

Proxy Inference Methods: survey of literature

Nahim Bin Zahur*
1st November 2009

Introduction

This is an exercise in studying proxy inference techniques- but that immediately begs the question- what is a proxy inference technique? The name suggests a link with the concept of a proxy variable, which can be defined as “a measurable variable that is used in place of a variable that cannot be measured.”¹ Thus, from a broad perspective, a proxy inference technique is simply a method of making an inference about a variable of interest using one or more proxy variables.²

However, given that we are primarily interested in studying and evaluating these methods as used in the field of targeting in social programs, our study of proxy inference techniques has to focus on those techniques that are of relevance to targeting, while excluding other less relevant methods. This is why we begin with an overview of proxy inference techniques as used in the fields of poverty assessment and targeting of social programs. As we shall see, a review of these methods will suggest that proxy inference methods or techniques, as defined above, in fact typically consist of two separate steps that are analytically very different from one another. Specifically, a proxy inference method involves firstly a *conceptualization step* in which a measurable variable is selected to proxy the primary variable of interest (e.g. the use of consumption per capita to approximate welfare)³, and secondly a *proxy step* in which a set of other variables correlated with the chosen measurable variable is selected and a suitable statistical technique is applied on this set of variables so as to proxy the measurable variable. The present study will primarily focus on the proxy step, although a review of the conceptualization step will also be included.

The proxy means test approach

In a proxy means test, according to Grosh and Baker, “information on household or individual characteristics correlated with welfare levels is used in a formal

* The author is a student at the National University of Singapore. The work presented in this paper is only the first part of a literature survey that was initiated during his internship with the Economic Research Group under Dr. Sajjad Zohir.

¹ Upton, Graham and Cook, Ian, *A Dictionary of Statistics*, 2nd edition (revised), (Oxford University Press, 2008).

² Usually, it involves using values of several variables, thus necessitating the use of the phrase ‘inference technique’.

³ This first step is of course unnecessary if the variable we are interested in happens to be measurable.

algorithm to proxy household income, welfare or need.”⁴ The motivation for proxy means tests comes from the fact that means tests used to select households for social welfare programs (where the variable used to represent welfare, such as per capita consumption expenditure, is measured directly) might be best in theory but in practice can be difficult to verify, quite often biased (since applicants to the social programs will tend to understate their level of welfare) and costly to administer in poor countries. Thus the idea of indirectly measuring the welfare level, by “proxying a means test”, suggests itself as an alternative.⁵

In their paper, Grosh and Baker argue that consumption expenditure rather than income should be used as the welfare measure for which a proxy measure is to be obtained.⁶ They then use household survey data from Jamaica, Peru and Bolivia to estimate the relationship between consumption expenditure and a number of household characteristics; they do this by using ordinary least squares (OLS) regression to predict consumption expenditure using selected characteristics. By varying the selection of right-hand side variables as well as by varying other parameters, they obtain a number of models which they evaluate using two main approaches- by comparing the degree of “leakage” and “under-coverage”, and by calculating a poverty index for each model and comparing the indices.⁷ They also discuss practical issues involved with an actual proxy means test program used in Chile, the Ficha CAS system (which incidentally uses principal components analysis rather than OLS regression- see page 4 for a discussion of PCA approaches).⁸

A large number of proxy means tests exist in the literature, most of which are similar to the above in that they define consumption expenditure as the welfare measure and then use data from household surveys to estimate the relationship between household consumption and various household characteristics using regression or some other statistical technique. The variables (i.e. household characteristics) chosen to proxy the welfare measure can vary among different models/programs, especially since the type of data available in household surveys is different in different countries, but the criteria for choosing these variables tends to be similar. At the very least, the variables have to be relatively few in number and well-correlated with poverty (as measured using the welfare measure). Additional considerations are that the variables have to be easy to answer and to measure/observe, easy to verify and at the same time difficult to manipulate by the

⁴ Grosh, Margaret E. and Baker, Judy L., “Proxy Means Tests for Targeting Social Programs: Simulations and Speculation,” *LSM118*, Vol. 1 (1995) p. ix.

⁵ *Ibid.*, p.1.

⁶ *Ibid.*, p.6.

⁷ Grosh and Baker, p. 5-22.

⁸ *Ibid.*, p. 23-26.

household (which will naturally have an incentive to do so in order to qualify for the social program that is using the proxy means test).⁹

There is a rich variety of statistical techniques available, and hence the choice of statistical technique can vary widely. As in the simulation by Grosh and Baker, OLS regression is commonly used, although sometimes with refinements added so as to improve performance. For example, Ahmed and Bouis developed a proxy means test for targeting food subsidies in Egypt in which they use OLS regression to find variables correlated with per capita expenditure, but with one difference from Grosh and Baker- they developed two models, one with per capita expenditure as the dependent variable and the other using the natural log of per capita expenditure as the dependent variable instead, and found that the second model was superior in explaining the variability in per capita expenditure.¹⁰ Naryan and Yoshida, in constructing a proxy means test to target welfare benefits in Sri Lanka, used a similar method where they carried out OLS regression with the dependent variable as the natural log of per capita consumption expenditure, since this was “found to work well in the regressions.”¹¹

While OLS regression has traditionally been the most common choice of statistical technique, other forms of regression are increasingly being utilised in carrying out proxy means tests. Chen and Schreiner, for instance, used logistic regression to estimate a “poverty scorecard” for Bangladesh; their scorecard also differed from the proxy means tests outlined above in that it was aimed not at national governments but at smaller local organizations.¹² In spite of that, their scorecard is conceptually similar to the proxy means testing approach given that it is still based upon finding a set of variables that can be used to proxy consumption expenditure per capita; as such we can safely classify the poverty scorecard method as a particular type of proxy means test.¹³ Linear probability and probit models, which are similar to logit models in that they try to estimate the probability that a household’s expenditure is below a specified poverty line (as opposed to OLS regression which seeks to estimate what the household’s consumption expenditure is), are also used. For instance, the IRIS center uses the linear probability model (in conjunction with a two-step approach i.e. one scorecard is used to eliminate

⁹ Coady, David; Grosh, Margaret and Hoddinott, John, *The Targeting of Transfers in Developing Countries: Review of Experience and Lessons*, (IFPRI, 2004), p. 35. See also, Chen, Shiyuan and Schreiner, Mark, *A Simple Poverty Scorecard for Bangladesh*, (Microfinance Risk Management, 2009), p. 3.

¹⁰ Ahmed, Akhter U. and Bouis, Howarth E., “Weighing what’s practical: proxy means testing for food subsidies in Egypt,” *Food Policy*, Vol. 27 (2002), p. 519-540.

¹¹ Naryaan, Ambar and Yoshida, Nobuo, “Proxy Means Test for Targeting Welfare Benefits in Sri Lanka,” *PREM Working Paper Series SASPR-7* (July 2005), p. 1-30.

¹² Chen and Schreiner, p. 1-164.

¹³ Poverty scorecards have been developed for a number of other countries as well- see http://www.microfinance.com/#Poverty_Scoring for a list of examples.

those households which are highly likely to be living above the poverty line, followed by a second scorecard which is applied to the remaining sample¹⁴) to construct poverty scorecards in four countries- Bangladesh, Guatemala, Madagascar and Tajikistan, while a probit-based scorecard using a one-step approach is developed for Timor Leste. However, the dominant method for the IRIS center seems to be scorecards based on quantile regression models, with a one-step or two-step approach depending on the country, since such scorecards are preferred for 15 out of the 20 countries.¹⁵

There are also approaches which try to proxy welfare measures other than consumption expenditure. Income is an obvious alternative, although as mentioned earlier consumption is generally regarded as a better indicator of welfare than income. Nonetheless, an income proxy model for use by the USAID mission in Kenya was set up by Tschirley and Mathenge, using OLS regression as the statistical method.¹⁶ The dataset used was an income survey conducted in Kenya in 2000, which contained data on a number of income components (such as income from wages/salaried labor, income from different types of farming activities etc) as well as data on household characteristics. A novel feature of their analysis is that they developed separate regression models for each of their identified income components, and then obtained the overall predicted income by summing the predicted values of each of the income components.

A third alternative is to construct a welfare index, instead of relying on consumption or income as the measure welfare. Such indices are developed by first selecting a number of indicator variables that are thought to be correlated with welfare (the choice of indicators depends on the definition of the welfare index- a welfare index purporting to measure health, for instance, will consist of different indicators as opposed to a wealth index). The next step is to assign weights to these indicators- while this can be done arbitrarily, typically a statistical technique such as principal components analysis or PCA is used. PCA is a technique used to identify patterns in multidimensional data and re-express the data in terms of these patterns.¹⁷ The procedure used to identify these patterns is to find an orthonormal set of vectors, such that the first vector chosen is in the direction of maximum variance, the second vector chosen will be in the direction orthogonal to the first

¹⁴ IRIS Center, "Note on Assessment and Improvement of Tool Accuracy," (June 2005), http://www.povertytools.org/other_documents/AssessingImproving_Accuracy.pdf Accessed 10 September, 2009.

¹⁵ IRIS Center, "Accuracy Results for 20 Poverty Assessment Tool Countries," (December 2008), http://www.povertytools.org/other_documents/PAT_20_country_accuracy_results_Dec2008.pdf Accessed 11 September 2009.

¹⁶ Tschirley, David and Mathenge, Mary, "Developing Income Proxy Models for use by the USAID Mission in Kenya: A Technical Report," *Tegemeo Working Paper 7* (November 2003) p. 1-10.

¹⁷ Smith, Lindsay I, "A tutorial on Principal Components Analysis," (February 26, 2002), p. 12. http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf Accessed October 17, 2009.

vector that contains the maximum remaining variation in the data, and so on for the rest of the vectors. This set is then used as the new basis for the data which is re-expressed as linear combinations of these vectors, thus illustrating the extent to which each of these vectors represents the variation in the data.¹⁸ How is PCA related to the problem of finding a way of assigning weights when constructing an index? After carrying out PCA on a dataset consisting of data on the indicator variables, typically the first principal component (i.e. the vector in the direction of maximum variance) is chosen as the index, and since it is a linear combination of the indicator variables, the required weights are easily obtained. The main assumption underlying the use of PCA is that in such datasets, the welfare measure explains the maximum variance in the indicator variables (which is why the welfare measure can be approximated by the first principal component).¹⁹

Chile's proxy means program, the Ficha CAS system referred to earlier, is one example of the above- first a socioeconomic index was derived by carrying out PCA on the Continuous Nutrition Status Survey, and then around ten variables were chosen that were easy to collect and would best predict this index (this two-step approach would be puzzling but for the fact that the socioeconomic index was originally developed for a different purpose, and had to be simplified for the purpose of the proxy means test, thus explaining the need for the second step which reduced the number of variables required).²⁰ Another example is the DHS Wealth Index which derives its name from the fact that it uses data from the Demographic and Health Survey (DHS) that is carried out in a number of countries. Rutstein and Johnson discuss the rationale behind this index in their paper²¹ - they argue against the use of income or consumption as a measure of economic status and instead argue that wealth would be a better indicator. Because the DHS lacks asset or debt measures that could be used to directly compute wealth, the authors use a relative definition of wealth instead (i.e. a household's wealth is measured *relative* to other households in the sample). They choose some indicator variables from the DHS which are thought to be related to households' economic status, as measured by their wealth, in a particular way- for instance as wealth increases, the proportion of households owning TVs increases while the proportion of households owning motorcycles increases initially but decreases as wealth increases even further. They then use PCA on these variables to construct a

¹⁸ Shlens, Jonathon, "A tutorial on Principal Component Analysis," Version 3.01 (April 22, 2009), p. 5-6. <http://www.sn1.salk.edu/~shlens/pub/notes/pca.pdf> Accessed October 17, 2009.

¹⁹ Filmer, Deon and Pritchett, Lant H., "Estimating Wealth Effects Without Expenditure Data—Or Tears: An Application to Educational Enrollments in States of India" *Demography* Vol. 38, No. 1 (February 2001) p. 116-117. <http://www.bsos.umd.edu/socy/vanneman/socy699J/FilmerP01.pdf> Accessed 31 October 2009.

²⁰ Grosh and Baker, p. 24.

²¹ Rutstein, Shea Oscar and Johnson, Kiersten, "The DHS Wealth Index" *DHS Comparative Reports No. 6*, (Calverton, Maryland: ORC Macro, 2004), p. 1-14. <http://www.measuredhs.com/pubs/pdf/CR6/CR6.pdf> Accessed 13 September 2009.

wealth index that can be used to estimate the wealth of a household given information on the selected indicator variables.

One reason for the divergence between the PCA-based DHS wealth index, and the regression-based models discussed earlier, is the fact that their objectives are quite different. As Chen and Schreiner point out, the wealth index can only be a good proxy for long-term economic status;²² indeed, one of the arguments put forward by Rutstein and Johnson for their wealth index is that wealth is a more permanent measure of economic status than either income or consumption, which both tend to be volatile. Because the objective of the DHS wealth index is to decide how to distribute health services among the population, long-term economic status is more appropriate compared to short-run measures, while the fact that relative wealth is being measured is not a problem.²³ The regression-based models, however, are typically used to target cash transfers or other types of social programs in an effort to alleviate poverty; thus, both an absolute rather than relative measure of poverty is required, and the poverty measure has to consider short-term economic status as well since the short run directly affects the households' welfare level as much as the long run.

What this serves to show is that the criteria for evaluating proxy means tests have to be sensitive to the underlying objectives, and the criteria used for programs with differing objectives must necessarily be different at least in some respects. One further instance of this is that Grosh and Baker assume that the proxy means test would be used to identify households as poor and non-poor after which cash transfers would be targeted to the poor; thus they evaluate their simulation by considering how effective the proxy means test approach is at reducing poverty.²⁴ Chen and Schreiner, on the other hand, mention targeting as only one of the possible functions of their poverty scorecard- they mention that the scorecard can also be used for estimating the poverty likelihood of a household (i.e. "the probability that a household has per-capita expenditure below a given poverty line), for estimating the poverty rate of a group (where poverty rate is defined as "the average poverty likelihood among the households in the group") and for monitoring the change in poverty rates for a group over time.²⁵ When monitoring poverty is the goal, the criteria for judging a method is simply to judge how effective the method is at estimating poverty accurately (whereas in Grosh and Baker's formulation, accurate monitoring is secondary to the underlying goal of *reducing* poverty). Indeed, the name chosen by the IRIS Center for these methods, "poverty assessment tools", highlights the monitoring aspect of proxy means tests, and as such the criteria they use to choose between alternative methods is

²² Chen and Schreiner, p. 14-15.

²³ Rutstein and Johnson, p. 3-4, p. 7.

²⁴ Grosh and Baker, p. 9, p. 13.

²⁵ Chen and Schreiner, p. 3-4.

primarily the accuracy of the methods, which they evaluate using a measure called the BPAC or the Balanced Poverty Accuracy Criterion.²⁶

The BPAC is defined as the Poverty Accuracy minus the absolute difference between Under-coverage and Leakage rates, where poverty accuracy is the proportion of poor households who are correctly identified as poor, under-coverage rate is the proportion of poor households who are incorrectly identified as non-poor while the leakage rate is the number of non-poor households incorrectly defined as poor, expressed as a percentage of the total number of poor households. To make sense of the above measure, it is necessary to note first that though both under-coverage and leakage errors are inevitable with any proxy inference method, it is also often the case that decreasing the under-coverage rate will typically increase the leakage rate, and vice versa. For instance, in Grosh and Baker's simulation, increasing the cutoff poverty line (i.e. those households estimated below this line will be identified to be poor, while those above will be identified as non-poor) decreases leakage but at the expense of under-coverage.²⁷ Thus, assuming that both types of errors are equally not to be preferred, the accuracy measure has to first make sure that the two errors are roughly similar in size, so that one type of error is not minimized at the expense of the other; this is the rationale for including the absolute difference between under-coverage and leakage errors. It also has to ensure that both the errors are low in magnitude (which is done by including poverty accuracy, since poverty accuracy shares an inverse relationship with the two kinds of errors). Poverty accuracy, rather than total accuracy (which is simply the proportion of all households whose economic status is correctly identified) is used because it is poverty incidence and not the poverty status of the entire population that is of interest.²⁸

Poverty mapping

An approach that is related to proxy-means tests in many ways but is distinct enough to be considered a separate field in its own right is poverty mapping. The basic idea of a poverty map is to provide information on the extent of poverty in various geographical areas in a country. Generally these areas are small enough (in area or population) that having a single value to represent their level of poverty has some value, yet at the same time they have to be large enough so that sufficient data exists to estimate their level of poverty. When poverty mapping is used for the purpose of targeting services, the approach is known as geographical targeting.²⁹ While there are many types of poverty maps, there is one important general difference between proxy-means tests and poverty mapping- while

²⁶ IRIS Center (2005).

²⁷ Grosh and Baker, p. 16.

²⁸ IRIS Center (2005), p. 1-8

²⁹ Coady, Grosh and Hoddinott, p. 45

consumption is the dominant measure of welfare in the field of proxy-means tests, this is not the case when it comes to poverty mapping³⁰. A basic needs approach has been a popular alternative to using consumption expenditure as the measure of welfare, with “poverty defined in terms of access to basic services.”³¹

A fairly comprehensive review of poverty mapping is provided by Davis,³² where he identifies (among others) two major approaches to poverty mapping that are particularly relevant to our discussion of proxy inference methods. The first is the small area estimation approach, which is in some ways an extension of the consumption-based proxy means approach. The direct proxy-means tests approach is inappropriate for poverty mapping because it would involve using household surveys to construct estimated poverty rates for each geographical region, yet for some regions the sample size may be small giving rise to large standard errors that would render the poverty map unreliable. In such situations, the statistical technique called small area estimation can be used, where the model “borrows strength from related or similar small areas.”³³ In the context of poverty mapping, this means combining the data from household surveys with some other dataset (usually a census) which contains a sufficiently large sample for each of the “small areas” (i.e. geographical areas). How this can be done is described by Davis.³⁴ First, regression analysis is used to construct a model to estimate consumption expenditure (or some other welfare measure) based on both household-level characteristics as well as geographical-level characteristics. The regression model is then applied to the census data to calculate the probability that each household is living below the poverty line. Aggregating these probabilities for each geographical region of interest, the level of poverty in each region (measured by some poverty index such as the poverty gap or the poverty headcount³⁵) can be estimated. A variant of this method is when community-level data (rather than household-level data) is combined with the census data to construct the poverty map.

Davis also discusses the multivariate weighted basic-needs index approach,³⁶ which has similarities to the PCA-based indices approach discussed earlier. The welfare measure that is used to represent poverty is a basic-needs index which, as

³⁰ The dichotomy between geographical targeting and other forms of targeting posited in the literature appears to be artificial; and demands critical assessment.

³¹ *Ibid.*, p. 46.

³² Davis, Benjamin, *Choosing a Method for Poverty Mapping*, (Food and Agricultural Organization of the United Nations: Rome, 2003). See p. 5-27 for discussion of types of poverty mapping. <http://www.povertymap.net/publications/doc/CMPM%20DAVIS%2013%20apr03%20sec.pdf> Accessed 13 September 2009.

³³ Ghosh, M. and Rao, J. N. K., “Small Area Estimation: An Appraisal,” *Statistical Science*, Vol. 9, No. 1 (1994) p. 55-56.

³⁴ Davis, p. 5-12.

³⁵ See Ghosh and Baker, p. 10, Box 1 for a discussion of poverty measures.

³⁶ Davis, p. 12-16.

the name suggests, is the result of a weighting scheme applied to a number of variables which represent basic needs. Examples of such variables in Mexico's PROGRESA program (which uses a basic needs approach to geographical targeting) are percentage of households without access to electricity, without access to drainage, with access to running water and the household over-15 illiteracy rate.³⁷ Such indices can be constructed using PCA (using the procedure described earlier-one illustration is the PROGRESA program referred to above), factor analysis and even OLS regression.³⁸

Other approaches to poverty mapping include direct measurement, either from household surveys using proxy-means tests (if the sample sizes for each geographical area are large enough, in which case small area estimation is not required); or even from census data (either if income data is directly available in the census, in which case income is used to measure welfare; or if there are variables in the census that can be combine to create a basic needs index). There are approaches that combine qualitative data with secondary data, as well as participatory approaches which extrapolate local assessments of poverty.³⁹ Finally, relatively simple forms of geographical targeting exist which target a single variable that is used to represent the basic needs; this approach is useful when the objective of the program is more congruent with the single variable chosen rather than a broader measure of welfare.⁴⁰

Reexamining the concept of proxy inference methods

The above summary of the proxy means tests and poverty mapping represents the range and type of methods that we are studying under the heading of proxy inference methods. Before expanding to other fields that use similar methods in order to carry out a comparative analysis, however, it is necessary to reexamine the concept of proxy inference methods in light of the above review.

One issue that stands out with respect to this is the fact that proxy means tests actually involve two separate steps. We have already identified one of these steps, which is the proxying of welfare measures such as consumption using a set of other variables. But the thing that stands out from the review of proxy means testing approaches, and particularly when it comes to poverty mapping approaches, is *that there is no universal agreement on what is the true welfare measure that can indicate whether a household is poor or not*. Consumption expenditure, income, wealth indices and basic needs indices have all been used as a welfare measure. This suggests that before we can even talk about the statistical

³⁷ Coady, Grosh and Hoddinott, p. 46.

³⁸ Davis, p. 12-16.

³⁹ Ibid., p. 17-28 contains a discussion of these alternative methods.

⁴⁰ Coady, Grosh and Hoddinott, p. 46.

approach to proxying, we have to make a decision about the conceptualization of welfare i.e. which of these measures do we choose as a proxy for welfare? Only after we have carrying out this first *conceptualization step* can we move on to the next *proxy step* in which we proxy the selected measure.

Should we consider both of these steps in our study of proxy inference methods? I will argue that it is the proxy step that we should primarily focus on. This has to do with the purpose of this exercise- we are trying to estimate one unknown variable from a number of known variables, and there is a range of approaches available to us. So we need a set of criteria that can help us choose such an approach, and these criteria have to be general i.e. applicable beyond the fields we have been looking at. However, the choice of which welfare measure to use (the conceptualization step) is largely a matter of judgment using criteria that is probably applicable only to the field of developmental studies, and thus the comparative analysis that we attempt to carry out here will not help in making that choice. This is not to suggest that the conceptualization step is unimportant, but that the focus of the current exercise is on the second proxy step, although the question of conceptualization will also be considered.

Thus, from that perspective, we can define the proxy step of a proxy inference method as a method of making an inference about a *measurable* variable of interest using one or more proxy variables. This excludes the conceptualization step- because there is no universal agreement as to the measure of poverty, we can consider poverty or welfare as variables that cannot be measured, at least when contrasted with variables such as consumption expenditure which can be measured.

A model of the decision-maker's problem

Before concluding it is useful to model the problem faced by the decision-maker attempting to choose among various proxy inference methods. The decision-maker has an underlying *purpose* that we can assume is exogenously determined- for instance the purpose could be to target certain services to the poor. The decision-maker also has a number of available *choices* that could be used to achieve the purpose; each of these choices consists of the conceptualization choice (i.e. the choice of welfare variable) and, dependent on the first choice, a choice of the proxy method (which encompasses things such as the choice of statistical technique, the selection of indicator variables and a host of other choices). In deciding from among these choices, the decision-maker will seek to optimally achieve his purpose, and this he can do by using a set of *criteria* to evaluate competing choices. Possible criteria can be accuracy measures (for instance the BPAC measure), costs and feasibility, and are of course by their very definition dependent on the purpose.

The goal of this exercise therefore is to identify the possible choices (something which has been partially achieved in the above review), develop the set of criteria and then evaluate the choices based on these criteria, keeping in mind that the outcomes obtained are conditional on the purpose of the decision-maker. The fact that the purpose is exogenous does not present a major problem to this formulation, due to two reasons. First of all, in both the proxy means testing field and the field of poverty mapping, there does not appear to be a significant degree of variation when it comes to the purpose or objective of the decision-makers. Secondly, one goal of this exercise is to develop some criteria that are *independent* of the purpose, so that the conclusions based on those criteria will be universal and not just applicable for a narrow range of purposes.

Conclusion

This paper has summarized the two broad approaches- proxy means tests and poverty mapping- that we are interested in studying, and has provided a definition of proxy inference methods which can be utilized in order to select those approaches that would be suitable for the current exercise. The next paper will contain a descriptive analysis of the methods and approaches outside the realm of developmental studies that can be classified as proxy inference methods, followed by a classification of all the methods under consideration so that we can start evaluating them.

For comments and suggestions, please contact Nahim at nahimzahur@nus.edu.sg