the concentration index.

The final variable which makes a substantive contribution to the negative concentration index is overall health. As in the case of smoking, the chain of causation appears reasonably straightforward. Health is measured here as self-assessed health ranging from “excellent” to “poor” with higher values corresponding to poorer health. Thus poorer overall health increases the probability of LBW and since there is a well-documented gradient between health and income (e.g. Kakwani et al 1997) this translates into a contribution of over 10% to the concentration index.

Tables 7-8 show a similar decomposition for preterm and IUGR. Note that since overall LBW is comprised of preterm and IUGR, then the concentration index for LBW will be a rough weighted average of that for preterm and IUGR. The breakdown for preterm is quite similar to that for overall LBW, with one or two exceptions. First of all, the contribution of smoking to the concentration index is less, as the elasticity with respect to smoking is only about half as big in magnitude. The contribution of drinking is much less in the case of preterm reflecting once again a lower elasticity than in the case of overall LBW. Income makes a substantially greater contribution in the case of preterm, with a considerably higher elasticity.

Table 8 shows the breakdown for IUGR and here we observe greater differences with respect to overall LBW. First of all, the residual element is considerably larger here, with nearly one-quarter of the total concentration index unexplained. Looking at the individual variables, fathers' education exercises a greater role, with the sum of these variables contributing nearly 30% to the index. Higher fathers' education leads to lower rates of IUGR and higher fathers' education is also associated with higher income. There is a much diminished role for income *per se*. Compared to a contribution of 58% for overall LBW it now contributes only -6.8% i.e. IUGR is positively related to income, though the effect is small. There is also a greater role for local environmental and health variables.

The approach we have adopted so far has assumed that socio-economic inequality in the various forms of LBW arises from difference in characteristics only. An alternative way to express this is that in the regression based decomposition, it is assumed that the effect of each covariate on LBW is common across all levels of income. However, this may not be the case. For example, it is possible that the effect of ill-health on LBW may differ by income level. This could arise because richer mothers may have the resources to partially offset the effects of ill-health. Thus at a more general level it is possible that differing LBW by income level may arise owing to different characteristics and/or different *returns* to these characteristics (i.e. the returns differ by income).

This form of decomposition is the well-known Blinder-Oaxaca decomposition (see Blinder, 1973, Oaxaca, 1973) whereby in this case it is necessary to partition the sample into two groups by income. We partition them according to median equivalised income into the “rich” and “poor” (we also experiment with a partition at the 25th percentile, the “very poor” and the rest). The overall gap in LBW between the two groups will then be decomposed into a part arising from differences in characteristics and differences in returns to characteristics (i.e. the impact of these characteristics on LBW). Decompositions of this type will be sensitive to whichever group’s LBW is assumed to be the “norm”. This is a standard path-dependence (or index number) issue and in our application here there does not seem to be a compelling case to regard either the “rich” or “poor” as the reference or norm and so we adopt the procedure of Neumark (1995) who suggests using the vector of returns obtained from the pooled sample of “rich” and “poor” (see Oaxaca and Ransom, 1994, for a more detailed discussion of this issue). The results are presented in table 9 and indicate that for all forms of LBW and for both the poor-rich and very poor-rich partitions, around 90% (in some cases more) of the gap is explained by differences in characteristics, rather than returns to characteristics.

Thus overall these results suggest that the socioeconomic gradient in LBW arises primarily from differences in characteristics between lower and higher income groups, and not because there are different returns to characteristics at different levels of income. Applying the B-O decomposition along the lines requires some form of relatively crude partition, but given that the results still hold broadly when the very poor-rich partition is used, they seem fairly robust.

5. Conclusion

This paper has provided a new perspective on the issue of socio-economic inequalities for low birthweight in Ireland by calculating concentration indices for a representative sample of Irish women. A further innovation of the paper is that the indices are calculated for low birthweight arising from two different sources, preterm and intra-uterine-growth-retardation. For all forms of low birthweight the calculated concentration indices are in the -0.4 to -0.5 region, indicating fairly substantial concentration amongst the less well-off. The decompositions of the concentration index for the different sources of low birthweight show some uniformity, but there are also some differences. For example, income inequality appears to be less important for the case of preterm births, while fathers education and local environmental conditions appear to be more relevant for IUGR. Finally, the application of the standard Blinder-Oaxaca decomposition indicates that the socioeconomic gradient for all sources of birthweight appear to arise owing to different characteristics of rich and poor, and not because the return to characteristics differ between rich and poor.

References

Almond, D and J. Currie, 2011a. "Human Capital Development Before Age Five". *Handbook of Labour Economics*, 4, 1315-1486.

Almond, D and J. Currie, 2011b. "Killing Me Softly: The Fetal Origins Hypothesis". *Journal of Economic Perspectives*, 25, 153-172.

Black, S.E., P. Devereux and K. Salvanes, 2007. "From the Cradle to the Labour Market? The Effect of Birthweight on Adult Outcomes". *Quarterly Journal of Economics*, 122, 409-439.

Blinder, A. 1973. "Wage Discrimination: Reduced Form and Structural Estimates." *Journal of Human Resources*. 8: 436-455.

Bonzini, M., D. Coggon and K. Palmer, 2007. "Risk of prematurity, low birthweight and pre-eclampsia in relation to working hours and physical activity: a systematic review", *Occupational and Environmental Medicine*, 64, 228-243.

Currie, J., 2009. "Healthy, Wealthy and Wise: Socioeconomic Status, Poor Health in Childhood and Human Capital Development", *Journal of Economic Literature*, 47, 87-122.

Currie, J., 2011. "Inequality at Birth: Some Causes and Consequences", *mimeo*, Princeton University.

Delaney, L., M. McGovern and J. Smith, 2011. "From Angela's Ashes to the Celtic Tiger: Early Life Conditions and Adult Health in Ireland", *Journal of Health Economics*, 30, 1-10.

Erreygers, G., 2009a. "Correcting the Concentration Index", *Journal of Health Economics*, 28, 505-515.

Erreygers, G., 2009b. "Correcting the Concentration Index: A reply to Wagstaff", *Journal of Health Economics*, 28, 521-524.

Horton, N., and K. Kleinman, 2007. "Much ado about nothing: A comparison of missing data methods and software to fit incomplete data regression models", *American Statistician*, Vol. 61, 79-90.

Kakwani, N., A. Wagstaff and E. Van Doorslaer. 1997. "Socioeconomic Inequalities in Health: Measurement, Computation and Statistical Inference", *Journal of Econometrics*, 77, 87-103.

Koolman, X. And E. Van Doorslaer. 2004. "On the Interpretation of a Concentration Index of Inequality", *Health Economics,* 13, 649-656.

Kramer, M.S., L. Seguin, J. Lydon and L. Goulet, 2000. “Socio-economic disparities in pregnancy outcome: why do the poor fare so poorly?”, *Paediatric and Perinatal Epidemiology*, 14, 194-210.

McAvoy, H., J. Sturley, S. Burke and K. Balanda. 2006. *Unequal at Birth: Inequalities in the Occurrence of Low Birthweight Babies in Ireland.* Institute of Public Health in Ireland.

McGovern, M., 2011. “Still Unequal at Birth: Socioeconomic Status and Birthweight in Ireland”, *mimeo*.

Neumark, D. 1995. Employers’ Discriminatory behaviour and the Estimation of Wage Discrimination. *Journal of Human Resources, 23,* 279-295.

Niedhammer, I., C. Murrin, D. O’Mahony, S. Daly, J. Morrison and C. Kelleher. 2011. “Explanations for Social Inequalities in Preterm Delivery in the Propspective Lifeways Cohort in the Republic of Ireland”, *European Journal of Public Health*, 1-6.

Oaxaca, R. 1973. “Male-Female Wage Differentials in Urban Labour Markets.” *International Economic Review*. 14: 693-709.

Oaxaca, R., and Ransom, M. 1994. On Discrimination and the Decomposition of Wage Differentials. *Journal of Econometrics, 61*, 5-21.

Quail, A., J. Williams, C. McCrory, A. Murray and M. Thornton. 2011. *A Summary Guide to Wave 1 of the Infant Cohort (at 9 months) of Growing Up in Ireland.* Economic and Social Research Institute, Dublin.

Radloff, L.S. (1977) “The CES-D scale: A self report depression scale for research in the general population”. *Applied Psychological Measurement* 1: 385-401.

Wagstaff, A., 2005. “The bounds of the concentration index when the variable of interest is binary, with an application to immunisation inequality”, *Health Economics*, 14, 429-432.

Wagstaff, A., O. O Donnell, E. van Doorslaer and M. Lindelow. 2007. *Analysing Health Equity Using Household Survey Data*. World Bank.

Wagstaff, A., 2009. “Correcting the Concentration Index: A Comment”, *Journal of Health Economics*, 28, 516-520.

Table 1: Summary Statistics, Birthweight

	Mean	**St. Dev.**
Birthweight (grams)	3476.1	538.3
Low Birthweight (%)	5.79	
Preterm (%)	6.56	
IUGR (%)	2.10	

Table 2: Birthweight by Mothers Education

	Low Sec	**Leaving**	**DipCert**	**3rd Lev**	**Postgrad**
Bwgt (gms)	3389.3	3461.2	3500.0	3517.1	3543.5
LBW (%)	7.87	5.98	5.01	5.40	4.12
Preterm (%)	7.79	7.16	5.88	5.43	5.89
IUGR (%)	3.38	1.76	2.36	1.94	1.03

Table 3: Birthweight by Income Quintile

	1	**2**	**3**	**4**	**5**
Bwgt (gms)	3390.9	3448.2	3506.9	3516.4	3520.9
LBW (%)	8.23	5.98	5.07	5.15	4.48
Preterm (%)	8.84	7.20	5.38	5.83	5.48
IUGR (%)	3.14	1.99	1.90	2.06	1.44

Table 4: Birthweight by Social Class

	Professional, Managerial	**Non-manual, skilled manual**	**Semi-skilled, unskilled**	**No class**
Bwgt (gms)	3506.8	3478.9	3431.0	3342.8
LBW (%)	5.18	5.49	6.89	9.35
Preterm (%)	5.72	6.68	8.18	8.71
IUGR (%)	1.87	2.01	2.01	3.94

Table 5: Concentration Indices for various measures of LBW (robust standard error in brackets)

	N=10196
Low Birthweight	-0.461 *(0.120)*
Preterm	-0.411 *(0.103)*
IUGR	-0.512 *(0.234)*

Table 6: Decomposition of Concentration Indices, Low Birthweight

	Elasticities	Conc Index	Contrib.	Contrib.(%)
Age	-6.166	0.138	-0.851	184.2
Age^2	3.864	0.246	0.952	-206.1
Male	-0.016	0.001	-9.6E-06	0
Leaving	-0.026	-0.901	0.023	-5.1
Dip/Cert	-0.028	0.337	-0.009	2
3rd Lev	-0.008	1.101	-0.009	2
Postgrad	-0.035	1.564	-0.055	11.9
FLeaving	-0.034	-0.103	0.004	-0.8
FDip/Cert	-0.038	0.642	-0.024	5.3
F 3rd Lev	-0.029	1.272	-0.036	7.9
FPostgrad	0.001	1.708	0.002	-0.5
FUnreported	-0.036	-1.533	0.055	-11.9
F/T Work	0.140	0.705	0.099	-21.4
P/T Work	-0.071	-0.369	0.026	-5.6
Smoker	0.090	-1.299	-0.117	25.3
Occ Smoke	-0.016	-0.222	0.003	-0.7
Drink	-0.173	0.405	-0.070	15.2
BMI	-8.258	-0.050	0.411	-89
BMI^2	3.622	-0.113	-0.411	88.9
Log Y	-2.044	0.130	-0.265	57.4
Local 1	-0.453	0.064	-0.029	6.3
Local 2	0.470	-0.062	-0.029	6.4
Health	0.291	-0.166	-0.048	10.5
Ur Infect	0.014	-0.475	-0.007	1.5
Blood Press.	0.032	-0.017	-0.001	0.1
Pre-eclampsia	0.113	-0.235	-0.026	5.7
Depression	-0.027	-0.502	0.0134	-2.9
Stress	-0.046	-0.060	0.003	-0.6
Irish	-0.039	0.217	-0.008	1.8
Residual			**-0.056**	**12.2**
Total			**-0.462**	

Table 7: Decomposition of Concentration Indices, Pre-term

	Elasticities	Conc Index	Contrib.	Contrib.(%)
Age	-6.655	0.138	-0.918	223.4
Age^2	3.975	0.246	0.979	-238.3
Male	0.032	0.000	0	0
Leaving	0.009	-0.901	-0.008	1.9
Dip/Cert	-0.011	0.337	-0.004	1
3rd Lev	-0.012	1.101	-0.013	3.2
Postgrad	0.009	1.563	0.014	-3.4
FLeaving	0.036	-0.104	-0.004	1
FDip/Cert	-0.010	0.642	-0.006	1.5
F 3rd Lev	-0.0162	1.272	-0.021	5.1
FPostgrad	-0.004	1.708	-0.006	1.5
FUnreported	-0.031	-1.533	0.048	-11.7
F/T Work	0.127	0.706	0.09	-21.9
P/T Work	-0.021	-0.369	0.008	-1.9
Smoker	0.048	-1.299	-0.062	15.1
Occ Smoke	-0.004	-0.222	0.001	-0.2
Drink	-0.025	0.405	-0.01	2.4
BMI	-4.194	-0.050	0.209	-50.9
BMI^2	2.114	-0.113	-0.24	58.4
Log Y	-2.797	0.130	-0.363	88.3
Local 1	-0.077	0.064	-0.005	1.2
Local 2	0.358	-0.063	-0.022	5.4
Health	0.060	-0.166	-0.01	2.4
Ur Infect	0.003	-0.475	-0.002	0.5
Blood Press.	0.008	-0.017	0	0
Pre-eclampsia	0.092	-0.235	-0.022	5.4
Depression	0.054	-0.502	-0.027	6.6
Stress	-0.034	-0.060	0.002	-0.5
Irish	-0.206	0.217	-0.045	11
Residual			**0.026**	**-6.3**
Total			**-0.411**	

Table 8: Decomposition of Concentration Indices, IUGR

	Elasticities	Conc Index	Contrib.	Contrib.(%)
Age	-2.866	0.138	-0.395	77.1
Age^2	1.878	0.246	0.462	-90.2
Male	-0.043	0.000	0	0
Leaving	-0.139	-0.901	0.125	-24.4
Dip/Cert	0.002	0.337	0.001	-0.2
3^{rd} Lev	-0.046	1.101	-0.05	9.8
Postgrad	-0.089	1.563	-0.139	27.1
FLeaving	-0.102	-0.103	0.011	-2.1
FDip/Cert	-0.088	0.641	-0.056	10.9
F 3^{rd} Lev	-0.069	1.272	-0.088	17.2
FPostgrad	-0.002	1.707	-0.003	0.6
FUnreported	0.010	-1.532	-0.016	3.1
F/T Work	0.100	0.706	0.071	-13.9
P/T Work	-0.041	-0.369	0.015	-2.9
Smoker	0.130	-1.299	-0.168	32.8
Occ Smoke	-0.033	-0.222	0.007	-1.4
Drink	-0.368	0.405	-0.149	29.1
BMI	-7.715	-0.050	0.384	-75
BMI^2	2.782	-0.113	-0.316	61.7
Log Y	0.273	0.130	0.035	-6.8
Local 1	-0.748	0.064	-0.048	9.4
Local 2	0.050	-0.063	-0.031	6.1
Health	0.403	-0.166	-0.067	13.1
Ur Infect	0.031	-0.475	-0.015	2.9
Blood Press.	0.055	-0.017	-0.001	0.2
Pre-eclampsia	0.065	-0.235	-0.015	2.9
Depression	-0.115	-0.502	0.058	-11.3
Stress	-0.270	-0.060	0.016	-3.1
Irish	-0.082	0.217	-0.018	3.5
Residual			**-0.122**	**23.8**
Total			**-0.512**	

Table 9: B-O Decomposition of LBW gap (in %)

	Partition by median		Partition by 25^{th} percentile	
	Charact.	**Returns**	**Charact.**	**Returns**
LBW	89.4	10.6	96.6	3.4
Pre-term	98.4	1.6	112.2	-12.2
IUGR	103.7	-3.7	88.3	11.7

Table A1: Low birth-weight regressions (N=10196), Robust standard errors in parenthesis

	LBW	Preterm	IUGR
Age	-0.013	-0.015	-0.002
	(0.007)**	(0.007)**	(0.004)
Age^2	0.000	0.000	0.000
	(0.000)**	(0.000)**	(0.000)
male	-0.002	0.004	-0.002
	(0.006)	(0.006)	(0.004)
Leaving	-0.007	0.001	-0.009
	(0.010)	(0.010)	(0.007)
Dipcert	-0.011	-0.005	-0.000
	(0.011)	(0.011)	(0.008)
3rd Level	-0.006	-0.006	-0.005
	(0.012)	(0.011)	(0.008)
Postgrad	-0.020	0.001	-0.015
	(0.013)	(0.013)	(0.008)*
Father Leaving	-0.008	0.009	-0.008
	(0.010)	(0.010)	(0.006)
Father Dipcert	-0.017	-0.005	-0.014
	(0.010)	(0.011)	(0.006)**
Father 3rd Level	-0.011	-0.007	-0.009
	(0.011)	(0.011)	(0.007)
Father Postgrad	0.001	-0.003	-0.000
	(0.013)	(0.013)	(0.008)
Father Education Unreported	-0.011	-0.010	0.001
	(0.011)	(0.011)	(0.007)
Working	0.027	0.020	0.006
	(0.006)***	(0.006)***	(0.004)*
Daily smoker	0.029	0.017	0.016
	(0.009)***	(0.009)*	(0.007)**
Occasional smoker	-0.012	-0.003	-0.009
	(0.009)	(0.011)	(0.005)*
Drinker	-0.007	-0.002	-0.005
	(0.002)***	(0.003)	(0.002)***
BMI	-0.019	-0.011	-0.006
	(0.006)***	(0.006)*	(0.003)**
BMI^2	0.000	0.000	0.000
	(0.000)***	(0.000)**	(0.000)*
Log Inc	-0.012	-0.018	0.001
	(0.007)*	(0.007)***	(0.004)
Local Conditions 1	0.002	0.000	0.001
	(0.001)	(0.001)	(0.001)
Local Conditions 2	0.003	0.003	0.001
	(0.001)***	(0.001)**	(0.001)
Health	0.009	0.003	0.004
	(0.004)**	(0.004)	(0.003)*
Urinfect	0.014	0.008	0.007
	(0.009)	(0.009)	(0.006)
Depression	-0.001	0.002	-0.001
	(0.001)	(0.001)	(0.001)*
Parental Stress	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Irish	-0.001	-0.014	-0.002
	(0.008)	(0.009)*	(0.005)
R-squared	0.018	0.011	0.013
* sig at 10%; ** sig at 5%;	*** sig at 1%		

List of Variables used for Decomposition

Education	Omitted category is no formal education, primary education and lower secondary (leaving school at or before 16 approx). Remaining categories are Upper Secondary (including technical and vocational qualifications), Diploma/Cert (i.e. non-degree third level), 3rd Level (including a professional qualification equivalent to a degree) and postgraduate (including postgrad cert/diploma, Masters, PhD).
Working (part-time and full-time)	Response to question: did you work full-time, part-time or not at all immediately before you became pregnant with study child?
Smoking	Response to question: do you currently smoke daily, occasionally or not at all?
Drinking	Constructed on basis of question: which of the following best describes how often you usually drink alcohol? Respones are coded 0-6 based on: never, less than once a month, 1-2 times a month, 1-2 times a week, 3-4 times a week, 5-6 times a week, every day.
Income	Response to question: if you added up all the income sources from all household members what would be the total household net income after deductions for income tax and PRSI only?
Local 1	Based on response to question about incidence of four factors (rubbish and litter, homes and gardens in bad condition, vandalism, public drunkenness and drug-taking). Responses are coded 1-4 based on: not at all common, not very common, fairly common, very common. Aggregate score is used.
Local 2	Based on response to degree of agreement with statements concerning safety to walk after dark, safety for children to play outside, safe parks and playgrounds, intend to continue living in the area, are settled and part of the community. Responses are coded 1-4 based on: strongly agree, agree, disagree, strongly disagree. Aggregate score is used.
Health	Based on response to question: in general, how would you say your current health is? Responses are coded 1-5 based on: excellent, very good, good, fair, poor. Actual value is used.
Urinary Infection, Blood Pressure, Pre-eclampsia	Coded 0-1 based on response to question: were there any of the following complications with the pregnancy
Depression	8-item short version of the Centre for Epidemiological Studies Depression Scale (see Radloff, 1977).
Stress	Aggregate of responses to a series of stress related questions. Responses are coded 1-5 based on strongly agree, agree, not sure, disagree, strongly disagree. See Quail et al (2011)
Irish	Based on response to question: are you a citizen of Ireland

UCD CENTRE FOR ECONOMIC RESEARCH – RECENT WORKING PAPERS

WP11/16 Robert Gillanders: 'The Effects of Foreign Aid in Sub-Saharan Africa' August 2011
WP11/17 Morgan Kelly: 'A Note on the Size Distribution of Irish Mortgages' August 2011
WP11/18 Vincent Hogan, Patrick Massey and Shane Massey: 'Late Conversion: The Impact of Professionalism on European Rugby Union' September 2011
WP11/19 Wen Fan: 'Estimating the Return to College in Britain Using Regression and Propensity Score Matching' September 2011
WP11/20 Ronald B Davies and Amélie Guillin: 'How Far Away is an Intangible? Services FDI and Distance' September 2011
WP11/21 Bruce Blonigen and Matthew T Cole: 'Optimal Tariffs with FDI: The Evidence' September 2011
WP11/22 Alan Fernihough: 'Simple Logit and Probit Marginal Effects in R' October 2011
WP11/23 Ronald B Davies and Krishna Chaitanya Vadlamannati: 'A Race to the Bottom in Labour Standards? An Empirical Investigation' November 2011
WP11/24 Wen Fan: 'School Tenure and Student Achievement' November 2011
WP11/25 Mark E McGovern: 'Still Unequal at Birth - Birth Weight, Socioeconomic Status and Outcomes at Age 9' November 2011
WP11/26 Robert Gillanders: 'The Mental Health Cost of Corruption: Evidence from Sub-Saharan Africa' November 2011
WP11/27 Brendan Walsh: 'Well-being and Economic Conditions in Ireland' December 2011
WP11/28 Cormac Ó Gráda: 'Fetal Origins, Childhood Development, and Famine: A Bibliography and Literature Review' December 2011
WP12/01 Mark E McGovern: 'A Practical Introduction to Stata' January 2012
WP12/02 Colm McCarthy: 'Ireland's European Crisis: Staying Solvent in the Eurozone' January 2012
WP12/03 Colm McCarthy: 'Improving the Eurosystem for Old and New Members' January 2012
WP12/04 Ivan Pastine and Tuvana Pastine: 'All-Pay Contests with Constraints' February 2012
WP12/05 David Madden: 'Methods for Studying Dominance and Inequality in Population Health' February 2012
WP12/06 Karl Whelan: 'ELA, Promissory Notes and All That: The Fiscal Costs of Anglo Irish Bank' February 2012
WP12/07 Olivier Bargain, Eliane El Badaoui, Prudence Kwenda, Eric Strobl and Frank Walsh: 'The Formal Sector Wage Premium and Firm Size for Self-employed Workers' March 2012
WP12/08 Brendan Walsh: 'The Influence of Macroeconomic Conditions and Institutional Quality on National Levels of Life Satisfaction' March 2012
WP12/09 Ronald B Davies and Rodolphe Desbordesz: 'Greenfield FDI and Skill Upgrading' March 2012
WP12/10 Morgan Kelly and Cormac Ó Gráda: 'Change Points and Temporal Dependence in Reconstructions of Annual Temperature: Did Europe Experience a Little Ice Age?' March 2012
WP12/11 Morgan Kelly and Cormac Ó Gráda: 'The Waning of the Little Ice Age' April 2012
WP12/12 Morgan Kelly and Cormac Ó Gráda: 'Agricultural Output, Calories and Living Standards in England before and during The Industrial Revolution' April 2012
WP12/13 Arnaud Chevalier and Orla Doyle: 'Schooling and Voter Turnout - Is there an American Exception?' April 2012

UCD Centre for Economic Research Email economics@ucd.ie