# A Detailed Decomposition for Limited Dependent Variable Models

Jörg Schwiebert*


## Abstract

In this paper, we consider a detailed decomposition method for limited dependent variable models. That means, we propose a method to decompose the differential in the (limited dependent) outcome variable between two groups into the contributions of the explanatory variables. We provide a theoretical derivation of the detailed decomposition and show how this decomposition can be estimated consistently. In contrast to decomposition approaches already presented in the literature, our method leads to a unique decomposition and accounts for the nonlinearity of the underlying econometric model in a rather intuitive way. Our results can be applied to the most common limited dependent variable models such as probit, logit and tobit models.


**Keywords:** Decomposition methods, detailed decomposition, explained differential, limited dependent variable models, probit, logit, tobit.
**JEL codes:** C40, J31, J70.

*Leibniz University Hannover, Institute of Labor Economics, Königsworther Platz 1, 30167 Hannover, Germany, Tel.: 0511/762-5657, E-mail: schwiebert@aoek.uni-hannover.de

# 1 Introduction

Decomposition methods in economics have been a nascent field of research over the last years. Recently, in the fourth volume of the *Handbook of Labor Economics* a full chapter has been devoted to this topic (Fortin et al., 2011).

In this paper, we consider a detailed decomposition method for limited dependent variable models, such as probit, logit and tobit models. In contrast to models which are linear in parameters and explanatory variables, a detailed decomposition in limited dependent variable models is not straightforward and comes along with some difficulties, as shown below. Approaches already presented in the literature to tackle these difficulties are not satisfactory as they do not lead to a unique decomposition or do not take into account the nonlinearity of the model. On the contrary, we propose a decomposition approach which leads to a unique decomposition and accounts for the nonlinearity of the model in a rather intuitive manner.

Our decomposition approach is in the spirit of the famous Oaxaca-Blinder decomposition method. The Oaxaca (1973) and Blinder (1973) decomposition is a well-known and often applied technique to decompose the mean differential in some outcome variable between two groups into a part which is due to differences in observable characteristics (*explained differential*) and another part which is due to differences in unobservable characteristics (*unexplained differential*). A typical example is an analysis of the mean wage differential between, e.g., men and women or white and black people. Under some conditions, the unexplained differential can be attributed to discriminatory behavior of firms, households or other economic institutions; hence the Oaxaca-Blinder decomposition has often been applied to analyze the impact of discrimination.

The Oaxaca-Blinder decomposition in its original version can be applied to econometric models which are linear in parameters and explanatory variables. An extension to limited dependent variable models has been suggested by Bauer and Sinning (2008), for instance. However, Bauer and Sinning only provide a decomposition into the total explained and unexplained differential. We proceed further and consider a *detailed* decomposition of the explained differential, which means that we seek to decompose the explained differential into the contribution of each explanatory variable. In case of wage differentials, a detailed decompositions allows the researcher to make statements like "10 percent of the mean wage differential between men and women can be explained by differences in educational attainment, 20 percent by differences in working experience", and so on.

In this paper, we focus our attention on the *explained* differential only since the unexplained differential is hard to interpret. In the linear Oaxaca-Blinder decomposition, the unexplained differential is given by differences in coefficients multiplied by a vector of characteristics of a particular group (e.g., men or women). In nonlinear models such as limited variable models, however, differences in coefficients could also be the result of a misspecified model. Moreover, nonlinear models typically involve nuisance parameters (such as a variance parameter); a detailed decomposition of the unexplained differential would also have to attribute differences in nuisance parameters to the effects of specific factors. A detailed decomposition is then hard to justify economically. A further critique which applies to linear and nonlinear decompositions has been pointed out by Jones (1983). As Jones has shown, a detailed decomposition of the unexplained differential is not unique if there are dummy variables among the list of explanatory variables. The detailed decomposition then depends on the reference category chosen for the dummy variable, hence the decomposition is not unique.

On the contrary, a detailed decomposition of the *explained* differential assumes an identical model structure for the analyzed groups. That means, we relate the mean differential in the outcome variable only to differences in explanatory variables, but holding constant the model structure. In case of the Oaxaca-Blinder decomposition that means we consider differences in (mean) explanatory variables, evaluated at a constant coefficient vector of *one* particular group.

A detailed decomposition in linear models is rather straightforward, since the mean differential in the outcome variable can directly be attributed to the mean differential in the explanatory variables. This, however, is not true for limited dependent variable models. Fairlie (1999, 2005) and Yun (2004) have proposed approaches to obtaining detailed decompositions in such models. As will be shown below, Fairlie's decomposition is path-dependent, which means that the decomposition relies on the ordering of explanatory variables. Since different orderings imply different decomposition results, Fairlie's approach has the drawback that it does not lead to a unique decomposition.

Yun (2004) seeks to tackle the difficulties associated with the nonlinear model structure by two linearizations, thus bringing the model back to the linear case where mean differences in the outcome variable can directly be related to mean differences in the explanatory variables. However, such a procedure has the drawback that it ignores the nonlinear model structure. For instance, if the outcome differential is located in the tails of the distribution or in case of large differences in the explanatory variables (see Fortin et al., 2011, p. 52), such a linearization is likely to be inadequate.[1]

Our approach is based on a linearization using marginal effects, hence we explicitly account for the nonlinearity of the model in a way which is familiar from the general analysis of limited dependent variable models. Fortin et al. (2011) have already mentioned such a possibility (without providing details, though), but have also remarked that the contribution of each variable derived in such a decomposition would not add up to the total differential. This remark is only partly true. By applying the mean value theorem, we will show that there is exactly one marginal effect which not only leads to a detailed decomposition that adds up to the total differential, but which also leads to a unique decomposition and which has a very appealing interpretation.

The remainder of the paper is structured as follows. In section 2, we set up the econometric framework. In section 3 we derive the detailed decomposition theoretically, whereas section 4 shows how to estimate the detailed decomposition. In section 5, we compare our decomposition method to the approaches of Fairlie (2005) and Yun (2004). Finally, section 6 concludes the paper.

[1] This is the same argument why one should at all use a limited dependent variable model.

## 2 Econometric Framework

We consider the following latent representation of a limited dependent variable model:

$$y_i^* = x_i'\beta + \varepsilon_i, \tag{1}$$

where $i = 1, \ldots, n$ indexes individuals, $y^*$ is the latent dependent variable, $x$ is a vector of explanatory variables associated with a coefficient vector $\beta \in \mathbb{R}^{K+1}$ and $\varepsilon$ is a zero-mean error term. We assume that $x$ contains a constant term in its first component and $K$ "real" explanatory variables. We denote the observable dependent variable by $y$ which is functionally related to $y^*$. For instance, in a binary choice model we would have that $y = 1(y^* > 0)$, where $1(\cdot)$ denotes the indicator function. Furthermore, we let $d$ be a group indicator, taking a value of one if an individual belongs to a certain group and zero otherwise.

We make the following assumptions:

Assumption 1: *$\{(y_i, x_i)\}_{i=1}^n$ is an i.i.d. sample from some underlying distribution.*

Assumption 2: *$E[y_i|x_i] = G(x_i'\beta, \psi)$, $\forall i = 1, \ldots, n$ almost surely, where $G : \mathbb{R} \times \Psi \to \mathbb{R}$ is a known (link) function which (a) is continuous, (b) is differentiable, (c) depends on $x$ only through $x'\beta$; $\psi \in \Psi$ denotes a vector of nuisance parameters.*

We need these assumptions for deriving our proposed detailed decomposition in the next section. Note that Assumption 2 covers some well-known limited dependent variable models such as probit, logit and tobit. For these three models, we have the following link functions:

- Probit: $G(x_i'\beta, \sigma) = \Phi(x_i'\beta/\sigma)$;
- Logit: $G(x_i'\beta, \sigma) = \Lambda(x_i'\beta/\sigma) = \frac{\exp\{x_i'\beta/\sigma\}}{1+\exp\{x_i'\beta/\sigma\}}$;
- Tobit with truncation from the left at zero: $G(x_i'\beta, \sigma) = \Phi(x_i'\beta/\sigma)\left(x_i'\beta + \sigma\frac{\phi(x_i'\beta/\sigma)}{\Phi(x_i'\beta/\sigma)}\right)$,

where $\sigma = \sqrt{E[\varepsilon_i^2|x_i]}, \quad \forall i = 1, \ldots, n$; $\Phi(\cdot)$ and $\phi(\cdot)$ denote the standard normal cumulative distribution function and density function, respectively.

Note that Assumption 2 holds for all individuals, i.e. irrespective of whether $d$ is equal to one or zero. Since we are only concerned with a detailed decomposition of the explained differential, we can ignore issues such as group-dependent parameters or other group-dependent model structures.

When reviewing the Oaxaca-Blinder decomposition for the linear model in the context of discrimination, Oaxaca and Ransom (1994) suggest that the explained differential should be evaluated not at the coefficient vector of one particular group, but at the coefficient vector in the absence of discrimination. We generalize this point of view and consider the framework in equation (1) and Assumption 1 and 2 to represent a model structure in the absence of discrimination.

This point will also be important for the economic interpretation of our proposed detailed decomposition, which will be derived in the next section.

# 3 Derivation of the Detailed Decomposition

In this section we derive (and define) the detailed decomposition. We begin with a formal notation of the total explained differential, which we define as

$$\Delta = E[y_i|d_i = 1] - E[y_j|d_j = 0]. \tag{2}$$

This definition has also been proposed by Fortin et al. (2011, p. 52). The explained differential is thus given by the expected difference in the outcomes of each group. By the law of iterated expectations and Assumption 2, it follows that

$$E[y_i|d_i = 1] - E[y_j|d_j = 0] = E[G(x_i'\beta, \psi)|d_i = 1] - E[G(x_j'\beta, \psi)|d_j = 0]. \tag{3}$$

Since observations are i.i.d. (due to Assumption 1), we can write

$$E[G(x_i'\beta,\psi)|d_i=1] - E[G(x_j'\beta,\psi)|d_j=0] \tag{4}$$

$$= E[G(x_i'\beta,\psi)|d_i=1,d_j=0] - E[G(x_j'\beta,\psi)|d_i=1,d_j=0] \tag{5}$$

$$= E[G(x_i'\beta,\psi) - G(x_j'\beta,\psi)|d_i=1,d_j=0]. \tag{6}$$

In order to obtain a detailed decomposition of the explained differential, we linearize the term in the expectations operator by applying the mean value theorem. This yields

$$E[G(x_i'\beta,\psi) - G(x_j'\beta,\psi)|d_i=1,d_j=0] \tag{7}$$

$$= E[g((x_{ij}^*)'\beta,\psi)(x_i-x_j)'\beta|d_i=1,d_j=0], \tag{8}$$

where $g(u,\psi) = \partial G(u,\psi)/\partial u$ and $(x_{ij}^*)'\beta$ is a scalar lying on the line segment joining $x_i'\beta$ and $x_j'\beta$. Note that $x_{ij}^*$ can also be represented as

$$x_{ij}^* = \lambda x_i + (1-\lambda)x_j \tag{9}$$

for some $\lambda \in (0,1)$.

Due to the linearization, we define the contribution of each variable to the explained differential as follows:

Definition 1: ***Detailed Decomposition***. *The contribution of a variable $x_k$ to the explained differential is given by $c_k = E[g((x_{ij}^*)'\beta,\psi)\beta_k(x_{i,k}-x_{j,k})|d_i=1,d_j=0]$, $\forall k=1,\dots,K$.*

Note that the mean value theorem guarantees that the contributions of the variables add up to the total explained differential. Furthermore, note that Definition 1 implies that the contribution of each variable is given by the difference in explanatory variables multiplied with the marginal effect of this variable. Hence, as suggested by Fortin et al. (2011) our decomposition approach evaluates differences in variables between two groups at the marginal effects of these variable, thus taking into account the nonlinearity of the underlying model.

But we have a specific marginal effect. In general, marginal effects could be evaluated at any value of the explanatory variables. However, it makes an intuitive sense that we choose the marginal effect evaluated at $x_{ij}^*$ in order to define the detailed decomposition. To see this, note that, by equation (9), the marginal effect in Definition 1 is based on a convex combination of the explanatory variables of two individuals which belong to different groups. Suppose for the moment that one group consists of males and the other one of females. The convex combination may be interpreted so as to represent a *synthetic individual*, so that the marginal effect in Definition 1 is the marginal effect of a synthetic individual which is a combination of the male and female individual. Now suppose that our synthetic individual is initially endowed like the female individual ($j$). Then, after receiving the difference $x_i - x_j$, the marginal effect implies a change of the synthetic individual from the female ($j$) to the male individual ($i$) in terms of the value of the link function $G$. Given that our model represents a situation in the absence of discrimination, this marginal effect can thus be interpreted as the *marginal effect in the absence of discrimination*. This is a generalization of the suggestion by Oaxaca and Ransom (1994) that the explained differential in linear models should be evaluated at a coefficient vector which would be prevalent in the absence of discrimination.

# 4 Estimation of the Detailed Decomposition

The detailed decomposition proposed in Definition 1 is, of course, a theoretical one and represents a population concept (due to the expectations operator). In this section we show how the detailed decomposition can be estimated. Furthermore, we prove consistency and asymptotic normality of our proposed estimator. Finally, we discuss how inference can be conducted if the parameter vector $\theta$ is estimated as well.

Let $\mathcal{D} = \{i : d_i = 1\}$ denote the set of individuals belonging to the group with $d = 1$

and $m = \sum_{i=1}^n 1(d_i = 1)$ be the corresponding number of group members. We propose to estimate $c_k = E[g((x_{ij}^*)'\beta, \psi)\beta_k(x_{i,k} - x_{j,k}), \psi)|d_i = 1, d_j = 0]$ by

$$\hat{c}_k = \frac{1}{m(n-m)} \sum_{i \in \mathcal{D}} \sum_{j \notin \mathcal{D}} g((x_{ij}^*)'\beta, \psi)\beta_k(x_{i,k} - x_{j,k}), \psi). \tag{10}$$

Thus, we take all possible pairs between members of both groups, compute the detailed decomposition as in Definition 1 for each pair and then average over these pair-specific decompositions to obtain an approximation to the theoretical expectation in Definition 1.

Note that the estimator $\hat{c}_k$ contains $x_{ij}^*$ which follows from the mean value theorem. However, it is not necessary to calculate $x_{ij}^*$ explicitly. It suffices to calculate $g((x_{ij}^*)'\beta, \psi)$ by

$$g((x_{ij}^*)'\beta, \psi) = \frac{G(x_i'\beta, \psi) - G(x_j'\beta, \psi)}{(x_i - x_j)'\beta}. \tag{11}$$

Hence, in practice it is not complicated to calculate $\hat{c}_k$ for each explanatory variable.

To prove consistency and asymptotic normality of our estimator, note that $\hat{c}_k$ is in fact a 2-sample $U$-statistic (cf. Serfling, 1980, ch. 5), which belongs to the class of generalized $U$-statistics. Define $h_k(x_i, x_j; \theta) = g((x_{ij}^*)'\beta, \psi)\beta_k(x_{i,k} - x_{j,k}), \psi)$. Then, we can establish the following theorems:

Theorem 1: ***Consistency***. *Under Assumptions 1, 2, and $E|h_k(x_i, x_j; \beta)| < \infty$, we have that $\hat{c}_k \xrightarrow{p} c_k, k = 1, \ldots, K$.*

*Proof.* Follows immediately from Theorem A of Serfling (1980, p. 190), and its extension to generalized $U$-statistics (see Serfling, 1980, p. 191).

Theorem 2: ***Asymptotic Normality***. *Under Assumptions 1, 2, $E[h_k(x_i, x_j; \theta)^2] < \infty$, and $m/N \to \delta, \delta \in (0, 1)$ we have that $\sqrt{n}(\hat{c}_k - c_k) \xrightarrow{d} \mathcal{N}(0, \eta_k^2), k = 1, \ldots, K$, where*

$$\eta_k^2 = \delta^{-1}\{E[h_k(x_1, x_2; \theta)h_k(x_1, x_3; \theta)] - c_k^2\} + (1-\delta)^{-1}\{E[h_k(x_1, x_3; \theta)h_k(x_2, x_3; \theta)] - c_k^2\}. \tag{12}$$

*Proof.* Follows immediately from Theorem 12.6 of Van der Vaart (1998, p. 166).

In our derivation of asymptotic normality we assumed that the parameter vector $\theta$ is deterministic or known in advance. However, in practice $\theta$ must be estimated. For the link functions listed in section 2, estimates can be obtained by using the probit, logit or tobit model; estimation routines for these models are contained in any standard statistical software package. However, estimation of $\theta$ alters our asymptotic distribution results for $\hat{c}_k$, since we have to adjust the variance term for the uncertainty due to estimating $\theta$ rather than assuming its value to be known. Since an adjustment of the asymptotic distribution is complicated, we suggest to employ the bootstrap rather than asymptotic distribution results in order to conduct inference.

We suggest the following procedure: At first, one obtains estimates of the detailed decomposition according to equation (10). Standard errors can then be obtained by bootstrapping. That means, we sample with replacement from the original sample and compute the $\hat{c}_k$'s a large number of times (say $P$ times). Let $\hat{c}_{kp}$ denote the estimate of $c_k$ in the $p$th bootstrap replication. Then, the standard error of $\hat{c}_k$ can be estimated as

$$\hat{se}(\hat{c}_k) = \sqrt{\frac{1}{P}\sum_{p=1}^{P}(\hat{c}_{kp} - \hat{c}_k)^2}, \tag{13}$$

where $\hat{c}_k$ is the estimate based on the original sample.

# 5 Comparison to Existing Decomposition Methods

In this section, we compare our proposed decomposition method to competing approaches proposed by Fairlie (1999, 2005) and Yun (2004). Both authors derive their detailed decompositions in a finite-sample-context and do not provide population considerations (as expressed by the expectations operator on Defintion 1 above). We briefly discuss their methods and show that our proposed detailed decomposition overcomes the main drawbacks of these approaches.

We begin with Fairlie (1999, 2005) who proposes what is called a *sequential decomposition*. Fairlie analyzes a detailed decomposition for binary choice models, i.e., probit and logit models. For simplicity, we consider the case of only two explanatory variables ($K = 2$). For notational ease, we index individuals with $d = 1$ by $w$ and individuals with $d = 0$ by $b$. Fairlie's procedure works as follows:

1. Reduce the size of the larger group (by randomly selecting individuals) so that both groups have the same size $l$.

2. Rank observations by their predicted probability that $y$ is equal to one (i.e., by $G(x'\beta)$) *within* each group.

3. Match observations from both groups which have the same rank.

4. Let $x_{v,i,j}$ denote the value of variable $x_j$ for an individual $i$ from group $v \in \{w, b\}$. Fairlie's sequential decomposition would then be given by

$$\Delta \approx \frac{1}{l}\sum_{i=1}^{l}\{G(\beta_0 + x_{w,i,1}\beta_1 + x_{w,i,2}\beta_2, \psi) - G(\beta_0 + x_{b,i,1}\beta_1 + x_{b,i,2}\beta_2, \psi)\} \tag{14}$$

$$= \underbrace{\frac{1}{l}\sum_{i=1}^{l}\{G(\beta_0 + x_{w,i,1}\beta_1 + x_{w,i,2}\beta_2, \psi) - G(\beta_0 + x_{b,i,1}\beta_1 + x_{w,i,2}\beta_2, \psi)\}}_{\hat{c}_1} \tag{15}$$

$$+ \underbrace{\frac{1}{l}\sum_{i=1}^{l}\{G(\beta_0 + x_{b,i,1}\beta_1 + x_{w,i,2}\beta_2, \psi) - G(\beta_0 + x_{b,i,1}\beta_1 + x_{b,i,2}\beta_2, \psi)\}}_{\hat{c}_2}, \tag{16}$$

where the index $i$ runs over the matched observations from both groups.

Hence, the contribution of a variable is given by the average change of the link function $G$ if the variable of interest is changed while holding all other variables constant. Note that the decomposition ensures that the sum of the contributions of the explanatory variables is equal to the individual gap in the values of the link function between the matched individuals. A disadvantage is, however, that the contributions of each variable depend on the ordering of variables. If the order of variables during the decomposition is interchanged, different decomposition results will be obtained, so that the decomposition is not unique. This problem, which is known as *path-dependency* (cf. Fortin et al., 2011, p. 27), is a drawback of any sequential decomposition. On the contrary, our approach derived in section 3 is not a sequential one, implying that our decomposition results are unique in the sense that they do not depend on the ordering of variables.

However, our decomposition results are not only unique in this sense. As mentioned before, Fairlie's methodology is based on a matching procedure for individuals from both groups. However, the matching procedure is arbitrary and lacks a theoretical foundation. Our approach, on the other hand, is theoretically founded and uses all between-group-pairs (recall equation (10)) of individuals, thus avoiding an arbitrary matching procedure.

The decomposition approach proposed by Yun (2004)[2] is based on two linearizations to bring the model back to the linear case, where a detailed decomposition is straightforward. His decomposition is given by

$$\begin{aligned}
\Delta &\approx \frac{1}{m}\sum_{i\in\mathcal{D}} G(x_i'\beta) - \frac{1}{n-m}\sum_{i\notin\mathcal{D}} G(x_i'\beta) && (17)\\
&= G(\bar{x}_w'\beta) - G(\bar{x}_b'\beta) + R_M && (18)\\
&= (\bar{x}_w - \bar{x}_b)'\beta g(\bar{x}_w'\beta) + R_M + R_T, && (19)
\end{aligned}$$

where $\bar{x}_w = \frac{1}{m}\sum_{i\in\mathcal{D}} x_i$ and $\bar{x}_b = \frac{1}{n-m}\sum_{i\notin\mathcal{D}} x_i$; $R_M$ and $R_T$ denote appropriate remainder terms. The contribution of a variable $x_k$ is given by

$$\hat{c}_k = \frac{(\bar{x}_{w,k} - \bar{x}_{b,k})'\beta_k g(\bar{x}_w'\beta)}{(\bar{x}_w - \bar{x}_b)'\beta g(\bar{x}_w'\beta)} \left\{ \frac{1}{m}\sum_{i\in\mathcal{D}} G(x_i'\beta) - \frac{1}{n-m}\sum_{i\notin\mathcal{D}} G(x_i'\beta) \right\}. \tag{20}$$

[2]For the explained differential, Even and Macpherson (1990, 1993) used the same decomposition methodology as derived by Yun (2004) in order to explain the decline of unionism in the United States. However, they just stated the decomposition method without providing a formal derivation.

Note that Yun develops weights based on equation (19) which are given by

$$\frac{(\bar{x}_{w,k} - \bar{x}_{b,k})'\beta_k g(\bar{x}'_w\beta)}{(\bar{x}_w - \bar{x}_b)'\beta g(\bar{x}'_w\beta)}. \tag{21}$$

These weights are then multiplied with the observed differential $\frac{1}{m}\sum_{i\in\mathcal{D}} G(x'_i\beta) - \frac{1}{n-m}\sum_{i\notin\mathcal{D}} G(x'_i\beta)$ in order to yield the contribution of a variable $x_k$.

However, note that the weights in equation (21) reduce to

$$\frac{(\bar{x}_{w,k} - \bar{x}_{b,k})'\beta_k}{(\bar{x}_w - \bar{x}_b)'\beta}, \tag{22}$$

so that we have the same weights as in decompositions for linear models since the nonlinear component $g(\bar{x}'_w\beta)$ cancels out. Put differently, the Yun procedure ignores the nonlinear model structure. As mentioned in the introduction, this may be problematic if the outcome differential is located in the tails of the distribution or in case of large differences in the explanatory variables (see Fortin et al., 2011, p. 52).

We illustrate this point by means of a small numerical example. Our model is given by

$$y_i^* = -6 + x_1 + x_2 + \varepsilon_i \tag{23}$$

$$y_i = 1(y_i^* > 0) \tag{24}$$

$$\varepsilon_i \sim \mathcal{N}(0,1). \tag{25}$$

Hence, we consider a probit model with a link function given by $G(x'_i\beta) = \Phi(x'_i\beta)$, where $\Phi(\cdot)$ is again the standard normal cumulative distribution function. We set $n = 2,000$ with 1,000 individuals belonging to each group. Moreover, for the group with $d = 1$ we specified that $x_1 \sim \mathcal{N}(6,1)$ and $x_2 \sim \mathcal{N}(3,1)$. For the group with $d = 0$ we took $x_1 \sim \mathcal{N}(2,1)$ and $x_2 \sim \mathcal{N}(2,1)$. Hence, the group with $d = 1$ has larger mean values for both explanatory variables, in particular with respect to the variable $x_1$.

We simulated this model with 1,000 replications and performed our proposed decomposition and Yun's method. We then averaged over the 1,000 replications in order to obtain results. Over these replications, the (averaged) mean of our dependent variable $y$ is 0.959 for the group with $d = 1$ and 0.124 for the group with $d = 0$, so that the averaged differential in the outcome variable is given by 0.835. The averaged decomposition results are given in the following table:

| | $\hat{c}_1$ | $\hat{c}_2$ |
|---|---|---|
| Our decomposition | .6912 | .1432 |
| Yun's decomposition | .6675 | .1669 |

Hence, we see that our proposed decomposition and Yun's method yield different results. In this example, the differences are not large. Nevertheless, they indicate that the Yun method may be too rough since it does not properly account for the nonlinearity of the underlying model.

# 6 Conclusion

In this paper we derived a detailed decomposition (of the explained differential) for limited dependent variable models. We first defined the detailed decomposition theoretically and then showed how the theoretical decomposition can be consistently estimated using sample data. We also proposed a bootstrap procedure for obtaining standard errors for the decomposition results. Unlike existing approaches discussed in the literature to perform detailed decompositions in nonlinear econometric models, our method leads to a unique decomposition and accounts for the nonlinearity of the model in a rather intuitive way (i.e., by using marginal effects to evaluate differences in explanatory variables). Moreover, in light of the suggestion by Oaxaca and Ransom (1994) that the explained differential should be evaluated at a parameter vector which would be prevalent in the absence of discrimination, our decomposition approach provides a natural extension of this idea to nonlinear models.

A detailed decomposition of the explained differential in a limited dependent variable model is important because it allows to relate differences in non-continuous outcome variables to differences in characteristics. For instance, one can analyze which characteristics contribute most to the differential in, say, labor force participation rates between men and women. Another field of research where our method can be applied is an analysis of the erosion of union membership over time, where the erosion can be attributed to changes in the characteristics of the workforce (see Fitzenberger et al., 2011). Hence, our method cannot only be applied to group differences at a given point in time, but it can also be used to analyze changes over time (where two points in time serve as “groups”). However, such applications are left for future research.

# References


Bauer, T. and Sinning, M. (2008). An extension of the Blinder-Oaxaca decomposition to nonlinear models. *AStA Advances in Statistical Analysis*, 92(2):197–206.

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, 8(4):pp. 436–455.

Even, W. E. and Macpherson, D. A. (1990). Plant size and the decline of unionism. *Economics Letters*, 32(4):393–398.

Even, W. E. and Macpherson, D. A. (1993). The decline of private-sector unionism and the gender wage gap. *Journal of Human Resources*, 28(2):279–296.

Fairlie, R. W. (1999). The absence of the african-american owned business: An analysis of the dynamics of self-employment. *Journal of Labor Economics*, 17(1):pp. 80–108.

Fairlie, R. W. (2005). An extension of the Blinder-Oaxaca decomposition technique to logit and probit models. *Journal of Economic and Social Measurement*, 30(4):305–316.

Fitzenberger, B., Kohn, K., and Wang, Q. (2011). The erosion of union membership in germany: determinants, densities, decompositions. *Journal of Population Economics*, 24(1):141–165.

Fortin, N., Lemieux, T., and Firpo, S. (2011). *Decomposition Methods in Economics*, volume 4A of *Handbook of Labor Economics*, chapter 1, pages 1–102. Elsevier.

Jones, F. L. (1983). On decomposing the wage gap: a critical comment on Blinder's method. *Journal of Human Resources*, 18(1):126–130.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14(3):693–709.

Oaxaca, R. L. and Ransom, M. R. (1994). On discrimination and the decomposition of wage differentials. *Journal of Econometrics*, 61(1):5–21.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York, NY.

van der Vaart, A. W. (1998). *Asymptotic Statistics.* Cambridge University Press, New York, NY.

Yun, M.-S. (2004). Decomposing differences in the first moment. *Economics Letters*, 82(2):275–280.