SINCE 1956

QUALITY. RELEVANCE. IMPACT

NCAER National Data Innovation Centre

Student Grantee Report Number 05,

July 2020

A Simplified Measure of Nutritional Empowerment

Using Machine Learning to Abbreviate the Women's Empowerment in Nutritional Index (WENI)

Shree Saha and Sudha Narayanan

A SIMPLIFIED MEASURE OF NUTRITIONAL EMPOWERMENT: USING MACHINE LEARNING TO ABBREVIATE THE WOMEN'S EMPOWERMENT IN NUTRITIONAL INDEX (WENI)

Shree Saha

PhD Scholar, Indira Gandhi Institute of Development Research, Mumbai

shree@igidr.ac.in

Sudha Narayanan

Associate Professor, Indira Gandhi Institute of Development Research, Mumbai

sudha@igidr.ac.in

Acknowledgement and Disclaimer: This research is supported by the National Council of Applied Economic Research through the NCAER National Data Innovation Centre. The views presented in this report are those of the authors and not those of NCAER or its Governing Body. Funding for the NCAER National Data Innovation Centre is provided by Bill & Melinda Gates Foundation.

Suggested citation: Saha, Shree and Sudha Narayanan (2020). "A Simplified Measure of Nutritional Empowerment: Using Machine Learning to Abbreviate the Women's Empowerment in Nutritional Index (WENI)", *Student Grantee Report Number 05*, Report submitted to NCAER National Data Innovation Centre, New Delhi: National Council of Applied Economic Research.

A simplified measure of nutritional empowerment

Using machine learning to abbreviate the Women's Empowerment in Nutrition Index (WENI)*

Shree Saha, Sudha Narayanan

Indira Gandhi Institute of Development Research (IGIDR), Mumbai, India.
e-mail: shree@igidr.ac.in

Draft Paper - Please do not circulate or cite

July 14, 2020

Abstract

Measuring empowerment is both complicated and time consuming. A number of recent efforts have focused on how to better measure this complex multidimensional concept such that it is easy to implement. In this paper, we use machine learning techniques, specifically LASSO, to abbreviate a recently developed measure of nutritional empowerment, the Women's Empowerment in Nutrition Index (WENI) that has 33 distinct indicators. Our preferred Abridged Women's Empowerment in Nutrition Index (A-WENI) consists of 20 indicators. We further validate the A-WENI via a field survey in a new context, the western Indian state of Maharashtra. We find that the 20-indicator A-WENI is both capable of reproducing well the empowerment status generated by the 33-indicator WENI and predicting nutritional outcomes such as BMI and dietary diversity. Using this index, we find that in our survey sample, on average, only 51.2% of mothers of children under the age of 5 years are nutritionally empowered, whereas 86.1% of their spouses are nutritionally empowered. We also find that only 22.3% of the elderly women are nutritionally empowered. These estimates are broadly consisted with those based on the 33-indicator WENI. The A-WENI will reduce the time burden on respondents and can be incorporated in any general purpose survey conducted in rural context. Many of the indicators in A-WENI are often collected routinely in contemporary general purpose household surveys and hence capturing nutritional empowerment does not entail significant additional burden. Developing A-WENI can thus aid in an expansion of efforts to measure nutritional empowerment; this is key to understanding better the barriers and challenges women face and help identify ways in which women can improve their nutritional well-being in meaningful ways.

Keywords: empowerment, nutrition, machine learning, LASSO, gender, South Asia, WENI

JEL Classification: J16, D63, I00, C55

*This project was funded by a Student Research Grant of the Data Innovation Centre, National Council Of Applied Economic Research (NCAER), New Delhi, India.We thank, without implicating, Linden McBride, Abhinav Motheram and Santanu Pramanik for their valuable inputs The field work was conducted with the support of Pragati Abhiyan, Ashwini Kulkarni, in particular. We thank Mr. Gal Reddy Chitty for his crucial technical support in the field.

1 Introduction

The United Nations has identified the achievement of gender equality and empowerment of all women and girls as one of the Sustainable Development Goals. While the notion of equality is somewhat straightforward and perhaps better understood, empowerment is harder to define. A widely accepted definition is based on Kabeer (1999), who characterizes empowerment as an individual's capacity to make strategic choices in life when such capacity has been absent previously. Empowerment is often conceptualized as encompassing a number of different aspects such as agency, autonomy, self-direction, self-determination, liberation (Ibrahim and Alkire, 2007; Malapit et al., 2019; Kabeer, 1999). An individual's agency itself refers to multiple aspects including, for instance, intrinsic agency (power within), instrumental agency (power to) as well as collective agency (power with) and influence (power over). Others emphasize that the opportunity structure faced by individuals and an individual's access to resources are key elements of empowerment (Kabeer, 1999; Alsop and Heinsohn, 2005). Empowerment is thus complex, dynamic and multidimensional. While there is no consensus on a single definition of empowerment, there is currently a significant congruence of views on what it represents. For example, most agree that empowerment is multidimensional and includes not just agency but the terms on which resources can be accessed and norms that govern such access as well as knowledge of these. There is also agreement that it is a process and empowerment in some spheres can occur alongside disempowerment in other spheres. It is important too to distinguish empowerment in specific realms, economic, health and nutrition, for example, from a generalized notion of empowerment.

While empowerment is desirable for its own sake, it can also have instrumental value in forwarding the well-being of individuals, especially crucial for women, both in terms of economic and non-economic attributes such as health and nutrition (Kadiyala et al., 2014). Economists, for instance note that although economic development and women's empowerment go hand in hand and that development typically leads to empowerment, it is important nevertheless to design policy explicitly for empowering women (Duflo, 2012). The call for gender equality and empowerment of all women and girls is therefore a call for policy action geared explicitly towards empowering women and girls.

Yet, our ability to track progress on women's empowerment, to examine its relationship with women's wellbeing and to assess the empowering effect of public policy depends crucially on our ability to reliably measure empowerment (Alsop and Heinsohn, 2005; Malhotra et al., 2005). Current evidence suggests a large number of measures that vary based on the concept or definition of empowerment, the domain in which empowerment is studied (economic empowerment, empowerment in agriculture, empowerment in livestock, nutritional empowerment, etc.), methods used in the selection and aggregation of indicators and data used (Pratley, 2016; Laszlo and Grantham, 2017; Pereznieto and Taylor, 2014). Reviews note for example that there are over 40 measures for economic empowerment and as many as 181 measures for empowerment relating to health and nutrition (Pratley, 2016; Laszlo and Grantham, 2017). Further, critics argue that many of these measures are not sufficiently well-grounded in theory. Many are based on data reduction techniques that often select one variable at the expense of others because they are all correlated, although they represent very different aspects of empowerment. Still other measures are not comprehensive enough for the purpose it is supposed to serve (Richardson, 2018; Malhotra et al., 2005; Heckert and Fabic, 2013). Apart from debates on how best to measure empowerment, a related challenge in measuring empowerment is that, owing to its complexity, it is typically cumbersome to implement (Alsop and Heinsohn, 2005). This often deters efforts to incorporate empowerment measures in general surveys of individuals and households, which capture instead just a few proxy indicators of empowerment.

There have been several recent attempts to develop simplified (sets of) indicators of empowerment. Ibrahim and Alkire (2007), for example, have focused on identifying an internationally comparable set of indicators that are relatively easy to collect. Others have created useful libraries of specific survey questions that can elicit information on different dimensions of empowerment (Laszlo and Grantham, 2017; Glennerster et al., 2018). Still others have tested whether the measures of empowerment are appropriate for the context and captures the underlying concepts well (Hannan et al., 2020; Yount et al., 2019). Somewhat differently, developers of the Women's Empowerment in Agriculture Index (WEAI), an important and widely used index for rural communities, have designed a more user-friendly Abbreviated Women's Empowerment in Agriculture Index (A-WEAI) that reduces interview time by 30% as compared with the WEAI (Malapit et al., 2019).

This paper contributes to these efforts, responding to the need to develop leaner indices of empowerment that reduce the time and cost burden associated with data collection, especially in resource-constrained settings. Our focus is on a specific measure of empowerment, the Women's Empowerment in Nutrition Index (WENI). Developed by an interdisciplinary team of researchers, the WENI aims to capture empowerment in the realm of nutrition in rural contexts (Narayanan et al., 2019). The WENI, in its original form, was developed in the context of rural south Asia and comprises 33 indicators, collected via a special purpose survey. Our effort in this paper is to develop a leaner version of this index with fewer indicators that can be more easily incorporated in general purpose household surveys for rural contexts. Our effort is motivated by the data gaps in current large scale surveys that capture neither women's nutrition nor their empowerment status particularly well. An abridged WENI (henceforth A-WENI) in our view would contribute to filling this gap.[1]

The central challenge in abbreviating any index is to be able to simplify the measure without compromising on either the richness or the spirit of the original index and achieving this balance with minimal procedural subjectivity. As Richardson (2018) cautions, data reduction techniques can often conflate conceptually distinct indicators merely because they are correlated and this can

[1]The choice of the name is intended to parallel WEAI's evolution and is analogous to A-WEAI - the abbreviated Women's Empowerment in Agriculture Index.

undermine the conceptual basis and richness of such indicators. In general, abbreviating an existing index involves selecting a subset of indicators in ways that do not affect the overall performance of the index (Malapit et al., 2019). Current efforts to minimize the number of indicators captured in an aggregate empowerment index typically exclude a subset of indicators that are highly correlated with other preferred indicators that are already in the index, using Principal Components Analysis, (Confirmatory) Factor Analysis and so on. In this paper, we propose the use of machine learning or statistical learning techniques, in particular, supervised learning techniques, as a reliable, objective approach to simplify complex, multidimensional, multi-indicator indices of empowerment.

The application of machine learning in the area of economic development has expanded in recent years. Many of the applications thus far have focused on variable selection for obtaining better estimates or predictions of poverty (Afzal et al., 2015; McBride and Nichols, 2018) or food insecurity (Lentz et al., 2019), for example. In the choice of indicators for constructing an index, Kshirsagar et al. (2017) apply LASSO techniques to identify variables that go into an index to be able to identify the poor and hence facilitate programme targetting. Machine learning approaches have so far not been explored in the context of complex multidimensional indicators such as empowerment.

Drawing on data used to create the 33-indicator WENI, we apply LASSO (least absolute shrinkage and selection operator), a popular machine learning technique to identify the candidate constituent indicators of A-WENI. We then validate it afresh out of sample, implementing a survey in rural Maharashtra, India, for this purpose. Our findings suggest that the WENI can be abridged to 20 indicators, while remaining faithful to the concept of nutritional empowerment and still being able to reproduce well the nutritional empowerment status of the 33-indicator WENI. Further, the A-WENI remains a good predictor of nutritional status of individuals, including BMI and dietary diversity scores, as with the original WENI. In this paper, we detail the process by which machine learning techniques can contribute to developing sophisticated yet simple measures of empowerment, based on more complex preexisting measures and elaborate on our findings from the validation of A-WENI in rural Maharashtra.

This paper is organized as follows. In Section 2 we describe in detail the WENI as it was originally developed. In Section 3 we elaborate on the machine learning technique we propose to use and then apply it to identify, develop and validate a simplified measure of nutritional empowerment, i.e. A-WENI. We then validate this abridged index in a field sites in Nashik, Maharashtra and examine if it predicts nutritional outcomes (Section 4). Section 5 deals with robustness checks and Section 6 concludes the discussion, highlighting the usefulness and limitations of A-WENI.

2 The Women's Empowerment in Nutrition Index (WENI)

The Women's Empowerment in Nutrition Index (WENI) was created in response to a perceived need for a measure of empowerment that is salient to women's nutritional well-being, as opposed to generic measures of empowerment that don't often pertain to nutrition or correlate well with nutritional status (Narayanan et al., 2019).[2] The effort also explicitly shifts the focus of the relationship of women's empowerment to their own nutritional status rather than that of their children. The WENI project focused on conceptualizing nutritional empowerment and then developing and validating WENI specifically for the rural South Asian context.

Accordingly, nutritional empowerment is defined as "the process by which individuals acquire the capacity to be well fed and healthy, in a context where this capacity was previously denied to them"(Narayanan et al., 2019, p. 2). This process entails "acquiring knowledge about, and a say over, nutritional and health practices; gaining access to and control over intake of adequate and nutritious food; and being able to draw support from both family and other institutions to secure and maintain an adequate diet and health"(Narayanan et al., 2019, p. 2).

2.1 WENI: Women's Empowerment in Nutrition Index

To operationalize this conceptualization of nutritional empowerment, the WENI researchers amalgamate insights from two distinct streams of literature - the literature on women's empowerment, drawing heavily on Kabeer (1999), including four dimensions - knowledge, resources, agency and achievements - and the literature on the drivers of nutrition following the UNICEF (1990) framework for child nutrition and identifying three domains - food, health and institutions - that are salient for women's nutrition).[3] Narayanan et al. (2019) then propose a WEN Grid to organize the domains and dimensions into a matrix to guide the identification of factors that constitute nutritional empowerment. WENI is conceived of as a metric that aggregates measures of these factors into a single number and is described in detail below.[4] In its original formulation, nutritional outcomes like BMI and anaemia, as also dietary diversity, are considered (nutritional) achievements relating to nutritional empowerment but not explicitly a part of WENI itself. Nutritional achievements are therefore used only to validate WENI and to assess the associative strength between nutritional empowerment and nutritional outcomes.

It is useful to note that the WENI seeks to capture the status of an individual at a particular point of time, even though nutritional empowerment is conceptualized as a process. The idea is that WENI can be measured at different points of time, enabling us to track progress in eliminating the

[2]The Women's Empowerment in Nutrition Index (WENI) project was funded by Competitive Research Grants to Develop Innovative Methods and Metrics for Agriculture and Nutrition Actions (IMMANA Grants) of the United Kingdom's Department for International Development/UK Aid and an award from the Policy Research Institute, the Lyndon B. Johnson School of Public Affairs at University of Texas, Austin.

[3]While Kabeer (1999) defines empowerment in the dimensions of resource, agency, and achievements, where knowledge is a part of resources, to operationalize nutritional empowerment, Narayanan et al. (2019) considers knowledge as a separate dimension, given its importance in deciding the nutritional achievements.

[4]While the WENI, an index based on this Grid allows us to track progress in nutritional empowerment and is a good predictor of nutritional outcomes, the WEN Grid itself provides insights by domain-dimension (DD) which can be a helpful diagnostic tool in identifying areas that require policy action. For more on this, see (Narayanan et al., 2019).

barriers to empowerment. Another feature of WENI is that those it focuses on women specifically, the empowerment measure created can be used to identify empowerment status of any adult.[5]

2.2 Computing WENI

The WENI is constructed using 33 indicators covering the seven DD (Table 1 lists these). These 33 indicators straddle several themes and typically there are multiple indicators representing different aspects that capture a single theme (described in detail in subsequent paragraphs). To compute WENI, each indicator is first converted to a binary variable, where 1 represents being empowered with respect to the specific indicator and 0 otherwise. A detailed discussion of these are found in Narayanan et al. (2019), also available in the Toolkit associated with this project (reference to come) and are therefore not presented here. A score is then computed for each DD indicating the proportion of indicators on which the individual is deemed to be empowered for that DD. The DD-specific score thus ranges between 0 and 1. These DD-specific scores are then averaged over the seven DDs to generate the index scores, weighting each DD equally. This score, the WENI, thus ranges between 0 and 1 (both inclusive). A cut-off for the aggregated index, 0.5 in this case, is set and on the basis of the cut-off, individuals who have scores less than 0.5 are classified as nutritionally disempowered and those with scores above the cut-off are classified as empowered. [6]

The WENI was originally constructed and validated in five Indian states - Bihar, Odisha, Tamil Nadu, Kerala and West Bengal over the years 2017-18 in two phases. The WENI was first developed for Bihar and Odisha. Starting with survey-based data, for 971 individuals, comprising approximately 128 indicators representing various DD as part of the WEN Grid, 33 indicators were eventually selected to be part of the WENI, using a mix of statistical, normative and qualitative analysis.

2.3 WENI Indicators and Themes

As mentioned earlier, WENI indicators for an individual span seven domain-dimensions or DDs - namely food-knowledge (FK), food-resources (FR), food-agency (FA), heath-knowledge (HK), health-resources (HR), health-agency (HA) and institutions (I). Recall that these are constructed by combining three domains (food, health and institutions) and three dimensions(knowledge, resources and agency). Narayanan et al. (2019) consider institutions as a separate domain incorporating factors such as legal rules, general community norms, not pertaining specifically to food, health or fertility, that represent nutritional empowerment.

[5]WENI also introduces an additional domain of fertility for women in the reproductive age between 15 and 49, as it is a critical aspect of nutritional empowerment for young women. In this paper, we do not incorporate Fertility and focus instead on indicators that are relevant for all individuals.

[6]At the DD level, a similar cutoff of 0.5 indicates whether an individual is empowered or disempowered in that specific DD. Several issues around the sensitivity of the WENI to cutoffs and the form of indicators are addressed in Narayanan et al. (2019) and are beyond the scope of this paper.

These indicators were chosen based on formative desk and field-based qualitative research to represent salient themes that capture different aspects of nutritional empowerment (Narayanan et al., 2019). Table 1 presents the list of indicators used in constructing WENI and also gives the list of themes by DD.

The food and health knowledge (FK and HK, respectively) DDs consists of indicators measuring knowledge of nutrition and health. The food-agency is composed of two separate themes, say in productive activities and control over income and expenditure. The DDs of food-resources (FR) and health-resources (HR) comprises three and four themes respectively. Indicators are categorized under each of these themes. For example, in the health-resource (HR) DD, the theme "Support when ill and health seeking "is measured by whether an individual has sought treatment when ill and whether they get assistance when they are sick. The indicators in the domain of institutions (I) capture different social and legal norms and they range from whether the individuals are members of any group, whether they receive information about government schemes or faces restrictions on movement to the individual's participation in public spheres (for example, speaking in public or contesting elections). In this paper, we maintain the classification of indicators into DDs and its themes as in the original WENI so that these are treated as inherited and fixed.

Based on these variables, a 33-indicator WENI (Table 1) can be computed for all individuals; the information is gathered via a special purpose survey that takes between 30 and 90 minutes to administer. [7] Given the number of indicators to compute the index, the burden on respondent's time (especially young mothers who are simultaneously involved in child care and domestic chores) is quite heavy. This is true of most empowerment surveys, for example the pilot 1.1 of WEAI took about 62 minutes in Bangladesh Malapit et al. (2019). The survey time and complex nature of data collected are likely serious deterrents to uptake and can result in the systematic exclusion of these types of measures from household surveys. Our attempt therefore is to remedy this problem and we use the 33-indicator WENI as a starting point for our work to abridge the index.[8]

[7]For young women in the reproductive age, the WENI uses an additional domain for Fertility with 10 more indicators (thus, a total of 43 indicators for WENI women). In this paper, however, we focus on generating A-WENI for any individual and therefore do not explicitly address the issue of fertility for younger women.

[8]Although our goal was to reduce survey time, we do not explicitly include the time burden of items as a formal criterion in the process of abridging the WENI.

Table 1: List of Indicators in Original WENI and their Description

No.	Variable name	Themes	Variable Description
1	FAagrisay	Say in productive activities	Has some say in agricultral activities(=1)
2	FAmajminsayent		Atleast some say in major or minor household enterprise decisions(=1)
3	FAassearnconsentown	Control/say in income and expenditure	Earnings from asset owned by respondent has not been used without consent
4	FAcashcontrol		Has cash as independent source of money & some control over how to spend it(=1)
5	FAdecisionpaidwrkbin		Decision to undertake or not undertake paid work own (=1)
6	FAnorestorwillstop		Faces no food restriction or from own will or can stop when wants(=1)
7	FKcalcium		Knowledge on calcium
8	FKiodine		Knowledge on iodine
9	FReatlast	Eating norms	Eat last only rarely/sometimes/never(=1)
10	FRfinsupportagrihhenter	Access to support and to assets	Some financial support/aid in HH enterprise or agriculture(=1)
11	FRjobandpds		Someone in household has both jobcard and rationcard(=1)
12	FRland		Own land in your name
13	FRpaidwork	Participation in income generating activities	Does paid work as employee(=1)
14	FRselfemployment		Has farm/non farm own employment(=1)
15	FRsourceincomediversitybin		Atleast 3 diverse sources of income out of 5 (=1)
16	FRwrkoutalone		Social norms permit women to work outside the village, alone
17	HAalonefortreatment		Can go alone to health centre for treatment if need be(=1)
18	HAhdecideownhealth		Can make decision on own health(=1)
19	HAhealthvisitpermission		No expectation to take permission from family before visiting health centre(=1)
20	HKanemia		Knowledge on anemia
21	HKmalaria		Knowledge on malaria
22	HKors		Knowledge on ors
23	HRdrinktoivent	Access to improved water, sanitation and smoke free kitchen	Household has access to improved water & improved toilet & ventilation(=1)
24	HRhassistwhensick	Support when ill and health seeking	Get assistance in household chores when ill(=1)
25	HRhoursmktworkbin	Work/energy expenditure and working conditions	Marketable work (paid or hh enterprise) > 8 hours (=1)
26	HRriskinjhealth		No risk of injury or major health problem in any activity(=1)
27	HRintensityany	Support in work	Does no paid/unpaid activity that is back breaking or heavy(=1)
28	Ianymemberownaccord		Member of any group with own accord(=1)
29	Idoveil		Never practice Ghonghat/Burkha/Pallu/Purdah(=1)
30	Ifreedommove		Has mobility to go to bank or post office or family alone(=1)
31	Imobileinformationgovt		Individual gets information on government schemes from mobile phone(=1)
32	Inoviolenceorsupport		Experiences no physical abuse or if does, has support within family (=1)
33	Iparticipatedany		Politically active in any activity in past 5 years(=1)

FA-Food Agency; FK-Food Knowledge; FR- Food Resources;HA-Health Agency; HK-Health Knowledge; HR- Health Resources; I- Institution

3 Designing an Abridged WENI

Our aim is to create a leaner WENI, with fewer indicators without compromising on its ability to reproduce the nutritional empowerment status of the 33-indicator WENI, while also predicting nutritional outcomes. The survey module of such an abridged index can then be more conveniently incorporated into a general purpose survey for rural communities and help measure nutritional empowerment with few additional resources. The process of eliminating some indicators rather than others however needs to have a sound basis and as far as possible, devoid of subjectivity. At the same time, it should remain faithful to the normative rationale for the choice of these indicators and consistent with the original 33-indicator WENI. We propose to use machine learning techniques, elaborated in the next section, to identify a subset of indicators that will comprise the A-WENI.

The abridged set of indicators then undergoes several sensitivity analyses and validation tests on the data from five states such that it reproduces the nutritional empowerment status generated by the 33-indicator WENI as closely as possible. The use of nutritional empowerment status as the variable we seek to predict or reproduce deserves justification and the reader might wonder why we don't choose a subset of indicators that predicts nutritional status itself. First, the central goal is to identify an alternative, leaner A-WENI that reproduces the empowerment status predicted by the 33-indicator WENI. By design therefore, A-WENI, which is derived from WENI, should lead us to identify whether an individual is nutritionally empowered the way the parent index WENI would. Further, using nutritional status as the outcome to predict would move us away from the largely normative framework of the original WENI to a purely instrumental approach that identifies those indicators that predict nutritional status. Thus, as with the original WENI, nutritional status does not factor in explicitly in the designing the index and is used only for the purpose of validation.

Once a plausible A-WENI is identified, we then validate it afresh with new data from Nashik, Maharashtra in western India. Our proposed methodology consists of several steps: indicator selection based on a set of pre-determined criteria, constructing the abridged index using the selected indicator list, predicting performance of the reduced set of indicators on existing data and validating the index based on data from a new field site. Finally, we test the ability of the abridged index to predict nutritional status in the new field site, comparing it to the 33-indicator WENI. The following sections discuss each of these steps in detail.

3.1 LASSO techniques

While there are several ways to reduce the number of indicators or variables, for example, principal components or factor analysis, backward and forward regression, we use the concept of supervised machine learning algorithms. Though this technique is used frequently in computer sciences, genome studies and financial markets, they have not been used much in the field of development, until recently.

Machine learning techniques are usually involved with the task of making predictions and can be classified broadly into two categories - unsupervised learning and supervised learning. The former tries to uncover the data structure based on association and classification, without any prior knowledge of the data. The latter, in contrast, predicts the outcome based on past events (Wuest et al., 2016; Schrider and Kern, 2018).[9] The advantage of machine learning over other statistical methods arises from its accurate predictions, ability to deal with high-dimension data along with its ability to simulate data in the absence of actual data. Most importantly machine learning allows the use data as they are "in nature, rather than in a way we represent it in a model"(Schrider and Kern, 2018). [10]

Supervised machine learning serves to predict future outcomes based on what has been learnt from the past (or alternatively predicting outcomes outside a sample based on what has been learnt from a particular sample). Starting the analysis from a known dataset called the training data, the algorithm creates an inferred function that then predicts the outcome/future events. We can compare this output with the actual and modify the model accordingly to get better results or predictions. Once the model is trained adequately on the training data, it can predict an outcome for any new data input. Even though there are several supervised machine learning algorithms, we use a technique called LASSO (Least Absolute Shrinkage and Selection Operator) for certain reasons.[11]

Unlike other machine learning algorithms, LASSO can perform both variable selection and regularization in order to improve the predictability and interpretation of statistical models it produces (Tibshirani, 1996). Similar to the classical Ordinary Least Squares (OLS) which minimizes the sum of squared deviations between observed and model predicted values, the LASSO in addition imposes penalty if coefficients are far from zero (Ahrens et al., 2020) (refer to equation 1). The lasso minimizes the mean squared error subject to a penalty on the absolute size of coefficient estimates.

To implement LASSO, we estimate the following model on the data from two states Bihar and Odisha.

$$\hat{\beta}_{Lasso}(\lambda) = \arg\min \frac{1}{n} \sum_{i=1}^{n} \left(nut\_emp_i - x_i'\beta \right)^2 + \frac{\lambda}{n} \sum_{j=1}^{p} \varphi_j \left| \beta_j \right| \tag{1}$$

[9] The examples of the unsupervised machine learning include principal component analysis, clustering etc. (Schrider and Kern, 2018).

[10] Schrider and Kern (2018) claims that even with simulation of data, machine learning produces models that are more robust to model misspecification in comparison to the traditional models.

[11] In addition to LASSO, we also use the Random Forest algorithm and elastic LASSO to select indicators as a robustness check. Section 5 discusses this in detail.

Here, nut_emp_i is the individual's nutritional empowerment status and X_i is the vector of 33 indicators. The tuning parameter λ controls the penalty level and ϕ_j are predictor-specific-penalty loadings. Supervised learning allows us therefore to reproduce the nutritional empowerment status constructed on the basis of the 33-indicator WENI.

Unlike the OLS, however, the LASSO imposes a penalty on the absolute size of coefficients estimates. It shrinks some coefficients and sets others to zero (Tibshirani, 1996; Ahrens et al., 2020). This way LASSO reduces model complexity and assists in feature selection, while keeping all predictors in the model. This feature of LASSO is contrast with forward/backward selection models where we are unable to find the impact of the removed variable on the outcome. However, like all regularized regression methods, LASSO too relies on tuning/penalty parameters that control the degree of penalization. In this case, there are three approaches to choosing the penalty level (λ); we discuss briefly each of these methods in the following paragraphs.

- *Data driven approach (CVLasso)*: This is the classical approach using cross-validation method of re-sampling data to optimize out-of-sample prediction performance. Also known as the k-fold cross validation, this method involves partitioning a dataset into approximately equal k-folds and estimating the model on all modules (training set) except the k^{th}-fold which is treated as the validation set. Predictive performance for a range of λs is assessed using the validation data. This process is repeated till all modules (thereby all data points) have been used for validation once. However, since the model is estimated k times, this approach is computationally intensive and time consuming. This method is useful for small datasets which are difficult to partition into testing and validation sets.

- *Information criterion approach (LASSO with IC)*: According to this approach, typically the selection of λ can be made using different information criteria like the Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (AICc), Bayesian Information Criterion (BIC) and Extended Bayesian Information Criterion (EBIC). Though the computation of the information criteria is easy, data-driven, and its theoretical properties well known, they are less robust to violations of independence and homoscedasticity assumptions. [12]

- *Theory driven approach (RLasso)- "Rigorous" penalization*: The name rigorous penalization is rooted in its strong theoretical framework. This approach requires three conditions to be satisfied to guarantee that a LASSO is consistent in terms of prediction and parameter estimation. Rigorous lasso provides the additional benefits of dealing with heteroscedasticity and

[12] The information criteria can be classified based on loss efficiency and selection efficiency. The former refers to the criterion that selects the smallest average squared error attainable by all candidate models. Model selection criterion on the other hand requires the true model to be selected with the probability tending to 1 as n tends to infinity. Thus, the former aims at prediction at and the latter at identification of true model. The AIC falls under the first category while the BIC in the latter. However, both of them perform poorly if the number of independent variables are more than the number of observations. To address this shortcoming, AICc and EBIC was constructed.

places high priority on controlling over-fitting and thus produces very parsimonious models. But given the focus on controlling overfitting, the cross-validation model may overtake in terms of prediction tasks.[13]

3.2 Selecting indicators

In this paper, we implement LASSO using survey data from five states that were used to construct and validate the 33-indicator WENI. The Bihar and Odisha dataset (where the WENI was first tested), consisting of 971 individuals, forms our training dataset and the data from the other three states, Tamil Nadu, Kerala and West Bengal (1427 individuals) forms our validation set. The original index is computed based on the procedure discussed in Section 2.2, using 33 indicators. The nutritional empowerment score so computed ranges from 0 to 1 and individuals are classified as nutritionally empowered if they have a score of 0.5 and above. Those who are nutritionally empowered are assigned 1 and those who are not are assigned a value of 0.

Using LASSO, we then identify a subset of indicators (from among the 33 indicators used to compute WENI) which 'best'predicts the original nutritional empowerment status of individuals, as generated using the 33-indicator WENI. In this paper, we implement LASSO using all the three approaches of tuning the λ parameter and display the results below. For each of these approaches, we compute both in-sample and out-sample Root-Mean Squared Error (RMSE). We choose the list of indicators that is identified by the method which has the lowest out-sample RMSE and in addition covers at least 50% of the themes from each DD. This additional criterion that indicators selected to be part of the A-WENI have to cover at least 50% of the themes in each DD, though subjective, aims to retain the spirit of WENI and prevents the index from being lopsided (Table 1 lists these 'themes' in each DD, which in turn are represented by indicators).

As a first step, we estimate LASSO models based on default penalty parameters. We find that using default penalty parameters, the indicator list for Rlasso, CVLasso contains 16 and 32 indicators respectively. While for Lasso with information criterion like AIC. AICC, BIC and EBIC the number of indicators chosen using the default penalty parameters is 32, 31, 27 and 27 respectively (Table 2; Table 3). We find that for most cases (except RLasso) the number of indicators is quite large. Even though RLasso generates the smallest indicator list (16 indicators), it fails both the criterion of lowest RMSE and our preference that the abridged set of indicators should include at least 50% of the 'themes' in each DD (listed in Table 1).

[13]The two main parts of the rigorous lasso is the restricted eigenvalue condition and the penalization level. The first condition offers respite from the full rank condition required in OLS context, which can be a problem in case of high dimensionality of data. For the second condition, λ has to be large enough to control all noise in the data but at the same time has to be small enough to avoid shrinkage bias.

Table 2: Performance of Indicators

	RLasso	CVLasso	Lasso/aic	Lasso/aicc	Lasso/bic	Lasso/ebic	OLS
Default Indicator Performance							
No. of Selected regressors	16	32	32	31	27	27	33
No. Dropped regressors	17	1	1	2	6	6	0
In sample RMSE	0.29	0.26	0.26	0.26	0.27	0.27	0.27
Out of sample RMSE	0.33	0.31	0.31	0.31	0.31	0.31	0.31
Sample Size	971	971	971	971	971	971	971
20 Indicator List Performance							
No. of Selected regressors	20	20	20	20	20	20	-
No. Dropped regressors	13	13	13	13	13	13	-
In sample RMSE	0.28	0.28	0.28	0.28	0.28	0.28	-
Out of sample RMSE	0.32	0.32	0.32	0.32	0.32	0.32	-
Sample Size	971	971	971	971	971	971	-

Note: The training set is Bihar and Odisha. Out of sample RMSE is calculated on the three short survey states with a sample of 1427 individuals.

The penalty level (λ) decides the number of indicators that get chosen. To generate an indicator list where the number of indicators is fixed, we adjust the penalty levels. In this case, we adjust the penalty levels for all the three approaches such that all of them generate a 20 indicator list (Table 2). [14] [15] We find that all the approaches generate the same set of indicators and covers at least 50% of the themes in each DD. Even though the indicator list is similar for all the three approaches, we prefer the cross-validation Lasso technique, due to its more rigorous approach of choosing variables and its better performance (as compared with Lasso technique based on information criterion) in small datasets. A detailed list of all the indicators generated across the different approaches is presented in Table 3.

Additionally, we also adjust the penalty levels for all the three approaches to generate a fifteen indicator list to compute WENI. Though the indices created based on these lists perform well in terms of validation (Table 3), they do not cover 50% of themes in each DD, diluting the conceptual basis for WENI and therefore not considered further.

3.3 The 20-indicator A-WENI

Based on the above, our preferred candidate A-WENI consists of 20 indicators (Table 3) that by design, a subset of the 33 indicators used in constructing the original index (Table 1). Those 33 indicators involved 59 unique questions is the survey, while only 28 questions to derive the 20

[14] All the four Information Criterion (AIC, AICC, BIC, EBIC) also generate the same 20 indicators as CVLasso and RLasso.

[15] This type of approach of either fixing the penalty level to achieve a required number of variables or enforcing that a model always includes a certain variable is common (See Kshirsagar et al. (2017) as an example, to understand how the LASSO estimation can be programmed such that certain variables are always included in the model).

indicators in A-WENI. In this sense, although we did not explicitly use time or number of questions as explicit criteria, our machine learning exercise appears to have identified indicators that require fewer distinct questions overall.

The abridged list of indicators represents a mix of questions, many of which are already captured in several nationally representative surveys relating to health, such as the National Family Health Survey (NFHS) or general purpose surveys such as the India Human Development Survey (IHDS). To illustrate, we compare the questions used to generate indicators for A-WENI with those covered in IHDS-2 and NFHS-4. While IHDS-2 includes some questions related to knowledge of health (reasons for occurrence of malaria, anemia and usage of ORS) NHFS-4 does not ask explicitly about respondents knowledge of food and health.[16] Both IHDS and NFHS have some form of questions related to control of income/cash, access to land, permission to visit health facilities alone (or with permission) (both the surveys capture all these indicators, though some of them are in posed differently than the WENI survey).[17] In addition to these indicators, like WENI, IHDS-2 also enquires about eating norms, a question missing from the NFHS-4.

However, there remains a small list of indicators that feature in WENI, but not in the surveys we compare (5 indicators in IHDS-2 and 7 in NFHS-4). The list varies between NFHS-4 and IHDS-2. Like the NFHS-4, the IHDS-2 does not include questions regarding knowledge of calcium, iodine, but NFHS-4 also additionally lacks information regarding the knowledge of anemia amongst individuals. Additionally, the IHDS-2 does not enquire about the availability of assistance when individual is sick, or whether the individual receives information related to government schemes on their phones. The IHDS-2 also does not explicitly enquire about domestic violence at respondent's home, but asks about the common practice in the village, the WENI questionnaire however, asks this information specifically of all individuals. The NFHS-4 on the other hand does not enquire about the eating norms, individual's membership in any group (for example self-help groups, political or religious group etc.), about the practice of veiling amongst women and the respondent's say in agricultural activities. [18]

[16]For example, NFHS-4 tests for presence of iodine in salt, where as WENI explicitly asks whether the respondent knows about iodine.

[17]Examples of questions posed differently include, for example, whether the individual has freedom to visit bank/post-office or family. The IHDS-2 asks it differently without explicit mention of bank/post-office. Other examples include knowledge of anemia, while the WENI questionnaire explicitly enquires about the individuals knowledge of anemia, the IHDS-2 asks if the woman suffered from it during her last child-birth. Similarly while the IHDS questionnaire asks if the individual uses a phone for SMS/email, WENI explicitly asks if information related to government schemes are received through mobile phones. Even in terms of eating norms, while WENI asks every individual whether they eat last in the family, IHDS enquires regarding the usual eating practices in the family, example whether everyone eats together or men eat earlier than women.

[18]The indicator 'HRdrinktoivent'in WENI, is a combination of household and individual response. While drinking water sources and ventilation in kitchen is asked at the HH level, type of toilet facility used is asked at the individual level to account for systematic differences there might be in the use of the toilet within the household. In the construction of original WENI, this combination was chosen based on factor analysis done in reducing the initial number indicators from 128 to 33, over others that attempted to capture all components at the individual level. It is possible that due to the high correlation of HH responses with individual responses, the former was picked up by the factor analysis. To maintain comparability, in this paper we follow the same steps of indicator construction as

Given the substantial overlap of questions with other surveys, we expect the measurement of nutritional empowerment via A-WENI would entail little extra time or effort and can be incorporated in these surveys.

We now use the list of 20 indicators generated using CVlasso, (Table 3) to compute A-WENI, using the same methodology for index construction as elaborated in Narayanan et al. (2019). Even though both CVLasso and Lasso with Information criterion give the same results in terms of the indicators chosen as well as RMSEs, we prefer to use the indicators identified via CVLasso due to its superior prediction abilities. Further, because the cross-validation under CVLasso uses subsets within the data, such that all data points have been used as a test data at least once, it is a preferred alternative.

Once the A-WENI is generated, we identify those who are nutritionally empowered and those who are not, based on A-WENI. We then validate A-WENI using data from the five states.

WENI. However, we recommend that this indicator can be constructed at an individual level, where possible.

Table 3: Indicator List

		Default Case		15-Indicator List		20 Indicator List
Sl No	Original WENI	Rlasso	Lasso BIC/EBIC	RLasso	CVLasso/Lasso IC	Rlasso/CVLasso/Lasso-IC
1	FAagrisay	FAagrisay	FAagrisay	FAagrisay		FAagrisay
2	FAassearnconsentown	FAassearnconsentown	FAassearnconsentown			FAassearnconsentown
3	FAcashcontrol	FAcashcontrol	FAcashcontrol	FAcashcontrol	FAcashcontrol	FAcashcontrol
4	FAdecisionpaidwrkbin		FAdecisionpaidwrkbin			
5	Famajminsayent		FAmajminsayent			
6	FAnorestorwillstop		FAnorestorwillstop			
7	FKcalcium	FKcalcium	FKcalcium	FKcalcium	FKcalcium	FKcalcium
8	FKiodine	FKiodine	FKiodine	FKiodine	FKiodine	FKiodine
9	FReatlast	FReatlast	FReatlast	FReatlast	FReatlast	FReatlast
10	FRfinsupportagrihhenter		FRfinsupportagrihhenter			
11	FRjobandpds		FRjobandpds			
12	FRland	FRland	FRland	FRland	FRland	FRland
13	FRpaidwork					
14	FRselfemployment					
15	FRsourceincomediversitybin					
16	FRwrkoutalone					
17	HAalonefortreatment	HAalonefortreatment	HAalonefortreatment	HAalonefortreatment	HAalonefortreatment	HAalonefortreatment
18	HAhdecideownhealth	HAhdecideownhealth	HAhdecideownhealth	HAhdecideownhealth	HAhdecideownhealth	HAhdecideownhealth
19	HAhealthvisitpermission	HAhealthvisitpermission	HAhealthvisitpermission	HAhealthvisitpermission	HAhealthvisitpermission	HAhealthvisitpermission
20	HKanemia	HKanemia	HKanemia	HKanemia	HKanemia	HKanemia
21	HKmalaria	HKmalaria	HKmalaria	HKmalaria	HKmalaria	HKmalaria
22	HKors	HKors	HKors	HKors	HKors	HKors
23	HRdrinktoivent	HRdrinktoivent	HRdrinktoivent	HRdrinktoivent	HRdrinktoivent	HRdrinktoivent
24	HRhassistwhensick		HRhassistwhensick			HRhassistwhensick
25	HRhoursmktworkbin		HRhoursmktworkbin			
26	HRintensityany					
27	HRriskinjhealth		HRriskinjhealth			
28	Ianymemberownaccord		Ianymemberownaccord		Ianymemberownaccord	Ianymemberownaccord
29	Idoveil	Idoveil	Idoveil		Idoveil	Idoveil
30	Ifreedommove		Ifreedommove	Idoveil		Ifreedommove
31	Imobileinformationgovt	Imobileinformationgovt	Imobileinformationgovt	Imobileinformationgovt	Imobileinformationgovt	Imobileinformationgovt
32	Inoviolenceorsupport		Inoviolenceorsupport			Inoviolenceorsupport
33	Iparticipatedany					
Total Number of Indicators	33	16	27	27	15	20

Note: CVLasso/Lasso AIC Default chooses all indicators except Iparticipatedany. Lasso AICC Default chooses all Indicators except Ianymemberownaccord and Iparticipatedany.

3.4 Validating the A-WENI

The index constructed using 20 indicators, the A-WENI, must be validated within sample, before it can be used as a measure of nutritional empowerment. Validating the abridged index is a two-fold process; first we test if the new index on an average classifies individuals as nutritionally empowered or disempowered 'similar' to the original. Second, we test whether the abridged WENI is a good predictor of nutritional outcomes like the original index. The results of these two tests will determine the usability of the abridged index as a comparable substitute of the original 33-indicator WENI. As mentioned above, the A-WENI will be considered a reliable measure of nutritional empowerment only when it ranks individuals as empowered or disempowered in a way similarly to the original WENI. To test this, we compute the rank order correlation between the abridged and original WENI nutritional empowerment index, which is a continuous score, and use Kendall's tau-b to address issues related to ties in ranking. We find that the rank-order is preserved in the abridged WENI (Table 4). We also find a high positive correlation of 0.94 between the original nutritional empowerment score (WENI), and that of the abridged version (A-WENI), the continuous variable on which the thresholds are imposed for identifying whether an individual is nutritionally empowered or disempowered.

Table 4: Area under the ROC curve and Rank order correlation between the abridged WENI and the original.

	Rlasso-15	CVLasso-15/ Lasso with IC-15	20 - Indicator List
ROC Curve			
Area under the ROC curve	0.87	0.87	0.88
Standard error	0.01	0.01	0.01
Lower bound	0.86	0.86	0.87
Upper bound	0.88	0.88	0.9
Rank Order			
tau-a	0.37	0.37	0.38
tau-b	0.74	0.74	0.78
p-value	0	0	0
Kendall score	1054090	1058334	1091346
se-score	29142.6	29230.58	28520.65
Observations	2398	2398	2398

Note: The sample is the existing data from 5 states.

To further test whether the abridged WENI classifies individuals correctly into nutritionally empowered and disempowered categories (based on original WENI), we conduct a Receiver operating characteristic (ROC) analysis (Table 4). The ROC analysis quantifies the accuracy of diagnostic tests used to discriminate between two states/conditions. This discriminatory accuracy of a diagnostic test is measured by its ability to correctly classify observations into their actual states/conditions. In this case, we examine if A-WENI correctly classifies an individual into being nutritionally empowered or disempowered as determined by the 33-indicator WENI. We find that 88.29% of the classifications into empowered or disempowered by the abridged WENI are correct (Table 4).

Finally, we test if the abridged WENI index is a good predictor of nutritional outcomes. We test this using data from the five original field sites. We estimate a least squares (for continuous BMI, for the subsample who are not overweight or obese) and Probit regression model (for the binary variable, where normal BMI (between 18.5 and 25) is coded as one and underweight is coded as zero.[19] Additionally, we also estimate a probit regression to examine if the nutritional empowerment leads to higher probability of achieving the minimum dietary diversity.[20] We estimate the following models:

$$BMI_i = \alpha_0 + \beta_0 A\_WENI_i + \epsilon_i \tag{2}$$

$$Pr(18.5 < BMI < 25)_i = \alpha_1 + \beta_1 A\_WENI_i + \delta_i \tag{3}$$

$$Pr(MDD = 1)_i = \alpha_2 + \beta_2 A\_WENI_i + \zeta_i \tag{4}$$

Our dependent variable is nutritional outcome, measured by BMI (continuous, binary status(indicating whether or not a person has a normal BMI) and a logarithmic transformation) and our explanatory variable is the A-WENI. All the regressions have robust standard errors that correct for heteroscedasticity.[21] We also control for the physiological status of individuals, i.e., whether or not the WENI woman was pregnant at the time of interview, demographic group (relationship status, whether WENI woman, spouse, etc.) and age. The regression excludes household and individual level controls, because systematic differences in household socio-economic status (especially like wealth and education) in principle should reflect in the WENI variables themselves, adequately if not fully.

We find that A-WENI has significant positive correlation with BMI levels and minimum dietary diversity, indicating that higher values of nutritional empowerment are associated with better BMI (Table 5) and higher probability of the individual meeting the minimum dietary diversity norm.

---

[19] The sample for BMI as a continuous is restricted to those who are not overweight/obese(BMI<25), this is because we care about whether nutritional empowerment leads to less under-nutrition (Narayanan et al., 2019).

[20] The MDD is based on (FAO, 2016) The minimum dietary diversity (MDD) is computed as 1 if the individual consumes at least 5 out of the ten food groups. For the MDD estimations, we use the full sample of individuals regardless of their BMI status.

[21] In case of a larger sample, we can use the robust cluster errors, where we cluster at the village level, however, due to small data size, and only 7 villages (clusters) where the survey was purposively done, we refrain from using robust cluster errors.

Table 5: : Relationship between nutritional empowerment and nutritional status

	BMI	Normal BMI (=1)	Log BMI	Minimum Dietary Diversity (=1)
Abridged WENI Nutritionally empowered (=1)	0.688***	0.346***	0.035***	0.364***
	(-5.04)	(-4.18)	(-5.08)	(-4.22)
State (Kerala=1)	-0.332*	-0.550***	-0.020**	-0.577***
	(-1.66)	(-4.38)	(-1.97)	(-6.66)
State (West Bengal=1)	0.515***	-0.086	0.024***	-1.640***
	(-2.93)	(-0.64)	(-2.81)	(-16.30)
State (Odisha=1)	-1.078***	-0.662***	-0.053***	-2.476***
	(-5.69)	(-5.07)	(-5.64)	(-17.16)
State (Bihar=1)	-1.111***	-0.632***	-0.055***	-2.425***
	(-6.43)	(-5.10)	(-6.36)	(-18.97)
Spouse (=1)	0.504***	0.259**	0.026***	0.025
	(-3.39)	(-2.54)	(-3.49)	(-0.26)
MIL(=1)	0.149	0.038	0.008	0.233
	(-0.46)	(-0.19)	(-0.49)	(-1.41)
Older Woman(=1)	0.091	0.042	0.006	-0.276
	(-0.23)	(-0.17)	(-0.3)	(-1.39)
Age (completed years)	0.004	0.003	0	-0.006
	(-0.47)	(-0.49)	(-0.36)	(-1.23)
Constant	20.395***	0.886***	3.009***	0.609***
	(-69.35)	(-4.5)	(-202.62)	(-3.51)
R-squared	0.113	-	0.111	-
Adjusted R-squared	0.109	-	0.106	-
Chi-sq	-	112.626	-	794.084
N	1783	1783	1783	2342

Note:* significant at 10%, ** significant at 5%, *** significant at 1%. We control for different demographic groups by using dummies

4 Validating the A-WENI in a new context

Overall, we find that the abridged WENI not only correctly classifies individuals as being nutritionally empowered or disempowered, it is also a good predictor of nutritional outcomes. However, since the analysis is done using existing data, we validate the A-WENI in a completely new context, in Nashik, Maharashtra, where we conduct a survey focused specially on nutritional empowerment. Using data from this new survey we validate the both the classification and prediction power of our novel index.

4.1 The Survey

To conduct the survey, we collaborated with Pragati Abhiyan, a rights-based civil society organization that works in Nashik district in the western Indian state of Maharashtra. We surveyed 516 individuals in Nashik district, Maharashtra. The choice of the site was driven by the fact that the WENI had not been validated in the western part of India. The survey was held in February, 2020, and covered 13 villages spanning 5 administrative blocks in Nashik. We selected a mix of tribal and non-tribal villages to include communities with diverse social norms and resource constraints.[22] The sample was selected to include diverse contexts and should not be construed as representative

[22] The WENI project had covered sites from states in northern (Bihar), eastern (Odisha, West Bengal) and southern India (Kerala and Tamil Nadu).

of the region surveyed. As part of the survey, we interviewed young mothers with children below the age of five (209 individuals), a smaller sample of their male spouses and mother-in-laws (101 and 103 individuals respectively) and elder women above the age of 70 (103 individuals). This tablet-based survey was conducted in the local language (Marathi) and prior to the launching of the main survey we conducted extensive pretesting and pilot surveys to ensure that the questions were clearly framed, specific and understandable.

We use the same survey instruments as those used to validate the WENI in Kerala, Tamil Nadu and West Bengal, enabling us to compute both the original WENI with 33 indicators and the abridged WENI with 20 indicators. We implement the abridged WENI on the data collected in this survey to measure nutritional empowerment as well as to predict the nutritional outcomes. We use BMI and MDDS as measures of nutritional status, as with the earlier exercise for the states of Kerala, Tamil Nadu and West Bengal. The BMI is a widely accepted measure that is easy to compute and appealing since it covers the entire spectrum of both under and over-nutrition.[23]

4.2 The status of nutritional empowerment

Using the A-WENI we now calculate the prevalence of nutritional disempowerment across individuals.[24] We use the same methodology to compute the A-WENI index, as in the original WENI (Narayanan et al., 2019).

The patterns we find using A-WENI are broadly similar to those with the 33-indicator WENI and is broadly similar too to the patterns from the other states as well.[25] We find significant variation in index scores based on WENI woman, their spouses, mother-in-laws (MIL) and other older woman. While 51.2% of WENI women in the sample are nutritionally empowered, 86.1% of their spouses and 60.2% of MILs are nutritionally empowered (Table 6). However, only 22.3% of older woman are empowered. We find therefore that older woman fare much worse than other demographic groups. Table 6 and Figure 2 gives a clear picture of the status of the four groups across the DDs.

[23] As mentioned earlier, we focus primarily on the under-weight and normal BMI i.e. BMI < 25, as we focus, as with WENI, on the link between nutritional empowerment and under-nutrition (Narayanan et al., 2019). However, the same results hold for the full sample as well.

[24] We also compute the original index (with 33 indicators) to be able to compare the results with the abridged version (using 20 indicators).

[25] A comparison of different state-samples is beyond the scope of this paper.

Figure 1: Cumulative Distribution Function of the Original and Abridged WENI in Maharashtra

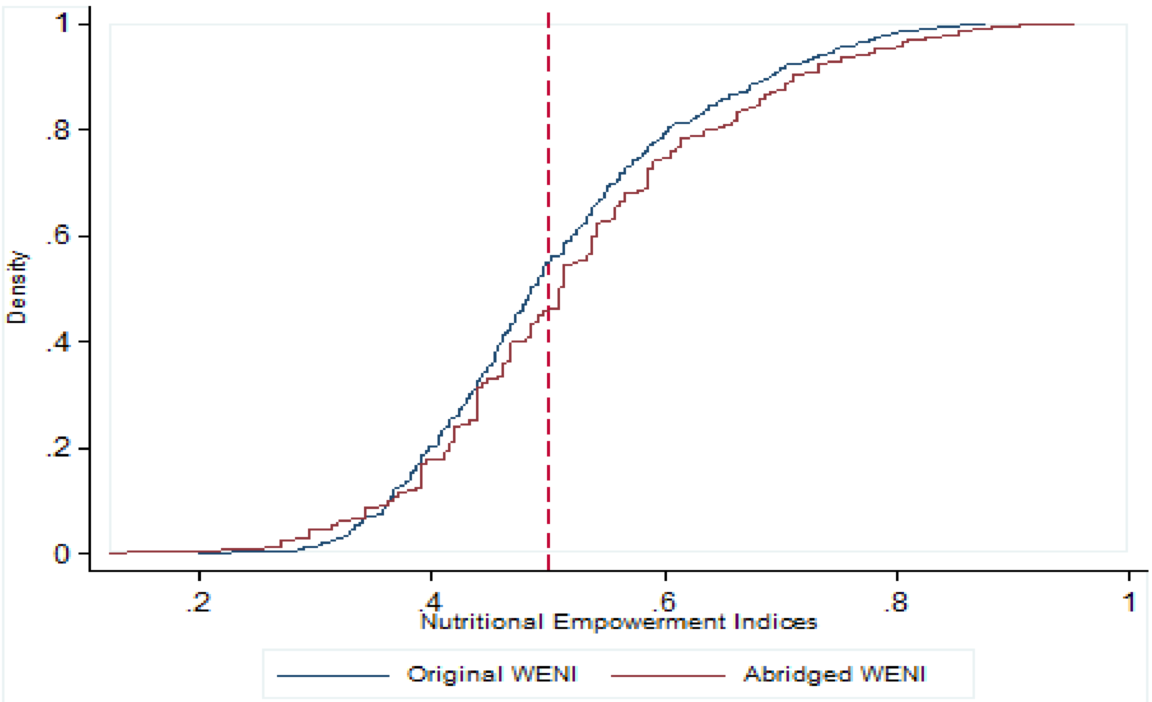
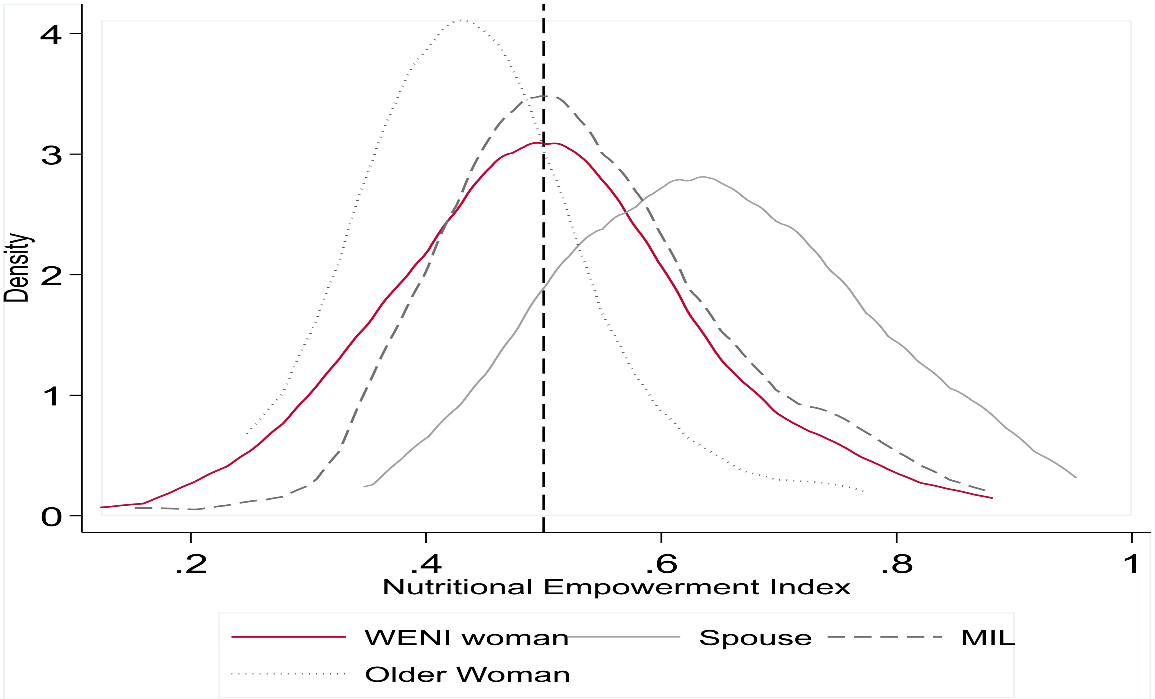


Table 6: Proportion of people empowered across Domain-Dimension in Maharashtra.

	WENI	Spouse	MIL	Older Women	Abridged WENI Total	Original WENI Total
Nutritional Empowerment	51.2	86.1	60.2	22.3	54.1	45.1
Food Knowledge	22.5	44.6	17.5	6.8	22.7	22.7
Food Resources	91.9	96	94.2	99	94.6	64.9
Food Agency	90	100	100	82.5	92.4	97.2
Health Knowledge	41.6	71.3	42.7	19.4	43.2	43.2
Health Resources	90.9	93.1	97.1	95.1	93.4	69.8
Health Agency	49.3	91.1	62.1	56.3	61.4	61.4
Institutions	50.2	99	56.3	24.3	55.8	87.8

Figure 2: Distribution of Abridged WENI by WENI Women, Spouses, Mothers -in-law (MIL) and Older Woman in Maharashtra



Note: The Kolomogorov-Smirnov statistic for equality of distribution of WENI woman vis-à-vis the male spouse is 0.4809 (significant at 1% level), mothers-in-law (MIL) is 0.1464 (significant at 10% level) and vis-à-vis older women is 0.3122 (significant at 1% level)

Table 7 gives us the indicator-wise details for each group for the 20 indicators used to construct the abridged WENI. We find that while about 71% of the WENI woman can visit the bank/post-office or family unaccompanied, 99% of the spouses can do so. We also find while all older woman practising some form of veiling self, 71.3% of the WENI woman perform veiling. In terms of access to assets, we find that while only 0.5% of the WENI women have land in their name, around 16% of their spouses have the same (Table 7 and Figure 3).This however is not surprising as young married women in rural India have little property in their name.

Table 7: Summary Statistics of Abridged WENI Indicators and Outcome Variables

Indicators and Outcome Variables	WENI	Spouse	MIL	Older Women	Total
Indicators					
Knowledge on calcium	0.16	0.36	0.15	0.06	0.18
Knowledge on iodine	0.14	0.30	0.06	0.04	0.14
Eat last only rarely/sometimes/never(=1)	0.92	0.95	0.92	0.99	0.94
Own land in your name	0.00	0.16	0.15	0.11	0.08
Has some say in agricultral activities(=1)	0.81	0.86	0.85	0.87	0.84
Earnings from asset owned by respondent has not been used without consent	0.96	0.96	0.99	0.89	0.95
Has cash as independent source of money & some control over how to spend it(=1)	0.74	0.74	0.66	0.39	0.65
Knowledge on anemia	0.66	0.86	0.86	0.57	0.72
Knowledge on ORS	0.31	0.51	0.30	0.06	0.30
Knowledge on malaria	0.38	0.71	0.41	0.17	0.41
Get assistance in household chores when ill(=1)	0.81	0.93	0.97	0.93	0.89
Household has access to improved water & improved toilet & ventilation(=1)	0.53	0.47	0.47	0.25	0.45
Can go alone to health centre for treatment if need be(=1)	0.65	0.98	0.71	0.60	0.72
No expectation to take permission from family before visiting health centre(=1)	0.50	0.91	0.69	0.81	0.68
Can make decision on own health(=1)	0.08	0.06	0.04	0.07	0.07
Member of any group with own accord(=1)	0.19	0.13	0.43	0.03	0.19
Never practice Ghonghat/Burkha/Pallu/Purdah(=1)	0.29	1	0.19	0	0.35
Individual gets information on government schemes from mobile phone(=1)	0.45	0.48	0.19	0.26	0.37
Experiences no physical abuse or if does, has support within family (=1)	0.98	0.99	0.97	0.99	0.98
Has mobility to go to bank or post office or family alone(=1)	0.71	0.99	0.95	0.90	0.85
Outcome Variables					
BMI	19.30	20.50	19.25	18.82	19.42
Proportion of people with normal BMI	0.57	0.75	0.60	0.51	0.60
Proportion of people fulfilling Minimum Dietary Diversity (MDD)	0.49	0.45	0.48	0.63	0.51

Across DDs we find that on an average, Food-Knowledge has the lowest empowerment rates and Food-resources have the highest rates of empowerment (Table 6). Only 22.7% are empowered in the Food-Knowledge DD (only 6.8% of the older women are empowered in this dimension) whereas 94.6% is empowered in the Food-Resource DD (99% of the older women are empowered in this DD).[26]

[26] In the Food-Knowledge DD we enquire about the knowledge of Calcium and Iodine. In Food-Resources DD, we ask in the respondent has access to land and whether eats last. The former captures whether the individual has access to any assets and the latter captures eating norms.

Figure 3: A Scorecard for the Abridged WENI by Domain-Dimensions in Maharashtra

The validation exercises suggest that the A-WENI is a good catch-all measure for tracking nutritional empowerment. However, for a richer understanding of the impediments to empowerment, especially in terms of specific DDs, the original WENI may be useful. We note here that the DD-specific empowerment status differs significantly between A-WENI and the original WENI in some cases (Institutions, Heath and Food Resources, in particular) (Table 6). for this reason, we recommend the 33-indicator WENI for a fuller account of the various DDs and for specific insights into the DDs.

4.3 A-WENI's rankings and predictive capacity

The abridged WENI (A-WENI) will be considered a good measure of nutritional empowerment if it ranks individuals in the same order as WENI. We thus conduct a rank order correlation test using the Maharashtra Survey data between A-WENI and the 33-indicator WENI. We find that the rank order is preserved and Kendall's tau-b score is 0.77, indicating high positive correlation. We also conduct a ROC analysis and find that 88.74% of the classifications into empowered or disempowered by the abridged WENI are correct (Appendix Table A1).

As with the earlier exercise for data from other states, for the new Maharashtra sample too we compute t-tests and proportion tests of nutritional outcomes based on empowerment status. We find that in both cases we reject the null hypothesis of equality between the two groups (Table 8). We then ascertain if, as with the original WENI, the A-WENI remains a good predictor of nutritional status. We use multiple approaches to test if nutritional empowerment status is associated with better nutritional outcomes.

We estimate models using both BMI and Minimum Dietary Diversity(MDD) as outcome variables and the A-WENI as the independent variable. We estimate a least squares regression and a probit regression models (equation 2,3 and 4) for the Maharashtra data only, with robust standard errors, that correct for heteroscedasticity. As with the earlier BMI regressions, we exclude the sample of overweight and obese individuals. The results hold for both the truncated and full samples. [27]

Table 8: Tests on nutritional outcomes (BMI) based on empowerment status in Maharashtra

Test of Equality of Means	Disempowered		Empowered		P-value
	Mean	Obs	Mean	Obs	
WENI Woman	19.40	102	20.55	107	0.01
Spouse	20.27	14	21.89	86	0.12
MIL	20.17	41	21.90	62	0.06
Older Women	19.73	80	21.91	23	0.03
All	19.70	237	21.38	278	0.00
Test of Equality of Proportions					
WENI Woman	0.53	102	0.67	107	0.03
Spouse	0.71	14	0.80	86	0.45
MIL	0.61	41	0.74	62	0.15
Older Women	0.52	80	0.83	23	0.00
All	0.55	237	0.74	278	0.00

[27]For MDD, we estimate the model on the full sample, including overweight/obese in the sample.

Table 9: Relationship between nutritional empowerment and nutritional status in Maharashtra

	BMI	Normal BMI (=1)	log BMI	Minimum Dietary Diversity (=1)
Abridged WENI Nutritionally empowered (=1)	0.993***	0.386***	0.051***	0.280**
	(-3.73)	(-2.94)	(-3.71)	(-2.3)
Spouse (=1)	0.715**	0.249	0.037**	-0.157
	(-1.97)	(-1.3)	(-2.01)	(-0.93)
MIL(=1)	-0.887	-0.54	-0.045	0.204
	(-1.17)	(-1.44)	(-1.13)	(-0.62)
Older Woman(=1)	-1.511	-1.060*	-0.077	0.859
	(-1.29)	(-1.78)	(-1.26)	(-1.63)
Age (completed years)	0.028	0.022*	0.001	-0.009
	(-1.18)	(-1.8)	(-1.08)	(-0.81)
Constant	18.129***	-0.544*	2.894***	0.037
	(-29.55)	(-1.75)	(-90.84)	(-0.13)
R-squared	0.073	-	0.074	-
Adjusted R-squared	0.063	-	0.063	-
Chi-sq	-	22.481	-	14.844
N	441	441	441	516

Note:* significant at 10%, ** significant at 5%, *** significant at 1%. We control for different demographic groups by using dummies

We find that the A-WENI is good predictor of nutritional outcomes. Higher values of A-WENI are associated with the statistically significantly higher values of nutritional outcomes (Table 9).[28]

These validation exercises offer confidence that relative to the 33-indicator WENI, the A-WENI does not produce statistically significantly different results either in terms of ranking individuals according to nutritional empowerment or in terms of predicting nutritional status.

5 Robustness Checks

As evident from the discussion so far, we find persuasive evidence that the A-WENI based on LASSO methodology performs well relative to the original WENI in many ways. The A-WENI is able to classify individuals into empowered or disempowered categories with acceptable levels of error and is a good predictor of nutritional outcomes, like the original WENI.

We perform a number of additional checks to boost the credibility of the A-WENI as a stand-alone measure of nutritional empowerment. First, we generate a 20-indicator list using a different supervised machine learning algorithm known as Random Forest(The list of indicators generated using random forest algorithm is presented in (Appendix Table 2). Even though it fulfills the criterion of representing at least 50% of the "themes "in each DD, it is a poor predictor of minimum

[28] One concern we have is that the BMI distribution is right-skewed. We therefore transform the data and use logarithmic BMI values. We find that our results are robust to the transformation. We also implement group-wise regressions to analyze the predictive power of A-WENI for different sub-groups are available with the authors.Apart, from the robustness measures mentioned here, we also test if A-WENI, is good measure of empowerment for women in the fertility age group. This is important as decisions relating to fertility is a critical aspect of nutritional empowerment (Narayanan et al., 2019). We find that A-WENI is a good predictor of the nutritional empowerment status created including the fertility module and can therefore be used a measure of empowerment even for women belonging to the fertility age group (Appendix Table 4).

dietary diversity. We also used the elastic net method, fine tuning it by adjusting the penalty values to identify 20 indicators. The 20 indicators identified are the same as those we identify in A-WENI.

Thus far, our focus was on whether A-WENI and WENI both identify the same individuals as being nutritionally empowered. We now test the ability of A-WENI to replicate the nutritional empowerment score (WENI) rather than just the 0-1 nutritional empowerment status based on the WENI. This is to ensure that the A-WENI replicates the underlying variable on which the threshold is imposed, given the prevailing argument against imposing thresholds on empowerment scores (Richardson, 2018). For this, we treat the Y variable in equation 1 as nutritional empowerment scores rather than empowerment status. We find that the list of indicators generated using nutritional empowerment scores is very similar to that generated by nutritional empowerment status (Appendix Table 2).[29]

Finally, we test the sensitivity of the A-WENI to thresholds (Appendix Table 3).We find that the results are robust in most scenarios.We increase the threshold by 0.05 at a time and find that statistical significance does not change significantly up to the threshold of 0.65, after which, the statistical significance is lost to as very few individuals are classified as empowered at such high cut-offs.This pattern is consistent with (Narayanan et al., 2019) using the 33 indicator WENI.

Finally, we test whether we can identify an even smaller set of indicators, for example, just 10, to test the scope for further abbreviation (Appendix Table 2). We find that both the in-sample and out of sample RMSE is considerably higher than 20-indicator A-WENI. A-WENI therefore appears to provide the best balance between the lowest RMSE and best prediction amongst all the cases.

These procedures collectively suggest that the A-WENI can offer a credible and practical alternative to the WENI and can be collected with little additional cost and effort.

Notwithstanding its obvious appeal, there are some caveats to its application. While the A-WENI offers a quick snapshot that facilitates comparisons of nutritional empowerment across communities or socio-demographic groups, it has limited use for conducting nutritional empowerment diagnostics to identify key barriers that individuals face in achieving nutritional status in a meaningful way. This is because despite its predictive power and its ability to reproduce the ordering of individuals based on the nutritional empowerment scores, there is nevertheless a loss of detail, relative to the WENI, that might lead one to overlook some key actionable barriers to nutritional empowerment. Abridging the WENI seems to trade away some of the 33-indicator-WENI's ability to offer a granular perspective of the obstacles and barriers challenging women's nutritional empowerment. To that extent, the A-WENI needs to be used cautiously.

[29] The new indicator list differs from the original by only two indicators. The indicators FAagrisay and Inoviolenceorsupport is replaced by FAmajminsayent and Iparticipatedany in the list of indicators based on nutritional empowerment scores. Although the factors are different, the new indicators belong to the same category of themes within their respective DDs as the indicators they replace.

6 Conclusion

We motivated this paper by emphasizing the growing need to track and measure empowerment and the imperative of having appropriate measures of empowerment. A key challenge is that given its complex nature, there appears to be a trade-off between measuring empowerment comprehensively to reflect its many dimensions, which makes it expensive both in terms of time and cost to conduct a full-fledged survey to measure it, and to keep it simple and in the process dilute the rich conceptualizations we have of empowerment. We focus on WENI a recent measure of empowerment in the realm of nutrition in an effort to resolve this dilemma.

Our paper proposed to use machine learning techniques, that are data driven and transparent ways of reducing the number of indicators in an empowerment index in ways that reproduce, more or less, the outcomes of the parent index. Whereas this approach has been used in poverty measurement, this is a first application to the class of empowerment indices. We demonstrate its use to reduce and develop an abbreviated version of a recently created indicator of women's empowerment in the realm of nutrition (the WENI).

The Abridged WENI or A-WENI consists of 20 indicators as opposed to the original 33. This reduction in number of indicators to a 20 indicator A-WENI, is our preferred recommendation as it covers at least 50% of the 'themes' originally selected for each DD, thus aligned closely to the spirit of the original WENI. While it is possible to reduce it even further to, say 15, as we explored, this latter leaves some dimensions with only one indicator, while performing more poorly relatively to A-WENI in its prediction of nutritional empowerment status.

There are however some caveats to incorporating the A-WENI in a general purpose survey. First, as emphasized, the WENI has been constructed for rural south Asian contexts and while it can be tested in other regions of the world, more work would be required to adapt the WENI to urban contexts. Second, there is a risk that incorporating the A-WENI in a general purpose survey might dilute the attention enumerators give to sensitive questions on intrahousehold issues and those such as domestic violence, that are critical for A-WENI. For this reason, if the A-WENI indicators are incorporated in a general purpose survey, enumerators need to be trained and advised appropriately since if indicators are missing in the construction of A-WENI, these need to be dropped them from the analysis.

We believe that the 20 indicators in A-WENI represent significant reductions in survey time (ranging from 20 to 30 minutes) while remaining faithful to the conceptual foundations underpinning WENI. The A-WENI's 20 indicators are based on 28 distinct survey questions relative to the 59 required for the 33-indicator-WENI.[30] Furthermore, many of these 20 indicators likely already form part of any general purpose survey, so that the application of A-WENI perhaps only demand the

[30]Based on our field experience, we estimate that this abridged WENI, even with questions that require some conversation around certain issues, should not take more than 30 minutes for a competent and well-trained enumerator. For future research, one can try to document the time taken for the short survey (even for each question) more systematically in future applications of the A-WENI.

inclusion of just a few new questions to such surveys. Thus reducing the number of indicators from 33 to 20 reduces the survey time considerably and can easily be incorporated in to a general purpose survey. Going from 20 to 15 is unlikely to hold significant additional gains in terms of time or resources.

The A-WENI can serve to fill a crucial gap, especially in many developing country contexts. Very often household surveys do not capture information at a gender dis-aggregated level and doing so can often be infeasible. Further, there has been a longstanding problem that most surveys tend to neglect issues relating to nutrition unless they are surveys specifically for health and nutrition, such as the Demographic Health Surveys (DHS). Consequently we know little about women's lives and the barriers they face, especially in terms of their own nutritional well-being. Metrics such as the A-WENI can be leveraged easily to plug this gap - to secure key information that captures women's empowerment in the realm of nutrition. Furthermore, its strong association with nutritional status precludes the need to collect anthropometric data, should there be serious resource or capacity constraints. The promise of A-WENI and the use of machine learning to help us better design measures of empowerment will be evident when these are tested and validated in other contexts outside of South Asia.This paper offers a way forward.

References

Afzal, M., Hersh, J., and Newhouse, D. (2015). Building a better model: Variable selection to predict poverty in Pakistan and Sri Lanka.

Ahrens, A., Hansen, C. B., and Schaffer, M. E. (2020). lassopack: Model selection and prediction with regularized regression in stata. *The Stata Journal*, 20(1):176–235.

Alsop, R. and Heinsohn, N. (2005). *Measuring empowerment in practice: Structuring analysis and framing indicators*. The World Bank.

Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic literature*, 50(4):1051–79.

FAO, F. (2016). Minimum dietary diversity for women: a guide for measurement. *Rome: FAO*, 82.

Glennerster, R., Walsh, C., and Diaz-Martin, L. (2018). A practical guide to measuring women's and girls' empowerment in impact evaluations. *Gender Sector, Abdul Latif Jameel Poverty Action Lab*.

Hannan, A., Heckert, J., James-Hawkins, L., and Yount, K. M. (2020). Cognitive interviewing to improve women's empowerment questions in surveys: Application to the health and nutrition and intrahousehold relationships modules for the project-level women's empowerment in agriculture index. *Maternal & child nutrition*, 16(1):e12871.

Heckert, J. and Fabic, M. S. (2013). Improving data concerning women's empowerment in sub-saharan africa. *Studies in Family Planning*, 44(3):319–344.

Ibrahim, S. and Alkire, S. (2007). Agency and empowerment: A proposal for internationally comparable indicators. *Oxford Development Studies*, 35(4):379–403.

Kabeer, N. (1999). Resources, agency, achievements: Reflections on the measurement of women's empowerment. *Development and Change*, 30(3):435–464.

Kadiyala, S., Harris, J., Headey, D., Yosef, S., and Gillespie, S. (2014). Agriculture and nutrition in india: mapping evidence to pathways. *Annals of the New York Academy of Sciences*, 1331(1):43–56.

Kshirsagar, V., Wieczorek, J., Ramanathan, S., and Wells, R. (2017). Household poverty classification in data-scarce environments: a machine learning approach. *arXiv preprint arXiv:1711.06813*.

Laszlo, S. and Grantham, K. (2017). Measurement of women's economic empowerment in grow projects: inventory and user guide.

Lentz, E., Michelson, H., Baylis, K., and Zhou, Y. (2019). A data-driven approach improves food insecurity crisis prediction. *World Development*, 122:399–409.

Malapit, H., Quisumbing, A., Meinzen-Dick, R., Seymour, G., Martinez, E. M., Heckert, J., Rubin, D., Vaz, A., Yount, K. M., Phase, G. A. A. P., et al. (2019). Development of the project-level women's empowerment in agriculture index (pro-weai). *World Development*, 122:675–692.

Malhotra, A., Schuler, S. R., et al. (2005). Women's empowerment as a variable in international development. *Measuring empowerment: Cross-disciplinary Perspectives*, 1(1):71–88.

McBride, L. and Nichols, A. (2018). Retooling poverty targeting using out-of-sample validation and machine learning. *The World Bank Economic Review*, 32(3):531–550.

Narayanan, S., Lentz, E., Fontana, M., De, A., and Kulkarni, B. (2019). Developing the women's empowerment in nutrition index in two states of india. *Food Policy*, 89:101780.

Pereznieto, P. and Taylor, G. (2014). A review of approaches and methods to measure economic empowerment of women and girls. *Gender & Development*, 22(2):233–251.

Pratley, P. (2016). Associations between quantitative measures of women's empowerment and access to care and health status for mothers and their children: a systematic review of evidence from the developing world. *Social Science & Medicine*, 169:119–131.

Richardson, R. A. (2018). Measuring women's empowerment: A critical review of current practices and recommendations for researchers. *Social Indicators Research*, 137(2):539–557.

Schrider, D. R. and Kern, A. D. (2018). Supervised machine learning for population genetics: a new paradigm. *Trends in Genetics*, 34(4):301–312.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288.

Wuest, T., Weimer, D., Irgens, C., and Thoben, K.-D. (2016). Machine learning in manufacturing: advantages, challenges, and applications. *Production & Manufacturing Research*, 4(1):23–45.

Yount, K. M., Cheong, Y. F., Maxwell, L., Heckert, J., Martinez, E. M., and Seymour, G. (2019). Measurement properties of the project-level women's empowerment in agriculture index. *World Development*, 124:104639.

Appendix

Appendix Tables

Table A1: Area under the ROC curve for Maharashtra

Indicator List	Observations	Area under the ROC curve	Standard error	Lower bound	Upper bound
RLasso-15	516	0.849922	0.015825	0.818905	0.880938
CVLasso-15	516	0.820258	0.016837	0.787258	0.853259
Lasso With IC-15	516	0.820258	0.016837	0.787258	0.853259
20- Indicator List	516	0.887426	0.013138	0.861676	0.913177

Table A2: List of Indicators based on different approaches

Sl. No.	20 Indicator List based on Nutritional Empowerment scores	15 Indicator List based on Nutritional Empowerment scores	10 Indicator list based on Nutritional Empowerment Status	20 Indicator list Based on Random Forest Algorithm
1	FRland	FKcalcium	FKcalcium	FKiodine
2	FReatlast	FKiodine	FKiodine	FRjobandpds
3	FKcalcium	FAcashcontrol	HRdrinktoivent	FRselfemployment
4	FKiodine	HRhassistwhensick	HKanemia	FReatlast
5	FAassearnconsentown	HRdrinktoivent	HKors	FRpaidwork
6	FAmajminsayent	HKanemia	HKmalaria	FRfinsupportagrihhenter
7	FAcashcontrol	HKors	HAalonefortreatment	FAdecisionpaidwrkbin
8	HRhassistwhensick	HKmalaria	HAhealthvisitpermission	FAassearnconsentown
9	HRdrinktoivent	HAalonefortreatment	HAhdecideownhealth	FAcashcontrol
10	HKanemia	HAhealthvisitpermission	Idoveil	HKors
11	HKors	HAhdecideownhealth		HKmalaria
12	HKmalaria	Ianymemberownaccord		HKanemia
13	HAalonefortreatment	Idoveil		HRintensityany
14	HAhealthvisitpermission	Ifreedommove		HRdrinktoivent
15	HAhdecideownhealth	Iparticipatedany		HRhassistwhensick
16	Ianymemberownaccord			HAalonefortreatment
17	Idoveil			HAhealthvisitpermission
18	Imobileinformationgovt			HAhdecideownhealth
19	Ifreedommove			Ifreedommove
20	Iparticipatedany			Idoveil

Table A3: Relationship between nutritional outcomes and nutritional empowerment for Maharashtra (across different thresholds)

	0.50			0.55			0.60			0.65		
	BMI	Normal BMI (=1)	MDD (=1)	BMI	Normal BMI (=1)	MDD (=1)	BMI	Normal BMI (=1)	MDD (=1)	BMI	Normal BMI (=1)	MDD (=1)
Abridged WENI Nutritionally empowered (=1)	0.993***	0.386***	0.280**	0.880***	0.427***	0.267**	0.866***	0.371**	0.340**	0.788**	0.367**	0.446***
	(-3.73)	(-2.94)	(-2.3)	(-3.11)	(-3.02)	(-2.13)	(-2.79)	(-2.3)	(-2.41)	(-2.25)	(-2.03)	(-2.91)
Age (completed years)	0.028	0.022*	-0.009	0.023	0.019	-0.009	0.02	0.018	-0.011	0.023	0.019	-0.011
	(-1.18)	(-1.8)	(-0.81)	(-0.99)	(-1.62)	(-0.88)	(-0.85)	(-1.52)	(-1.01)	(-0.96)	(-1.6)	(-1.02)
Spouse(=1)	0.715**	0.249	-0.157	0.713*	0.22	-0.171	0.782**	0.269	-0.185	0.826**	0.279	-0.208
	(-1.97)	(-1.3)	(-0.93)	-1.93	-1.14	(-0.99)	-2.17	-1.41	(-1.07)	-2.25	-1.46	(-1.20)
MIL(=1)	-0.887	-0.54	0.204	-0.735	-0.463	0.233	-0.623	-0.425	0.273	-0.68	-0.446	0.28
	(-1.17)	(-1.44)	(-0.62)	(-0.97)	(-1.24)	-0.7	(-0.82)	(-1.13)	-0.82	(-0.89)	(-1.18)	-0.84
Older Woman (=1)	-1.511	-1.060*	0.859	-1.423	-0.98	0.861	-1.297	-0.953	0.930*	-1.484	-1.021*	0.930*
	(-1.29)	(-1.78)	(-1.63)	(-1.19)	(-1.64)	-1.62	(-1.07)	(-1.58)	-1.74	(-1.23)	(-1.69)	-1.73
Constant	18.129***	-0.544*	0.037	18.457***	-0.427	0.116	18.632***	-0.345	0.171	18.623***	-0.35	0.179
	(-29.55)	(-1.75)	(-0.13)	(-31.27)	(-1.41)	(-0.42)	(-31.52)	(-1.14)	(-0.62)	(-31.19)	(-1.15)	(-0.65)
R-squared	0.073	-	-	0.065	-	-	0.06	-	-	0.055	-	-
Adjusted R-squared	0.063	-	-	0.054	-	-	0.05	-	-	0.045	-	-
Chi-sq	-	22.481	14.844	-	22.708	13.914	-	18.838	14.746	-	17.629	17.119
N	441	441	516	441	441	516	441	441	516	441	441	516

Note:* significant at 10%, ** significant at 5%, *** significant at 1%.

Table A3 (contd.): Relationship between nutritional outcomes and nutritional empowerment for Maharashtra (across different thresholds)

	0.70			0.75			0.80			0.85		
	BMI	Normal BMI (=1)	MDD (=1)	BMI	Normal BMI (=1)	MDD (=1)	BMI	Normal BMI (=1)	MDD (=1)	BMI	Normal BMI (=1)	MDD (=1)
Abridged WENI Nutritionally empowered (=1)	0.785*	0.288	0.339*	0.961*	0.44	0.23	1.419***	1.177**	0.234	1.406*	0	0.026
	(-1.87)	(-1.33)	(-1.93)	(-1.88)	(-1.46)	(-1.04)	(-2.93)	(-2.38)	(-0.87)	(-1.87)	(.)	(-0.07)
Age (completed years)	0.029	0.022*	-0.008	0.027	0.021*	-0.009	0.028	0.021*	-0.008	0.028	0.020*	-0.008
	(-1.2)	(-1.82)	(-0.74)	(-1.1)	(-1.75)	(-0.80)	(-1.16)	(-1.76)	(-0.76)	(-1.16)	(-1.71)	(-0.73)
Spouse(=1)	0.892**	0.321*	-0.149	0.955***	0.335*	-0.096	0.964***	0.347*	-0.09	0.986***	0.351*	-0.065
	(-2.45)	(-1.71)	(-0.87)	(-2.67)	(-1.8)	(-0.58)	(-2.7)	(-1.85)	(-0.54)	(-2.74)	(-1.86)	(-0.39)
MIL(=1)	-0.86	-0.524	0.203	-0.803	-0.506	0.226	-0.816	-0.499	0.216	-0.827	-0.484	0.208
	(-1.12)	(-1.40)	-0.61	(-1.03)	(-1.35)	-0.68	(-1.06)	(-1.34)	-0.65	(-1.08)	(-1.30)	-0.63
Older Woman (=1)	-1.791	-1.161*	0.764	-1.712	-1.133*	0.78	-1.756	-1.111*	0.755	-1.784	-1.103*	0.735
	(-1.48)	(-1.95)	-1.43	(-1.40)	(-1.88)	-1.46	(-1.47)	(-1.88)	-1.43	(-1.49)	(-1.86)	-1.39
Constant	18.525***	-0.384	0.139	18.603***	-0.361	0.17	18.582***	-0.362	0.161	18.612***	-0.335	0.16
	(-30.62)	(-1.26)	(-0.5)	(-30.51)	(-1.18)	(-0.61)	(-31.15)	(-1.20)	(-0.59)	(-31.09)	(-1.11)	(-0.58)
R-squared	0.052	-	-	0.052	-	-	0.054	-	-	0.05	-	-
Adjusted R-squared	0.042	-	-	0.042	-	-	0.044	-	-	0.039	-	-
Chi-sq	-	15.279	12.354	-	15.381	9.921	-	20.795	9.62	-	12.184	8.9
N	441	441	516	441	441	516	441	441	516	441	433	516

Note:* significant at 10%, ** significant at 5%, *** significant at 1%.

Table A4: Relationship between the original empowerment status of WENI Woman (including the fertility module) with the abridged WENI

	Empowerment status for WENI in 5 states	Empowerment status for WENI in Maharashtra
Abridged WENI Nutritionally empowered (=1)	2.039***	2.423***
	(-9.09)	(-24.4)
Age (completed years)	-0.038	-0.009
	(-1.47)	(-0.97)
Constant	0.219	-0.754***
	(0.34)	(-2.89)
Chi-sq	83.262	595.849
N	209	1145

Note:* significant at 10%, ** significant at 5%, *** significant at 1%.