I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16188

Oaxaca-Blinder Meets Kitagawa: What Is the Link?

Ronald L. Oaxaca
Eva Sierminska

MAY 2023

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16188

Oaxaca-Blinder Meets Kitagawa: What Is the Link?

Ronald L. Oaxaca
University of Arizona and IZA

Eva Sierminska
LISER, IZA, DIW Berlin and University of Arizona

MAY 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Oaxaca-Blinder Meets Kitagawa: What Is the Link?

Recently, papers have started combining the naming of two popular decomposition methods: the Oaxaca-Blinder method and the Kitagawa method, a popular method in demographics and sociology. Although the two approaches have the same objective in terms of decomposing outcome differences in some variable of interest between two populations, they are framed quite differently and do not overlap except in a special set of circumstances. Consequently, the combined labeling of these two approaches can be misleading. This note establishes the conditions under which the two methodologies are identical and when they are not. It also provides the citation history of the two methods and examples of "misuses" of the naming convention when the methods are not equivalent, accompanied by a proposal for the way forward.

JEL Classification: A10, B41, J0

Keywords: decomposition methods, economics, demography

Corresponding author:
Eva Sierminska
Luxembourg Institute of Socio-Economic Research
Campus Belval
11 Porte des Sciences
L-4366 Esch/Alzette
Luxembourg

E-mail: eva.sierminska@liser.lu

1 Introduction

In the 1950s, Evelyn M. Kitagawa, a professor of sociology and demography at the University of Chicago, published a paper "Components of a Difference between Two Rates" in the American Statistical Association Journal. The paper developed a method to decompose rates (or proportions) for two demograhic groups by controlling for differences between these two groups in selected characteristics. The method is referred to as "standardization" as the two rates are related to a standard population that has a specified age-sex-race composition.[1]

Some 20 years later, Alan Blinder published a paper in 1973 (JHR) using linear regression-based methods in which the mean difference in wages between two groups (a high-wage group and a low-wage group) is expressed as the sum of differentials attributable to differing endowments and those attributable to differing coefficients and an unexplained portion of the differential. In the same year, Ronald Oaxaca published a paper (IER) that estimates the average extent of discrimination faced by female workers in the United States.[2] By interpreting the component corresponding to gender differences in coefficients as the Becker discrimination coefficient and using the properties of ordinary least squares estimation, Oaxaca shows that the wage differential between two groups can be expressed as the sum of the estimated effects of differences in individual characteristics and the estimated effects of discrimination. Since then, these two papers have very often been cited together when referring to the Oaxaca-Blinder (OB) decomposition (typically conducted at the mean). As of March 31, 2022, the Blinder paper has been cited 8843 times and the Oaxaca paper 11418 times. Although published in the same year, these two papers do not cite each other, although Blinder does thank Oaxaca in the footnote for helpful comments and cites the conference volume version.

Recently, papers have started to combine the naming of the two methods referring to the Kitagawa-Oaxaca-Blinder (KOB) decomposition method. Although the two approaches have the same objective in terms of decomposing outcome differences in some variable of interest between two populations, they are framed quite differently and do not overlap except in a special set of circumstances. Hence, combined labeling of these two approaches

[1] So, for example, by applying the schedule of age-sex-race specific death rates for each of the groups to the age-sex-race composition of the standard population and then noting the difference between the two, with a certain confidence one can say that the remaining differences after standardization are due to factors other than the age, sex and race composition.

[2] A conference volume version chapter was published the same year by Princeton University Press.

can be misleading. Certainly, the combined label is inappropriate in any paper that actually does not use either the Kitagawa method or the OB method. However, the use of the combined label can also be misleading in circumstances when a paper uses only one of these methods while the other is not applicable.

This note establishes the conditions under which the two methodologies are identical and when they are not. First, in the next section we focus on the legacy of the Kitagawa method. Next, in section 3, we present an overview of each methodology, compare them, and offer a simple example of the set of circumstances in which they are identical. Section 4 is an overview of how the use of the KOB designation in published research papers may be incorrect or misleading. Section 5 summarizes the paper's conclusions.

2 The "Kitagawa" Legacy and the Recent Surge in Popularity

The Kitagawa "standardization" method is a popular decomposition method in demography and sociology. It is discussed in popular demographic textbooks including Preston et al. (2001) and Das Gupta (1993). The paper has received over 964 citations according to Google Scholar. Scopus indicates that citations to Kitagawa's seminal paper have been increasing annually over 1970 - 2022 (see Figure 1) and in particular over the last decade. As of December 20, 2022, there was a cummulative total of 415 citations. According to the Web of Science, the 1955 paper has been cited 392 times - mostly in sociological and demographic journals, but more recently in other type of journals.

We examine the citation history by field based on journals in Table 1 looking at the five previous decades and the last two years of this decade. The results are striking. As with Scopus, overall citations of the 1955 Kitagawa paper have doubled in the second decade of this century and in the last two years of this decade have reached those similar levels. In some fields, e.g., Ecology, Statistics, the citation rates remains more or less unchanged. In Medicine and Public Health the increase is not as striking as in the social sciences. Here, the citation rates in 2010-19 at least doubled compared to the previous decade. In Sociology/Political Science it tripled and in Economics grew seven-fold. In the first two years of 2020, the citations in the social sciences have exceeded those of the previous decade. Thus, the paper has been gaining popularity - particularly more recently, but only in selected fields - in the social sciences - and in Economics more so than in any other field.

3 The Two Decomposition Methods: Is there a link?

3.1 The Kitagawa Decomposition

Kitagawa never explicitly referred to her method as a decomposition, though it clearly was a decomposition method. Nor did she frame her method in terms of linear regressions. Kitagawa's decomposition technique was closely related to standardization methods that addresses such questions as "How much of the difference between death rates in A and B is attributable to differences in their age (or race and sex) distributions?" The development of Kitagawa's method was based on previous standardization methods used in sociology and demography, which identified for example, "influences of changes in age distribution" and "influences of changes in occurrence rates" and are included in Jaffe (1951) Handbook of Statistical Methods for Demographers on standardization methods. The main contribution of Kitagawa was to develop a technique called "components of a difference between two rates." The goal of this technique was to explain the difference between rates of two groups by decomposing the difference into differences in their specific rates and differences in their composition. This is the two-component method. Kitagawa also develops a three-component method in which the group difference in rates is decomposed into differences in their specific rates, differences in their composition, and differences attributable to the interaction between group differences in their specific rates and group differences in their composition.

The Kitagawa methodology is broader than the prior techniques for standardization rates, which were developed to summarize and compare differences in two (or more) sets of specific rates. Kitagawa's method provides counterfactual outcome rates based on assumed standard variables.

3.2 The Oaxaca-Blinder Decomposition

Unlike the original Kitagawa method, the OB decomposition is framed in terms of linear regression methods applied to the mean difference in some outcome variable between two groups. This mean group outcome difference is decomposed into the differences arising from group differences in the parameters and group differences in characteristics. The OB decomposition methodology can also include a term that captures the interaction between group parameter differences and group characteristics. Although the method has been significantly extended to decompositions associated with quantile regression, limited dependent variables, etc., these extensions can still fit comfortably under the OB heading.

3.3 Kitagawa vs OB: A Comparison of the Two Methods

Our investigation reveals that the Kitagawa decomposition method is a special case of OB, or equivalently, OB is a generalization of the Kitagawa decomposition method. In particular, the OB methodology is identical/equivalent to the Kitagawa methodology when the dependent outcome variables are binary, the covariates are categorical variables comprised of sets of indicator variables, and the OB decomposition uses the OLS estimated linear probability model (LPM). In short, all Kitagawa decompositions are OB, but not all OB are Kitagawa decompositions. A detailed example of the conditions described above is presented in the Appendix.

For the specialized circumstances associated with the Kitagawa methodology, manual computation becomes increasingly unwieldy in the presence of three or more categorical covariates. This would also be true with the regression approach associated with OB. It must be kept in mind that Kitagawa's seminal paper appeared in 1955. With the advent of personal computers, multiple regression became trivial and software implementing the Kitagawa method would also render the decomposition exercise trivial. Yet, the condition of a binary dependent variable and indicator explanatory variables would need to remain in place. Nevertheless, framing the decomposition in terms of a regression framework might seem more intuitive to some.

An interesting aspect of the Kitagawa methodology is that a researcher does not require the underlying micro data set to conduct the decompositions, only the requisite proportions/rates. The regression framework analog is that the researcher would require only the requisite sample moments. Such circumstances might arise if there are privacy concerns about identifying individuals. Disclosing only the sample proportions or sample moments would permit the researcher to conduct the decompositions without breaching confidentiality.

3.4 Exploring the link between the two methods: an Illustrative Example

It is helpful to illustrate the relationship between Kitagawa's decomposition methodology and the OB approach by examining Kitagawa's method in terms of a simple case in which there is one factor (explanatory variable) and the two-component decomposition (Kitagawa (1955), p.1182).[3] With some notational changes from that used by Kitagwa, suppose there are two population groups, A and B, for which the group mean difference in a binary indicator outcome variable Y is decomposed. The sample sizes for the two population groups are denoted by N^A and N^B. The explanatory variable X is a cate-

[3] Derivation details for the illustrative example are provided in the appendix.

gorical variable defined by a mutually exclusive set of K indicator variables such that $\sum_{k=1}^{K} X_{ik} = 1 \; \forall i$ in a group sample, where i refers to the ith individual. Accordingly, the group outcome proportion difference to be decomposed is given by

$$\overline{Y}^A - \overline{Y}^B = \frac{\sum_{i=1}^{N^A} Y_i^A}{N^A} - \frac{\sum_{i=1}^{N^B} Y_i^B}{N^B} = \tag{1}$$

$$= \frac{N_Y^A}{N^A} - \frac{N_Y^B}{N^B},$$

where Y_i^A, Y_i^B are the binary outcome indicators for the ith individuals in groups A and B, and N_Y^A, N_Y^B are the numbers of group A and B individuals for whom $Y_i = 1$.

Kitagawa's one-factor, two-component decomposition is given by

$$\overline{Y}^A - \overline{Y}^B = \text{Gross}\, X + \text{Residual}\, X, \tag{2}$$

where

$$\text{Gross}\, X = \sum_{k=1}^{K} \left(\frac{\overline{Y}_k^A + \overline{Y}_k^B}{2} \right) \left(\frac{N_k^A}{N^A} - \frac{N_k^B}{N^B} \right) \tag{3}$$

$$\text{Residual}\, X = \sum_{k=1}^{K} \frac{\left(\frac{N_k^A}{N^A} + \frac{N_k^B}{N^B} \right)}{2} \left(\overline{Y}_k^A - \overline{Y}_k^B \right). \tag{4}$$

Note that the group weights used to provide the decomposition reference group in eq (3) and eq (4) are the simple averages of the Y outcome weights and the X-specific compositional rates. Gross X measures how much of the group difference in the mean outcome arises from group compositional differences in the explanatory variable X. The Residual X component measures how much of the group difference in the mean outcome arises from group differences in the X-specific Y outcome rates.

As explained in the Appendix, $\overline{Y}_k^A$ and $\overline{Y}_k^B$ denote the outcome rates among individuals in the kth category of variable X for groups A and B (for $k = 1, ..., K$):

$$\overline{Y}_k^A = \frac{\sum_{i=1}^{N^A} Y_i^A X_{ik}}{N_k^A} = \frac{N_{Yk}^A}{N_k^A}$$

$$\overline{Y}_k^B = \frac{\sum_{i=1}^{N^B} Y_i^B X_{ik}}{N_k^B} = \frac{N_{Yk}^B}{N_k^B},$$

where $N_{Yk}^A = \sum_{i=1}^{N^A} Y_i^A X_{ik}$ and $N_{Yk}^B = \sum_{i=1}^{N^B} Y_i^B X_{ik}$ represent the number of individuals in groups A and B for whom $(Y_i^A \cdot X_{ik}) = 1$ and $(Y_i^B \cdot X_{ik}) = 1$, and N_k^A and N_k^b represent the numbers of individuals in each group for whom $X_k = 1$. Thus, N_{Yk}^A and N_{Yk}^B are the numbers of individuals in the X specific category for whom the outcome variable $= 1$, and $\left(\frac{N_k^A}{N^A}\right)$ and $\left(\frac{N_k^B}{N^B}\right)$ are the X specific compositional weights.

Next, we examine the OB decomposition (Blinder (1973); Oaxaca (1973)). With the same data set, the OB regression-based approach would specify LPMs for the two population groups:

$$\begin{aligned} Y_i^A &= \sum_{k=1}^{K} X_{ik}\beta_k^A + \varepsilon_i,\ i = 1, \ldots, N^A \\ Y_i^B &= \sum_{k=1}^{K} X_{ik}\beta_k^B + \varepsilon_i,\ i = 1, \ldots, N^B. \end{aligned} \tag{5}$$

Note that there is no separate constant term as the indicator variables sum to 1.

Typically, one counterfactually assigns one of the two population groups to be the reference group. In the present case, we follow Kitagawa's counterfactual and adopt the simple average of the two groups's estimated coefficients, which turn out to be identical to each group's K-specific rates seen in the Kitagawa decomposition:

$$\begin{aligned} b_k^* &= \frac{(b_k^A + b_k^B)}{2},\ k = 1, \ldots, K \\ &= \frac{(\overline{Y}_k^A + \overline{Y}_k^B)}{2}. \end{aligned} \tag{6}$$

Accordingly, the OLS decomposition is given by

$$\overline{Y}^A - \overline{Y}^B = \sum_{k=1}^{K} \overline{X}_k^A b_k^A - \sum_{k=1}^{K} \overline{X}_k^B b_k^B \tag{7}$$

$$= \underbrace{\sum_{k=1}^{K} \left(\overline{X}_k^A - \overline{X}_k^B\right) b_k^*}_{\text{Explained}} + \underbrace{\sum_{k=1}^{K} \overline{X}_k^A \left(b_k^A - b_k^*\right) + \sum_{k=1}^{K} \overline{X}_k^B \left(b_k^* - b_k^B\right)}_{\text{Unexplained}}.$$

In the Appendix, we show that the Kitagawa decomposition components in this example are identical to the OB decomposition components:

$$\text{Gross } X = \sum_{k=1}^{K} \left(\frac{\overline{Y}_k^A + \overline{Y}_k^B}{2} \right) \left(\frac{N_k^A}{N^A} - \frac{N_k^B}{N^B} \right) = \underbrace{\sum_{k=1}^{K} \left(\overline{X}_k^A - \overline{X}_k^B \right) b_k^*}_{\text{Explained}} \tag{8}$$

$$\text{Residual } X = \sum_{k=1}^{K} \frac{\left(\frac{N_k^A}{N^A} + \frac{N_k^B}{N^B} \right)}{2} \left(\overline{Y}_k^A - \overline{Y}_k^B \right) = \underbrace{\sum_{k=1}^{K} \overline{X}_k^A \left(b_k^A - b_k^* \right) + \sum_{k=1}^{K} \overline{X}_k^B \left(b_k^* - b_k^B \right)}_{\text{Unexplained}} . \tag{9}$$

The counterfactual assumption in the above example is a special case of the generalized decomposition described in Neumark (1988); Oaxaca and Ransom (1988) and Oaxaca and Ransom (1994). If one adopts either of the group's X-specific rates as the standard (reference) outcome rate, the Kitagawa decomposition would correspond to the OB decomposition in which the same group's estimated coefficients are used for the counterfactual. Also, one can conduct a three-component decomposition with a decomposition term that captures the interaction between coefficient differences and differences in group characteristics.

4 The Kitagawa-Oaxaca-Blinder decomposition method?

Recently, papers have started to combine the naming of the two methods by referring to the Kitagawa-Oaxaca-Blinder decomposition method or the pooled Kitagawa-Oaxaca-Blinder decomposition. As far as we know, no such canonical method exists. As we have shown in the previous section, the Kitagawa decomposition is a special case of the Oaxaca-Blinder decomposition.

As of December 21, 2022, we find 358 references to the KOB decomposition method in google scholar mostly coming from the last two years. Our investigation identified clear distinctions among papers that identify their decompositions as KOB. For example, authors use the OB method and mention the Kitagawa method as a precursor to the regression-based method (Group 1). Other authors use the OB method and refer to it as KOB but do not indicate how their decompositions incorporate the Kitagawa method (Group 2).

Before we examine these cases in more detail, it is important to articulate how one might arrive at different criteria for applying the KOB label. On one end of the spectrum, one could apply the KOB label only to decompositions that correspond to the original Kitagawa method. As we have shown, the original Kitagawa decomposition is identical to OB decompositions based on LPM linear regressions if dependent binary indicator variables are regressed on categorical variables comprised of sets of indicator variables. Thus, any OB decomposition that does not fit these specific circumstances is not KOB.

A more inclusive practice could be to apply the KOB label to any regression-based decomposition that involves decomposing group mean differences in a dependent outcome variable into a component that arises from differences in the group means of the explanatory variables and a residual term, or a residual term and an interaction term. Yet, this could be misleading in instances where it is not mentioned that Kitagawa is an early version of the more general OB method. In addition, many researchers may not be aware that Kitagawa did not frame or recognize her methodology as regression based.

We now examine the two cases and shed light on the context. Below, we provide examples of papers in the two categories. Our review is not a comment on the research contributions of the papers examined, but only an examination of the correspondence of the KOB designation to the decompositions reported.

Group 1: Use OB, refer to K as a precursor to the regression-based method, and use the KOB label. Some papers refer to the Kitagawa method as a precursor to the OB regression based method and then label the method KOB. For example, Kroger and Hartmann (2021) state in the Stata Journal:

"In large parts of the applied literature, these kinds of decompositions are known as Oaxaca-Blinder decompositions, after two of the three scholars who pioneered these approaches (Oaxaca 1973; Blinder 1973). We refer to this way of decomposing group mean differences as the Kitagawa-Oaxaca-Blinder (KOB) approach to reference the earliest and often overlooked contribution to this literature as well (Kitagawa 1955).

(...)

Before we review the existing approaches to the decomposition of change, it is useful to recapitulate what the original KOB decomposition does using cross-sectional data (Kitagawa 1955; Oaxaca 1973; Blinder 1973). (...)

We start with a basic linear regression model for an outcome Y and two groups A and B:"

Here, the authors acknowledge that they include Kitagawa's name due to her early contribution to the literature yet it is not correct to say that her contribution has been overlooked, particularly in light of the numerous Kitagawa citations in demography and sociology. What about the appropriateness of labeling the "original KOB decomposition," as a linear regression? As previously shown in this note, the original Kitagawa method is identical to the LPM regression approach only under a special set of conditions. In addition, Kitagawa did not recognize her method as being regression based and to our knowledge this has not been previously shown. Another difficulty is that the regression equivalence of Kitagawa's method is limited to the data restrictions unique to her approach. Without recognition of this point, it might erroneously appear that the original Kitagawa method can be readily applied to data sets that do not satisfy the Kitagawa data restrictions.

Another example is that of Nieuwenhuis et al. (2020) when they state in their article:

"The data were analysed using a decomposition technique originally developed by Evelyn Kitagawa (1955). This technique is better known as the Blinder-Oaxaca decomposition (Oaxaca 1973; ...), (...) The first step in the decomposition is estimating a regression model (with the same model specification) for each year separately."

The paper acknowledges that Kitagawa developed a decomposition method early on; however, without providing additional context, the paper suggests that in fact the two methods are the same. As far as the original Kitagawa decomposition is concerned, the two methods are not equivalent in general. The paper also suggests that the Kitagawa method was expressed as a linear regression. Again, Kitagawa never framed her method in terms of a linear regression. She states herself in the 1955 paper that ".. the difference between two crude rates is not the equivalent of a concept like total variance of a dependent variable in regression analysis (...), which will be increasingly "explained" as more independent variables are added to the regression equation." Furthermore, the paper does not recognize that Oaxaca-Blinder is a generalization of the Kitagawa method.

Group 2: Use OB and refer to it as KOB Our investigation of the literature identified a second group of papers in which there were casual uses of the terminology linking K and OB and there was no mention of how the methodology was linked to the original Kitagawa method. We provide three examples below:

"To investigate in how far different compositions as opposed to different point estimates for different coefficients to these characteristics between men and women drive the GWG (gender wealth gap), we employ Kitagawa-Oaxaca-Blinder decomposition at different points of the unconditional wealth distribution (Kitagawa 1955; Blinder 1973; Oaxaca 1973; ...)." (Waitkus and Minkus (2021)) or for example the paper in World Development by Songsermsawas et al. (2023):

"We calculate gender- and ethnicity-specific productivity gaps by decomposing the project impacts into endowment effects (the portion of the productivity gap attributable to the average differences in the levels of observable covariates between the two groups) and structural effects (the portion of the productivity gap driven by deviations in the each groups return for covariates from the corresponding average return, or returns to endowment) by using the Kitagawa-Oaxaca-Blinder (KOB) decomposition (Kitagawa, 1955; Oaxaca, 1973; Blinder, 1973).

(...)

We start with a linear model as follows:

$y_{ig} = X'_{ig}\beta_{ig} + \varepsilon_{ig}$,

where y_{ig} is the outcome of interest for household i in each group g (in this case, non-minority vs. minority, and male-headed and female-headed households) calculated for the reference period between May 2017 and April 2018 (12 months preceding data collection), X'_{ig} is the vector of production inputs, endowments, and other socio-economic attributes for household i in each group g, and ε_{ig} is the error term."

and finally a paper by Chen et al. (2022) :

"Our analysis uses methods pioneered in economics by Kitagawa (1955), Oaxaca (1973), and Blinder (1973) that decompose the differences in the main outcome of interest between two groups or time periods into a part due to differences in characteristics ("compositional effects") and a part explained by differences in the associations with characteristics ("changing coefficients").

The standard Kitagawa-Oaxaca-Blinder linear decomposition is based on a linear regression framework and requires coefficient estimates from the linear regression and sample means of the explanatory variables. "

Although we show that the original Kitagawa method can be expressed as a linear regression, it is not correct to suggest that Kitagawa framed her methodology as a "linear regression." In particular, none of these papers convey any sense of how the original Kitagaw method and the OB decompositions differ in general.

5 The Way Forward (in the Naming Convention)

So, what guidance would be useful regarding application of the KOB label? The KOB label could be reserved for a decomposition exercise that decomposes group differences in an indicator outcome variable in which the explanatory variables/factors are categorical variables comprised of mutually exclusive sets of indicator variables. In this sense, the Kitagawa decomposition is identical to the OB decomposition involving OLS estimation of LPMs. What if a decomposition is OB without meeting the data restrictions of the Kitagawa decomposition? One could certainly argue that Kitagawa and OB are both decomposition methods. However, in this case the the original Kitagawa decomposition approach does not apply and cannot be implemented. Consequently, the designation "KOB" is not appropriate without mentioning that Kitagawa is an early version of the more general OB method. In other words, it should be incumbent upon authors to explicitly recognize that the Kitagawa method and the OB method are not identical if the more restricted data criteria are not satisfied.

At the end of the day, none of this detracts from the originality and prescient contribution of Kitagawa's path breaking work, which has been appreciated and popularized in demography and sociology many decades before it became of interest in Economics.

6 Conclusions

To our knowledge, our paper is the first explicit recognition of the exact relationship between the Kitagawa decomposition and OB decomposition methodologies. The two methodologies are identical only in the very specific set of circumstances characterized by binary dependent outcome variables, all covariates are categorical variables comprised of sets of mutually exclusive indicator variables, and the OB decomposition uses OLS estimated LPMs. As the OB and OB-related methodologies do not require the restrictions on the dependent and independent variables inherent in the original Kitagawa methodology, the OB decomposition methodology is more general. Furthermore, there are abundant examples of uninformed, casual, or misleading labeling of decompositions as KOB.

The cross-fertilization of fields, in particular with that of economics, has given a new wave of popularity to the Kitagawa name. Yet, in our opinion, it would not give the original Kitagawa method justice if it were cited unknowingly and for the wrong reasons. In this paper, we show how the two methods are related and provide a suggestion on the naming convention moving forward.

References

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

Chen, Y., Sylvia, S., Wu, P., and Yi, H. (2022). Explaining the declining utilization of village clinics in rural China over time: A decomposition approach. *Social Science & Medicine*, 301:1–9.

Das Gupta, P. (1993). Standardization and decomposition of rates: A user's manual. Report P23-186, United States Census Bureau.

Jaffe, A. J. (1951). *Handbook of Statistical Methods for Demographers: Selected Problems in the Analysis of Census Data*. U.S. Bureau of Census. Statistical Research Division. U.S. Government Printing Office.

Kitagawa, E. M. (1955). Components of a difference between two rates. *Journal of the American Statistical Association*, 50(272):1168–1194.

Kroger, H. and Hartmann, J. (2021). Extending the Kitagawa–Oaxaca–Blinder decomposition approach to panel data. *The Stata Journal*, 21(2):360–410.

Neumark, D. (1988). Employers' discriminatory behavior and the estimation of wage discrimination. *Journal of Human Resources*, 23(3):279–295.

Nieuwenhuis, R., Lancker, W. V., Collado, D., and Cantillon, B. (2020). Trends in women's employment and poverty rates in oecd countries: A kitagawa–blinder–oaxaca decomposition. *Italian Economic Journal*, 6:37–61.

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

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

Oaxaca, R. L. and Ransom, M. R. (1988). Searching for the effect of unionism on the wages of union and nonunion workers. *Journal of Labor Research*, 9:139–148.

Preston, S. H., Heuveline, P., and Guillot, M. (2001). *Demography: Measuring and modeling population processes.* Blackwell Publishers Ltd.

Songsermsawas, T., Kafle, K., and Winters, P. (2023). Decomposing the impacts of an agricultural value chain development project by ethnicity and gender in nepal. *World Development*, 168:1–15.

Waitkus, N. and Minkus, L. (2021). Investigating the gender wealth gap across occupational classes. *Feminist Economics*, 27(4):114–147.

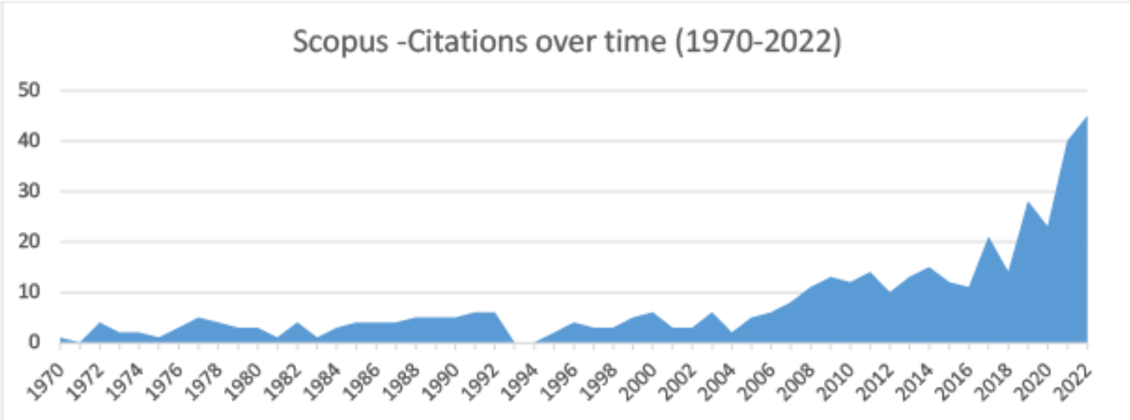
7 Tables and Figures

Table 1: Citations of Kitagawa (1955) over time by fields.

	Decades						Total	Rate increase in citations	
Field	1970-79	1980-89	1990-99	2000-09	2010-19	2020-23		2010-19	2020-23
Ecology/Geography	1	2	1	0	3	5	12		
Economics	0	2	0	1	7	17	27	7	2.4
Population/Demography	12	10	10	18	29	28	107	2	1.0
Social Science	8	3	3	8	13	23	58	2	1.8
Sociology/ Political Science	12	11	3	4	12	16	58	3	1.3
Medicine	0	12	8	11	27	10	68	2	0.4
Public Health	1	4	0	12	16	3	36	1	0.2
Statistics/ Quantative Methods	1	4	4	0	4	1	14		0.3
Other (e.g. Law, Science, Psychology)	0	0	3	2	4	3	12	2	0.8
Total	35	48	32	56	115	106	392		

Source: Web of Science (March, 2023)
Notes: Calculation based on journal classification available from authors upon request.

Figure 1: Citations of Kitagawa (1955) over time.



Appendix

A Example of the Kitagawa and Oaxaca-Blinder Decomposition

This appendix provides a detailed example of the conditions under which the Kitagawa and Oaxaca-Blinder decompositions are equivalent. The derivations below satisfy the data restrictions of the Kitagawa decomposition. At the level of the individual, the Kitagawa data restrictions are that the outcome variable of interest is an indicator variable, and the covariates are categorical variables comprised of sets of mutually exclusive indicator variables that sum to 1. We show that the Kitagawa decomposition is equivalent to the OB decomposition based on OLS-estimated linear probability models estimated with the same data.

Our example is a one factor (explanatory variable), two-component decomposition (Kitagawa 1955 p.1182). With some notational changes from the original Kitagawa paper, suppose there are two population groups, A and B, for which the group mean difference in a binary indicator outcome variable Y is to be decomposed (e.g. 0/1 not employed/ employed). The sample sizes for the two population groups are denoted by N^A and N^B. The one-factor, explanatory variable X is a categorical variable defined by a mutually exclusive set of K indicator variables such that $\sum_{k=1}^{K} X_{ik} = 1 \; \forall i$ in a group sample, where i refers to the ith individual (e.g. educational groups: no high school/ high school/ more than high school). Let Y_i^A and Y_i^B denote the outcome indicator variable for the ith individual in groups A and B.

The group mean proportions are given by

$$\overline{Y}^A = \frac{\sum_{i=1}^{N^A} Y_i^A}{N^A} = \frac{N_Y^A}{N^A}$$

$$\overline{Y}^B = \frac{\sum_{i=1}^{N^B} Y_i^B}{N^B} = \frac{N_Y^B}{N^B},$$

where N_Y^A, N_Y^B are the numbers of individuals in groups A and B for whom $Y_i = 1$. Accordingly, the group mean outcome proportion to be decomposed is given by

$$\overline{Y}^A - \overline{Y}^B = \frac{N_Y^A}{N^A} - \frac{N_Y^B}{N^B}.$$

Next, let $\overline{Y}_k^A$ and $\overline{Y}_k^B$ denote the outcome rates among individuals in the kth category of variable X for groups A and B:

$$\begin{aligned}\overline{Y}_k^A &= \frac{\sum_{i=1}^{N^A} Y_i^A X_{ik}}{N_k^A},\ k = 1, ..., K \\ &= \frac{N_{Yk}^A}{N_k^A}\end{aligned} \tag{1}$$

$$\begin{aligned}\overline{Y}_k^B &= \frac{\sum_{i=1}^{N^B} Y_i^B X_{ik}}{N_k^B},\ k = 1, ..., K \\ &= \frac{N_{Yk}^B}{N_k^B},\end{aligned} \tag{2}$$

where $N_{Yk}^A = \sum_{i=1}^{N^A} Y_i^A X_{ik}$ and $N_{Yk}^B = \sum_{i=1}^{N^B} Y_i^B X_{ik}$ represent the number of individuals in groups A and B for whom $(Y_i^A \cdot X_{ik}) = 1$ and $(Y_i^B \cdot X_{ik}) = 1$, or in other words, N_{Yk}^A and N_{Yk}^B are the number of individuals who fall in the kth indicator category and for whom the Y outcome indicator variable = 1. Now let N_k^A, N_k^b represent the numbers of individuals in each group for whom $X_k = 1$. Kitagawa shows that the overall outcome rate for each population group is equal to the sum of its own X specific outcome rates weighted by its own X composition $\left(\frac{N_k^A}{N^A}\right), \left(\frac{N_k^B}{N^B}\right)$:

$$\overline{Y}^A = \sum_{k=1}^{K} (\overline{Y}_k^A) \left(\frac{N_k^A}{N^A}\right)$$

$$\overline{Y}^B = \sum_{k=1}^{K} (\overline{Y}_k^B) \left(\frac{N_k^B}{N^B}\right).$$

Thus,

$$\overline{Y}^A - \overline{Y}^B = \sum_{k=1}^{K} (\overline{Y}_k^A) \left(\frac{N_k^A}{N^A}\right) - \sum_{k=1}^{K} (\overline{Y}_k^B) \left(\frac{N_k^B}{N^B}\right).$$

Kitagawa's two-component decomposition of the group difference in outcome rates consists of (1) a Gross X component which is that portion of the group outcome rate difference attributable to group differences in their X composition $\left(\frac{N_k^A}{N^A} - \frac{N_k^B}{N^B}\right)$, and (2) a Residual X component which is that portion of the group outcome rate difference attributable to group differences in their X specific outcome rates $\left(\frac{\overline{Y}_k^A + \overline{Y}_k^B}{2}\right)$ (Kitagawa 1955 p.1182):

$$\overline{Y}^A - \overline{Y}^B = \text{Gross } X + \text{Residual } X,$$

$$= \underbrace{\sum_{k=1}^{K} \left(\frac{\overline{Y}_k^A + \overline{Y}_k^B}{2}\right) \left(\frac{N_k^A}{N^A} - \frac{N_k^B}{N^B}\right)}_{\text{Gross X}} + \underbrace{\sum_{k=1}^{K} \frac{\left(\frac{N_k^A}{N^A} + \frac{N_k^B}{N^B}\right)}{2} \left(\overline{Y}_k^A - \overline{Y}_k^B\right)}_{\text{Residual X}}.$$

Note that the average of each group's X specific rates is used as the standard (reference) group outcome rate.

We now examine the OB regression-based decomposition. With the same data, the regression-based approach would specify linear probability models for the two population groups:

$$Y_i^A = \sum_{k=1}^{K} X_{ik}\beta_k^A + \varepsilon_i,\ i = 1, ..., N^A$$
$$Y_i^B = \sum_{k=1}^{K} X_{ik}\beta_k^B + \varepsilon_i,\ i = 1, ..., N^B.$$

Note that there is no separate constant term as the indicator variables sum to 1.

Because the indicator variables comprising X are mutually exclusive, the cross-product matrix for the OLS estimator of the β's is a diagonal matrix. This orthogonal design means that the OLS estimator simply corresponds to the separate simple regressions of Y_i

on each X_{ik} indicator variable:

$$b_k^A = \frac{\sum_{i=1}^{N^A} X_{ik} Y_i^A}{\sum_{i=1}^{N^A} X_{ik}^2}, \; k = 1, ..., K \tag{3}$$

$$= \frac{N_{Yk}^A}{N_k^A} = \overline{Y}_k^A.$$

$$b_k^B = \frac{\sum_{i=1}^{N^B} X_{ik} Y_i^B}{\sum_{i=1}^{N^B} X_{ik}^2}, \; k = 1, ..., K \tag{4}$$

$$= \frac{N_{Yk}^B}{N_k^B} = \overline{Y}_k^B.$$

We see that the OLS coefficient estimates correspond to Kitagawa's X specific outcome rates derived in eq(1) and eq(2). Furthermore, the sample indicator variable means correspond to the Kitagawa X compositions:

$$\overline{X}_k^A = \frac{\sum_{i=1}^{N^A} X_{ik}}{N^A}, \; k = 1, ..., K$$

$$= \frac{N_k^A}{N^A}$$

$$\overline{X}_k^B = \frac{\sum_{i=1}^{N^B} X_{ik}}{N^B}, \; k = 1, ..., K$$

$$= \frac{N_k^B}{N^B}.$$

Typically, one group is assigned as the reference group in the standard (OB) OLS decomposition. Here, we follow Kitagawa's counterfactual and adopt simple averages of the two groups's estimated coefficients, which is the same as the average of each group's K -

specific rates seen in the Kitagawa decomposition:

$$
\begin{aligned}
b_k^* &= \frac{(b_k^A + b_k^B)}{2},\ k = 1,...,K \\
&= \frac{(\overline{Y}_k^A + \overline{Y}_k^B)}{2}.
\end{aligned}
$$

Accordingly, the OLS decomposition is as follows:

$$
\begin{aligned}
\overline{Y}^A - \overline{Y}^B &= \sum_{k=1}^{K} \overline{X}_k^A b_k^A - \sum_{k=1}^{K} \overline{X}_k^B b_k^B \\
&= \underbrace{\sum_{k=1}^{K} \left(\overline{X}_k^A - \overline{X}_k^B\right) b_k^*}_{\text{Explained}} + \underbrace{\sum_{k=1}^{K} \overline{X}_k^A \left(b_k^A - b_k^*\right) + \sum_{k=1}^{K} \overline{X}_k^B \left(b_k^* - b_k^B\right)}_{\text{Unexplained}}.
\end{aligned}
$$

Note that simple algebra establishes the following relationship among the LPM coefficients: $\frac{b_k^A - b_k^B}{2} = b_k^A - b_k^* = b_k^* - b_k^B$. We now show with simple subsitution that the Kitagawa decomposition components correspond exactly to the OB decomposition components:

$$\underbrace{\sum_{k=1}^{K}\left(\frac{\overline{Y}_k^A+\overline{Y}_k^B}{2}\right)\left(\frac{N_k^A}{N^A}-\frac{N_k^B}{N^B}\right)}_{\text{Gross X}} = b_k^*\sum_{k=1}^{K}\left(\overline{X}_k^A-\overline{X}_k^B\right)$$

$$= \underbrace{\sum_{k=1}^{K}\left(\overline{X}_k^A-\overline{X}_k^B\right)b_k^*}_{\text{Explained}}$$

$$\underbrace{\sum_{k=1}^{K}\frac{\left(\frac{N_k^A}{N^A}+\frac{N_k^B}{N^B}\right)}{2}\left(\overline{Y}_k^A-\overline{Y}_k^B\right)}_{\text{Residual X}} = \sum_{k=1}^{K}\left(\frac{\overline{X}_k^A+\overline{X}_k^B}{2}\right)\left(b_k^A-b_k^B\right)$$

$$= \sum_{k=1}^{K}\overline{X}_k^A\left(\frac{b_k^A-b_k^B}{2}\right)+\sum_{k=1}^{K}\overline{X}_k^B\left(\frac{b_k^A-b_k^B}{2}\right)$$

$$= \underbrace{\sum_{k=1}^{K}\overline{X}_k^A\left(b_k^A-b_k^*\right)+\sum_{k=1}^{K}\overline{X}_k^B\left(b_k^*-b_k^B\right).}_{\text{Unexplained}}$$

Thus, Kitagawa's Gross K component portion of the group outcome rate difference attributable to group differences in their X composition is identical to the explained difference in the OB decomposition, and Kitagawa's Residual X component portion of the group outcome rate difference attributable to group differences in their X specific outcome rates is identical to the unexplained difference in the OB decomposition.

Under the special conditions shown above, it is straightforward to show that the equivalence between the Oaxaca-Blinder and Kitagawa decompositions holds whether the counterfactual standard corresponds to one of the two groups or is a weighted average.