**Institute for International Economic Policy Working Paper Series**
**Elliott School of International Affairs**
**The George Washington University**

# Measurement of Human Recognition: A Methodology with Empirical Applications in India and Kenya

**IIEP-WP-2011-10**

**Tony Castleman**
**Institute for International Economic Policy**
**Elliott School of International Affairs**
**George Washington University**
**Washington DC 20052**

**September 2011**

Institute for International Economic Policy
1957 E St. NW, Suite 502
Voice: (202) 994-5320
Fax: (202) 994-5477
Email: iiep@gwu.edu
Web: www.gwu.edu/~iiep

Measurement of Human Recognition: A Methodology with Empirical Applications in India and Kenya

Tony Castleman*

September 2011

**Abstract:** This paper develops and applies a methodology for measuring human recognition, which is defined as the acknowledgement provided to an individual by other individuals, groups, or organizations that he is of inherent value with intrinsic qualities in common with the recognizer. A framework is developed that organizes the sources of human recognition into various domains of an individual's life. The framework is used to develop an index of indicators that measures human recognition received in each of the domains and combines these domain-specific measures into a single overall measure of human recognition received. Two empirical applications of the index are presented with cross-sectional survey data from India and Kenya. Exploratory factor analysis is used to generate measures of human recognition with the index, and the resulting measures are used in multivariate regression models of nutritional status. Results from both datasets provide evidence that human recognition is a significant, independent, positive determinant of nutritional status, controlling for socio-economic characteristics. The method and applications demonstrate how latent, intangible aspects of development such as human recognition can be measured and indicate that further empirical work on the determinants and effects of human recognition is both feasible and needed.

JEL Codes: I12, I31, O15, I14,

Keywords: human recognition, nutrition, health, dehumanization, dignity, respect, domestic violence, measurement, India, Kenya, economic development, poverty

---

* Tony Castleman is Associate Research Professor of International Affairs and Associate Director of the Institute for International Economic Policy, Elliott School of International Affairs, The George Washington University. Email: tonyc@gwu.edu. Tel: 202-994-7722

The author thanks Stephen Smith, Shahe Emran, Sumit Joshi, and David Ribar for helpful comments.

*Measurement is the process by which a concept is linked to one or more latent variables and these are linked to observed variables*.

- Kenneth Bollen (1989)

## I. Introduction

As the objectives of economic development have expanded to encompass less tangible dimensions of people's lives, practitioners and researchers face the challenge of developing valid and reliable measures of these dimensions. Unlike tangible outcomes such as income, educational attainment, and infant mortality, straightforward measures do not exist for empowerment or social capital, two components of development that have received increased emphasis in recent years. Human recognition, the subject of this paper, is also a challenging concept to measure. Human recognition is defined as the acknowledgement provided to an individual by other individuals, groups, or organizations that he is of inherent value with intrinsic qualities in common with the recognizer, i.e. recognition as a fellow human being. Human recognition may be positive or negative. Positive recognition refers to viewing an individual as of value by virtue of being a human being, and negative human recognition refers to viewing an individual as lacking inherent value as a human being or not acknowledging this value. In many contexts the level of human recognition an individual receives is a function of the degree to which his basic needs and rights are acknowledged to exist and to be of consequence.

Human recognition can occur in many domains of an individual's life, and no observable composite indicator exists that captures the total level of recognition that an individual receives or possesses. Yet having effective measures of human recognition is a critical step to better understanding the role it plays in well-being, poverty, and development, and to enabling development programs and policies to address human recognition issues in the design and implementation of interventions. As Robert Lane wrote, “A discipline that does not have independent measures of its dependent variable…risks its standing as a scientific discipline” (Lane 1991, cited in Narayan 2005). Valid, reliable, and sensitive measures of human recognition are needed for the following applications:

1) To assess a population’s human recognition status either for research purposes or to inform the design of a program targeting this population, e.g., the level of human recognition received by members of a particular ethnic minority in an urban community.
2) To identify the determinants of human recognition for a population, e.g., the determinants of human recognition among married women living in a set of rural villages.
3) To understand the impacts that human recognition has on tangible outcomes, e.g., the effect changes in human recognition have on health outcomes.
4) To evaluate the effects that policy or program interventions have on human recognition levels, e.g., changes in human recognition among members of a microfinance group since the group was established.
5) To compare levels of human recognition across groups and contexts in order to increase understanding of different populations’ relative status and inform resource allocation decisions, e.g., comparing levels of human recognition among a group of urban slum residents to a group of rural village residents of similar economic status in the same country.

In response to these needs, this paper develops a framework that categorizes sources of human recognition into a set of domains and develops a methodology for measuring human recognition in each of the domains. Using data from the India National Family and Health Survey 2005-6 (NFHS-3) and the Kenya Demographic and Health Survey (DHS) 2003, the paper applies this methodology to measure human recognition received by approximately 2,500 women in Kenya and 26,000 women in India. The method is also applied in another paper using data from a randomized controlled trial in Kenya (Castleman 2011c).

The measures obtained are then used to empirically test hypotheses that receipt of human recognition is a significant determinant of nutritional status. Using the notation from the theoretical model of human recognition (Castleman 2011a), the following hypothesis is tested: $\frac{\partial h}{\partial R} \geq 0$. This expression is the partial derivative of health with respect to human recognition, and the hypothesis is that an individual's recognition level is a determinant of health status and that the relationship between recognition and health is positive. The hypothesis is tested with the India and Kenya datasets using regression models in which human recognition is a determinant of nutritional status, a key component of health. More extensive empirical work is carried out in Castleman 2011c; the primary objective of this paper is to develop and demonstrate the measurement methodology.

The next section describes the sources of human recognition and the challenges entailed in measuring recognition. Section III reviews related literature on measuring empowerment and social capital. Section IV presents a framework and index for measuring human recognition and describes a method for empirically applying the index.

Sections V and VI apply the index using the two survey datasets and empirically test the relationship between human recognition and nutrition outcomes. The final section discusses implications of the results and areas for further study.

## II. Human Recognition Sources and Measurement Challenges

<u>Sources of Human Recognition</u>

Many sources of human recognition exist because most institutions and activities that people are engaged in are potential arenas for human recognition transactions, especially activities involving substantial interpersonal interactions. Primary sources of human recognition can be grouped into three categories:

1. ***household*** and family relationships, roles, interactions, and behavior;
2. ***community*** norms and interactions among community members, including neighbors, community leaders, and friends;
3. ***organization and institution*** norms, interactions, and systems, such as in schools, places of employment, religious organizations and places of worship, health care facilities, and other service delivery points.

Culture and religion could be considered a fourth domain, and many human recognition transactions are rooted in cultural and religious factors and conditions. But cultural and religious traditions and norms generally operate through one of the three domains described above. For example, cultural traditions such as female genital cutting that involve provision of negative recognition occur in the household and community domains. Positive human recognition that otherwise marginalized individuals receive from their religious institutions and places of worship occurs in the domain of organizations and institutions. Therefore, in the framework of domains where human recognition transactions occur, culture and religion are structured as underlying the three primary domains listed above. The categorization of domains forms the basis of the framework for measuring human recognition that is presented in Section IV.

Development programs and policies can be significant sources of human recognition directly through organizations and institutions involved in implementation (domain 3) and indirectly through effects on family and community behavior (domains 1 and 2). The impacts that development interventions have on human recognition can be positive or negative and can be deliberate or inadvertent. As detailed elsewhere (Castleman 2011), program activities can increase or reduce the human recognition levels of targeted populations through the interpersonal interactions of implementers, organizational norms of implementing agencies, systems and processes applied in activities, interventions that directly influence human recognition transactions, and interventions that have indirect effects on individual, household, or community interactions. The hypothetical example in Castleman 2011 illustrated that if medical staff in a health clinic treat clients disrespectfully through objectifying or dehumanizing verbal communication or through lack of respect for privacy, this reduces clients' human recognition levels.

As modeled elsewhere (Castleman 2011a), receipt of human recognition has both psychic and material effects on the receiving individual's well-being. The psychic effects are direct or constitutive effects whereby the recognition itself affects the well-being of the individual receiving it. The material effects are indirect or instrumental effects whereby recognition influences the receiver's and/or the provider's behavior in ways that affect the receiver's well-being, such as health care seeking behaviors, violence, or other forms of abuse. The terminology of constitutive and instrumental effects is drawn from Sen's discussion of freedom (Sen 1999).

The populations targeted by development interventions are often those most vulnerable to low levels of human recognition, such as socially marginalized or economically impoverished groups. These vulnerable groups may also have significant scope to benefit from receipt of positive recognition, both through direct psychic effects and through recognition's effect on opportunities and capacities to improve material well-being through economic activities, fertility decisions, and protection from violence.

Challenges to Measuring Receipt of Human Recognition

Efforts to measure human recognition received by individuals face a number of challenges, many of which also apply to measurement of other intangible dimensions of development such as empowerment and social capital. Five key challenges are identified here.

1) First and foremost, *no direct measure of human recognition exists*; unlike income, fertility, education, or morbidity, people do not have observable, countable human recognition levels. It is a latent, unobserved variable. The challenge this poses to measurement is apparent, as it is not possible to use a single, directly observed indicator to measure human recognition.

2) While a number of observable indicators do exist that reflect human recognition transactions, these *indicators also reflect a number of factors other than recognition*. For example, suppose data on female genital cutting were used as a measure of human recognition in an empirical study of human recognition's effects on women's

employment, with a binary employed-not employed variable used as the dependent variable. A researcher could interpret the coefficient on the genital cutting variable as capturing the relationship between human recognition and employment among the population of women studied. But another researcher could interpret the same coefficient as capturing the relationship between ethnicity and employment because ethnicity is a strong determinant of the practice of female genital cutting. Even if ethnicity were controlled for in the regression, a third researcher could interpret the coefficient as capturing the relationship between the community's or household's degree of modernity and employment because within a given ethnic group, those adopting more modern lifestyles may be less likely to practice female genital cutting.

The problem here is that most existing, observable indicators of human recognition do not reflect only recognition; on the contrary, most existing indicators that strongly reflect human recognition also reflect other factors. Therefore, interpreting empirical results about such a variable in terms of human recognition can be problematic and difficult to defend. One way to address this challenge is to design indicators specifically aimed at capturing human recognition, such as self-reported recognition levels and incidence of specific interactions such as humiliation. (Castleman 2011c uses this approach.) The empirical application in this paper uses existing, observable variables from survey datasets to measure human recognition, which requires a method to deal with this challenge.

3) A third challenge to measuring human recognition is that *human recognition transactions occur in many different aspects of an individual's life*. Kishor's observation that "there are as many domains in which women can display evidence of empowerment

as there are domains in women's lives" (Kishor 2000) applies equally well to human recognition. Examining only one indicator – or even one domain – to measure human recognition may significantly mismeasure the actual level of recognition an individual receives because other human recognition transactions are not captured in the measure. One can be treated like a queen at work and beaten up at home.

4) Even after honing in on a particular human recognition transaction, *measurement will vary depending on the point of view from which it is measured*. The level of recognition provided, as self-reported by the *provider*, is likely to differ somewhat from the level of recognition received, as self-reported by the *receiver*. And the levels reported by the provider and the receiver are both likely to differ somewhat from the level reported by an objective third party, such as an observer of the interaction or a survey of the presence or absence of specific types of interactions (e.g. violence). The theoretical model of human recognition (Castleman 2011a) captures this difference between provider and receiver in the $\rho_{hi}$ term of the $r_{i_r} = \frac{1}{\sqrt{n}} \sum_{h=1}^{n} \rho_{hi} r_{hi}$ expression. The $\rho_{hi}$ term captures both differences among providers of recognition in the impact their recognition provision has on a given receiving individual, and differences among receivers of recognition in how they convert a given quantity of provided recognition into received recognition.

While the issue is addressed in the theoretical model, the challenge remains for empirical measures of human recognition, which depend on the point of view of the person reporting or measuring the level of recognition and this person's role in the interaction. The same human recognition transaction will have different valuations depending on who provides the information used to measure it. This challenge involves both a) the difference between subjective and objective measurement, e.g., an indicator of an individual's self-reported level of recognition vs. an indicator of the presence or absence of violence, and b) within subjective measurement, differences among the assessments of different parties, i.e., provider, receiver, and observer.

5) A final challenge relates to the variables used to measure human recognition in different contexts. *A valid and sensitive measure of human recognition in one context may be irrelevant or constant across individuals in another context*. Examples related to human recognition and empowerment include: female genital cutting is entirely absent from some cultures and occurs with high levels of variation in others; Mason points out that whether women can leave home by themselves reflects empowerment in some cultures but not in others (Mason, 2005); some observable indicators, such as the wearing of a veil, are interpreted to be empowering by women in some settings and disempowering by others (Narayan 2005). This suggests the need to be context-specific in selecting variables, and may pose difficulties for making empirical comparisons in human recognition levels across contexts.

## III. Experience Measuring Related Concepts

The concept of human recognition shares some characteristics with empowerment and social capital, but is distinct in a number of respects. (For more details on the commonalities and distinctions between empowerment and human recognition and between social capital and human recognition, see Castleman 2011.) Empowerment is defined as an increase in individuals' capacity to make key choices affecting their lives (Kabeer 2001, cited in Malhotra et al. 2002). Human recognition is an interactive process that underlies many dimensions of empowerment; empowerment occurs *within* individuals, and one process that brings about empowerment is human recognition, which occurs *between* individuals.

Social capital is defined by Fukuyama as "an instantiated set of informal values or norms shared among members of a group that permits them to cooperate with one another" (Fukuyama 1999, cited in Durlauf 2001). Both human recognition and social capital are inherently about interactions among individuals and groups, both affect economic and social development, and both are in turn influenced by development policies and programs. But the two are conceptually distinct. Social capital refers to networks and interactions that enhance trust and cooperation, which is distinct from the interaction of recognizing the inherent value of another individual as a fellow human being. Furthermore, social capital plays a purely instrumental role in development; the function of social capital, true to its name, is to enable production, which it does by supporting cooperation[1]. Human recognition plays both instrumental and constitutive roles, indirectly contributing to economic and social outcomes and directly improving utility and well-being.

While human recognition differs in several respects from empowerment and social capital, recent experience measuring these two concepts can be instructive for efforts to measure human recognition because – like human recognition – they are intangible concepts central to development that affect tangible, material outcomes. A review of recent work on measuring empowerment and social capital, summarized below, offers insights and lessons relevant to human recognition measurement.

[1] Note that there may be direct psychic benefits of social capital as well, but the treatment of social capital in development literature is mainly confined to its instrumental role enabling economic outcomes.

There has been a lag between realization of the important role these intangible concepts play in development and efforts to develop valid, standardized measures. A 2002 review of the women's empowerment literature, commissioned by the World Bank, concluded that while the Bank has identified empowerment as critical to poverty reduction and a key objective of development assistance, "to date neither the World Bank nor any other major development agency has developed a rigorous method for measuring and tracking changes in levels of empowerment" (Malhotra et al. 2002).

This is especially true at the individual and household level. At the national level, the United Nations established the Gender Empowerment Measure in 1995, an index of four national-level indicators of women's political and economic participation (UNDP 2004, Pillarisetti and McGillivray 1998). This measure can be used at the country level to gauge a country's progress over time or for cross-country comparisons, but cannot be applied at program, community, household, or individual levels.

The World Bank has compiled work undertaken by researchers in various disciplines on measuring different aspects of empowerment (World Bank 2005). Pointing out the challenges of measuring something that is essentially a process, Malhotra and Schuler discuss different approaches taken to measuring empowerment, including using proxy indicators such as education or employment, collecting qualitative data over time on the occurrence of specific events and actions, and examining changes in these data over multiple data points (Malhotra and Schuler 2005). Mason categorizes measurement approaches into four groups: measuring factors that lead to empowerment, measuring outcomes of empowerment, observing specific behaviors that indicate empowerment levels, and surveying self-reported levels of empowerment (Mason 2005).

Drawing from this categorization, the methods that apply best to measuring human recognition are measuring specific behaviors that involve human recognition transactions (e.g. violence), and surveying self-reported levels of human recognition.

As with human recognition, the fact that empowerment enters multiple aspects of an individual's life poses challenges for measurement. The World Bank-commissioned review recommends developing a framework of domains where empowerment occurs and identifying indicators of empowerment from each domain (Malhotra et al. 2002). By enabling selection of appropriate, context-specific indicators, this approach supports measurement of women's empowerment in different settings. Given the similarly multi-domain nature of human recognition, this is also a fruitful approach for measuring recognition and is the approach applied in Section IV.

Diener and Biswas-Diener examine the role empowerment plays in subjective well-being and how measurement of subjective well-being can be used to assess empowerment and its effects, including psychological empowerment (Diener and Biswas-Diener 2005). Building on their work and that of other authors (e.g. Kingdon 2006), subjective well-being can be used to assess the role human recognition plays in individual well-being (Castleman 2011c). Measuring subjective well-being is an important means of testing the hypothesis that human recognition affects overall utility, and of testing its impact on psychic utility, independent of its effects on utility through consumption and health outcomes.

Like human recognition, empowerment is both a means to other ends (e.g. income, reproductive health) and an important end unto itself. Khwaja discusses the implications that empowerment's role in a given model has for the measurement approach used, in particular for endogeneity issues if one seeks to demonstrate causality between empowerment and other outcomes (Khwaja 2005). This distinction is relevant to human recognition as well. Tests indicate that the recognition variable is endogenous in many of the empirical models used in this paper, and Khwaja's insights are valuable for understanding and addressing the endogeneity that can result from simultaneity or omitted variables.

Efforts to measure social capital in development settings are also relatively nascent and there has been limited focus on producing consistent, standardized measures. Researchers studying social capital in developing countries have used a range of different measures. Baliamoune-Lutz and Lutz (2004) developed an index of corruption measures as a measure of social capital in a cross-country study of social capital's role in well-being in Africa. Knack and Keefer (1997) have used survey responses about trust of others and civic norms related to honesty and cooperation. Narayan and Pritchett (1999) have used a combination of objective information about membership in groups and self-reported information about trust of strangers and government officials to measure social capital levels. The strategy of combining objective information about experiences and behaviors with self-reported survey responses is applied here to measure human recognition.

In a review of approaches used to measure social capital in community and family settings, Stone (2001) points out the need to differentiate outcomes of social capital, such as reciprocal acts and exchanges, from measures of actual social capital itself, such as the extent of trust itself within the unit under study. This distinction is relevant to measurement of human recognition as it is often easier to measure the observable outcomes of human recognition than it is to measure the recognition itself. The outcomes may be the result of multiple processes, only one of which is human recognition, which reduces the validity of these outcomes as individual measures of human recognition. Using multiple outcome measures in combination with measures of human recognition transactions themselves helps to mitigate this problem.

Grootaert finds that the multiplicity of ways researchers measure social capital stems from the multiplicity of ways that social capital is defined. Just as the World Bank review recommends for empowerment, Grootaert recommends that measurement of social capital be based on a conceptual framework that organizes the different components and roles of social capital (Grootaert 2001). This recommendation is adapted to measurement of human recognition in Section IV.

In separate articles, Durlauf and Moffitt analyze econometric challenges faced in empirical study of social interactions and social capital, primarily related to identification problems due to endogeneity caused by simultaneity or by codetermination of outcome variables and social capital (Durlauf 2001; Moffitt 2001). Moffitt offers some approaches to address these challenges, such as randomization and use of nonlinearities in models. As a social interaction, human recognition shares these same challenges in empirical work[2], and randomization, differencing, and instruments are used to address them in the empirical estimation in this paper and in Castleman 2011c.

As mentioned above, Malhotra et al. recommend the development of a framework of domains to identify appropriate context-specific indicators of women's empowerment

[2] As discussed in Castleman 2011, the simultaneity between individuals that Durlauf identifies as a problem for measuring many social interactions may not apply to human recognition because provision of human recognition tends to occur "down" the power hierarchy, but simultaneity between recognition and other characteristics may apply.

in each domain for a given population or context, and Grootaert suggests a similar approach for measurement of social capital. A common framework helps ensure conceptual cohesiveness among measurement in different contexts. Operationalizing measurement using such a framework requires identification of indicators and a method for combining multiple measures into a composite indicator. Williams begins this process for measurement of empowerment by using confirmatory factor analysis to identify a set of indicators that reflect the gender component of empowerment. The analysis uses data from a sample of women in Bangladesh, and Williams suggests that the model and empirical approach can be applied in other settings to identify and apply context-specific indicators to measure gender aspects of empowerment (Williams 2005).

Kishor applies exploratory factor analysis to combine indicators of women's empowerment in an empirical study of the impact that women's empowerment has on child health outcomes in Egypt. The empowerment indicators are categorized into groups based on different types of empowerment (Kishor 2000). By assessing the extent to which variation in a set of indicators is explained by common factors, factor analysis can combine multiple indicators into a single measure that captures the common factor.

Building on Kishor's and Williams' work, the empirical approach here uses exploratory factor analysis to combine multiple indicators of human recognition. While the approach draws on Kishor's and Williams' approaches to measuring empowerment, it differs in a number of respects:

- The concept being measured is human recognition instead of empowerment.
- The framework developed is specific to human recognition and is directly linked to a measurement index.

- While this paper uses existing indicators that are not specific to human recognition, the approach can also be applied using indicators specifically designed to measure human recognition (as is done in Castleman 2011c).
- The outcome variables examined here are women's own health outcomes, not their children's.
- Endogeneity is tested for and addressed in the empirical specifications.

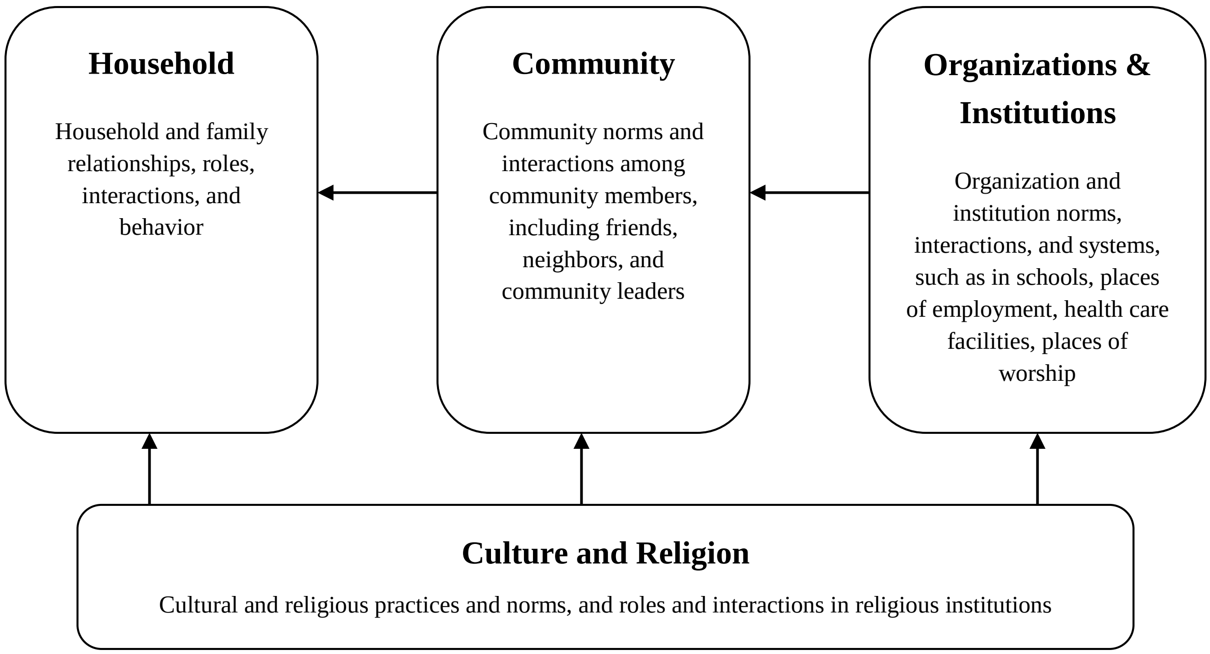
## IV. Measurement Methodology: A Human Recognition Index

Framework

Organizing into a simple framework the different domains in which human recognition transactions occur provides a structure for measurement of human recognition. This approach to measurement creates a structure for addressing the third challenge (multiple domains) and helps address the fifth challenge (variation across contexts). Such a framework of domains also helps concretize and illustrate the various sources of human recognition: household, community, organizations and institutions, and culture and religion. Figure 1 depicts the framework.

As discussed earlier, there are three primary domains in which individuals receive human recognition. Culture and religion are also sources of human recognition and they operate primarily through the three primary domains. As the horizontal arrows in Figure 1 indicate, the "larger" domains can also influence human recognition transactions in the "smaller" domains, such as when community norms about the role of women affect the human recognition women receive in their households.

**Figure 1: Framework of Human Recognition Domains**




Using this framework, specific sources of human recognition and measurable indicators of recognition within each domain can be identified for a given context. Table 1 provides illustrative examples of sources and indicators in each domain. Note that some of the indicators are objective measures of whether particular actions involving human recognition transactions have occurred, other indicators are self-reported levels of recognition *received* by individuals, and others are self-reported levels of recognition *provided* to individuals.

**Table 1: Illustrative Sources and Indicators of Human Recognition**

| Domain | Example Sources of Human Recognition | Example Indicators |
|---|---|---|
| Household | - Behavior during disagreements (e.g. violence, negotiation)<br>- Interactions in public (e.g. humiliation, protection, support)<br>- Decision processes about sexual behavior<br>- Degree and type of participation in household activities, e.g. meals, household work | - Occurrence/frequency of domestic violence<br>- Occurrence/frequency of forced sexual relations<br>- Occurrence of female genital cutting<br>- Self-reported (by individual) level of human recognition received in the household<br>- Self-reported (by others) level of human recognition provided to the individual by others in the household<br>- Self-reported (by others) permissibility of violence toward the individual |
| Community | - Interactions with neighbors<br>- Community rules and norms about participation in decision-making process | - Incidence/frequency of physical or verbal abuse of poorer/less powerful community members by wealthier/more powerful members<br>- Whether prohibited from common community facilities, e.g. collecting water from certain sources<br>- Self-reported (by individual) level of human recognition received in the community<br>- Self-reported (by community leader) view of the individual's rights and role in community activities |
| Organizations and Institutions | - Workplace disciplinary practices<br>- Availability of basic facilities in the workplace<br>- Employer rules/norms for when a worker is sick or hurt<br>- Teacher behavior and disciplinary practices<br>- Doctor/nurse attitudes towards patients | - Whether forced to work while sick or hurt<br>- Other specific workplace policies, e.g. working hours<br>- Self-reported (by employer) view of rights employees should be provided<br>- Incidence/frequency of cruel disciplinary actions by teachers<br>- Whether privacy respected when visiting a health care facility<br>- Self-reported (by individual) level of human recognition received in institutions |

An Index

Given these distinct domains of human recognition sources, a natural approach to measurement is to create a composite index that combines indicators from each of the three domains. A composite index measures the "aggregate" level of human recognition an individual is receiving from all the domains in her life. Since the purpose of this index is to aggregate human recognition from the various domains and since it cannot be assumed that there is necessarily strong correlation among the recognition levels across domains (e.g. being treated well at work and poorly at home), the simplest way to construct the index is a weighted sum of the levels of recognition received in each domain.

A general form of the index drawn from the theoretical model in Castleman 2011a is:

$$r_{i_r} = \omega_{\mathrm{ho}} \frac{1}{\sqrt{n_{ho}}} \sum_{h=1}^{n_{ho}} \rho_{hi} r_{hi} + \omega_{\mathrm{c}} \frac{1}{\sqrt{n_c}} \sum_{h=1}^{n_c} \rho_{hi} r_{hi} + \omega_{\mathrm{in}} \frac{1}{\sqrt{n_{in}}} \sum_{h=1}^{n_{in}} \rho_{hi} r_{hi}$$

where the $\omega$s are weights assigned to each domain that reflect the relative influence a given domain has on one's overall level of recognition; $r_{hi}$ is the amount of recognition provided to individual i by individual h; $\rho_{hi}$ is the provider-specific parameter that captures differences among providers and among receivers of recognition in how a given input of recognition affects an individual's received recognition; $n_k$ is the number of individuals with whom individual i has substantial human recognition transactions (as a receiver of recognition) in domain k for k=*ho*, *c*, and *in*; the subscripts represent the three domains.

Total recognition levels, also drawn from the theoretical model, are given by

$$R_i = f(r_{i_r}, \bar{r}_i) = f(r_{i_r} + \bar{r}_i)$$

where $\bar{r}_i$ is the base level of recognition individual i has at the beginning of the period of analysis. Issues and interpretations related to this function are discussed in Castleman 2011a.

In this general expression of the index, human recognition received in each domain is a function of the human recognition received in that domain from different individuals. This expression supports theoretical understanding of the index but does not directly correlate with how empirical measurement is implemented. Actual measurement of human recognition within a given domain relies on indicators of specific interactions and occurrences and/or individuals' self-reported receipt and provision of recognition. While self-reported receipt of recognition may involve mentally adding up one's various interactions, formal summation of individual interactions is not feasible because human recognition is not directly observable. Empirical measurement does use the weighted summation across domains as expressed in the index; the applications with data from Kenya and India illustrate this approach.

A more general expression of the index that does not restrict the types of indicators used is:

$$r_{i_r} = \omega_{ho} r_i^{ho} + \omega_c r_i^c + \omega_{in} r_i^{in}$$

where $r_i^k$ represent human recognition received in domain k.

The index addresses the main challenges to measuring human recognition; of the five challenges discussed above, the only one the index does not address is the lack of a direct measure of human recognition (*challenge 1*).

Since the index can include multiple indicators that all reflect human recognition, interpreting the composite index as measuring human recognition levels is more defensible than doing so for any single indicator (*challenge 2*). To continue the earlier example, while female genital cutting could be interpreted to be measuring ethnicity or modernism, there are much weaker grounds for interpreting a composite index comprised of female genital cutting, domestic violence, whether privacy is respected at a health center, and employee conditions at the workplace as measuring something other than human recognition. Each of these variables reflects human recognition and also reflects other factors, and factor analysis is used to draw out the common variation among these variables to measure the common factor of human recognition.

Furthermore, by aggregating recognition levels from each domain, the index can capture all the major domains in which human recognition transactions occur, assuming availability of sufficient data. This ensures that the index does not miss significant areas of an individual's life where human recognition is relevant (*challenge 3*).

Combining multiple indicators in the index also enables both subjective and objective indicators to be measured if data for both types of indicator exist. For example, data used in another paper (Castleman 2011c) include both an objective indicator of human recognition (whether HIV-infected individuals eat together with other household members) and an indicator of self-reported levels of recognition received from household members. Using factor analysis to combine both types of indicators in the index value helps to address possible differences between objective and subjective measures (*challenge 4*). If data from individuals providing human recognition to targeted subjects are available, the self-reported levels of human recognition they provide to subjects can also be included in the factor analysis and measured with the index. For example, the India NFHS-3 survey used in this paper includes data from husbands on their views about the permissibility of committing violence against one's wife, forced sexual relations, and wives' role in decision-making. These indicators are included in the measurement of human recognition their wives receive in the household, i.e. in the $r_i^{ho}$ term, to capture the level of recognition that the women's husbands perceive themselves to be providing to their wives.

The composite index also allows for selection of context-specific indicators of human recognition for each domain, as the illustrative indicators in Table 1 and the various indicators used in the applications (Table 2, and Table 8) illustrate. Thus, when applying the index in different contexts, appropriate indicators can be selected for each context (*challenge 5*). Comparisons of index values across contexts may not be meaningful when different indicators are used, but relative changes in index values could potentially be compared across contexts, as could correlations between index values and other variables.

In addition to supporting research on human recognition, this index can be useful for development programs as well. By capturing in one measure the human recognition levels of program beneficiaries, the index provides a relatively simple way for programs to assess the human recognition status and needs of a targeted population and to track the impacts a program has on human recognition. The process of identifying indicators to form the index for a particular program population may also assist program staff in understanding the critical issues and factors that affect the targeted population's human recognition and in considering a program's potential impacts on recognition and the role recognition plays in the program's desired outcomes, e.g. health, income, education. Depending on the type of program interventions, programs may choose to focus on and measure only one or two domains, rather than all three. To minimize the data collection burden on program staff and ensure sustained collection and use of human recognition information, collection and analysis of human recognition data could be integrated into existing program monitoring and evaluation systems. And when possible, human recognition can be tracked using data that are already being collected for other purposes, such as health care provider behaviors, in order to minimize additional data collection.

Application of the index does present practical challenges. The primary challenge is the data requirement. Existing datasets may not have sufficient information about indicators in every domain. Where data do exist on a sufficient number of indicators, they may often be cross-sectional data, yet for empirical research, panel data allow more fruitful analysis of the determinants and effects of changes in human recognition within individuals. While use of existing data or data being collected for other purposes should be maximized, especially for program purposes, in some cases obtaining sufficient data from all domains will require special effort or resources. On the other hand, even when data from all domains are not available, valuable insights and results can still be gained from examining one domain, as the analysis using the Kenya DHS data in this paper illustrates.

Another challenge to applying the full index is assignment of weights to the different domains, i.e. the $\omega_k$ values. Qualitative information from study subjects or other individuals about the relative import and value of various interactions can help estimate the weights, but since human recognition is an unobserved latent variable, ultimately the assignment of weights for each domain relies on the judgment of individuals applying the index. This is likely to lead to measurement error for values of total human recognition. This can be mitigated to a certain extent by estimating models using different combinations of weights to ensure greater robustness of results, as is done in the empirical applications undertaken here.

Empirical Method for Applying the Index

Applying the index requires a two-step process: the different variables used to measure recognition within each domain need to be combined to generate the level of recognition received in the domain, and these resulting domain-specific measures of recognition need to be combined to generate the total level of recognition received. For the latter process, the weighted values of recognition from each domain can simply be added together. Measures of recognition received in a given domain may not be correlated with measures of recognition from another domain, e.g. the case of being treated well at work and poorly at home.

For a number of reasons, measurement of human recognition within a given domain requires a different approach for combining the variables since they cannot just be summed. First, the variables often have different units and scales and adding them together is often not feasible or meaningful. For example, the following three measures of human recognition cannot be meaningfully summed together: incidences of domestic violence, self-reported level of respect received (on a 4-point scale), and number of reasons one's husband believes a wife is justified in refusing sexual relations. Secondly, it is likely that measures of human recognition *within* a given domain are correlated with each other. For example, as the India and Kenya survey data used in this paper illustrate, whether a woman has been humiliated by her husband in front of others is correlated with whether she has been physically beaten by her husband; in the India data the correlation is 0.41. Using econometric notation, this can be seen as follows:

$$humiliated_{\mathrm{i}} = \lambda_{1\mathrm{h}} r_i^{ho} + \beta_{\mathrm{h}} \mathrm{X}_{\mathrm{i}} + \delta_{\mathrm{ih}}$$

$$beaten_{\mathrm{i}} = \lambda_{1\mathrm{b}} r_i^{ho} + \beta_{\mathrm{b}} \mathrm{Y}_{\mathrm{i}} + \delta_{\mathrm{ib}}$$

where $r_i^{ho}$ is human recognition received by individual *i* in the household, $\mathrm{X}_{\mathrm{i}}$ and $\mathrm{Y}_{\mathrm{i}}$ are vectors of other determinants of $humiliated_{\mathrm{i}}$ or $beaten_{\mathrm{i}}$, and $\delta_{\mathrm{ih}}$ and $\delta_{\mathrm{ib}}$ are the error terms or unique factors.

Even if $\mathrm{X}_{\mathrm{i}}$ and $\mathrm{Y}_{\mathrm{i}}$ are entirely distinct and uncorrelated with each other and if $\delta_{\mathrm{ih}}$ and $\delta_{\mathrm{ib}}$ are uncorrelated, the covariance between the two measures will be positive, as $\mathrm{cov}(humiliated, beaten) = \lambda_{1\mathrm{h}}\lambda_{1\mathrm{b}}\sigma_{\mathrm{r}}^2 > 0$, where $\sigma_{\mathrm{r}}^2$ is the variance of $r_i^{ho}$. When there is strong correlation between the two variables, summing them to create a composite variable is not efficient because the common variation in the variables is not being exploited.

Thirdly, as discussed under challenge 2 above, some indicators used for human recognition reflect other factors in addition to human recognition, suggesting that a method other than simply aggregating the indicators within a domain is needed to capture and measure the parts of the indicators reflecting human recognition.

A method is needed to identify and estimate the common factor of human recognition from a set of indicators for a particular domain. Exploratory factor analysis can be utilized for this purpose to combine the indicators within a domain and draw out the common factor of human recognition based on the common variation among the indicators. This approach is similar to the method Kishor uses to estimate the effect of women's empowerment on child health and survival outcomes in Egypt. She uses exploratory factor analysis scores as explanatory variables in logit models for child immunization and child survival (Kishor 2000).

The factor analysis methods[3] used here evaluate the amount of variation across observations that the indicators have in common and use this information to identify the principal factors for a set of indicators. The square of the correlations among the indicators is initially used as an estimate of the proportion of each variable's variation that is explained by the identified common factors; these proportions are called communalities. Communalities are used to calculate the factor loadings, $\lambda$ in the example given above, for each indicator, essentially weighting the indicator based on the estimated contribution each identified factor plays in explaining the variation in the indicator. The sum of the squares of the factor loadings for a given measure equals the communality for that measure. In iterated principal factors the communalities are then re-estimated iteratively using only the factors that are significant to arrive at the final factor loadings. The factor loadings are then used to generate scoring coefficients that, together with the observed variable values, produce for each observation a single measure of human recognition that each individual receives in the given domain. The vector of scoring coefficients for a set of observations are generated by $\mathbf{s} = \mathbf{R}^{-1}\boldsymbol{\lambda}$, where $\mathbf{s}$ is the vector of scores, $\mathbf{R}^{-1}$ is the inverse of the correlation matrix among the measures, and $\boldsymbol{\lambda}$ is the vector of factor loadings based on the communalities. The resulting factor scores that estimate the measure of the factor for each observation are normalized to have

[3] Principal factors, iterated principal factors, and principal components factors methods yield similar results for the regression models estimated in these applications.

approximately mean 0 and standard deviation 1. The factor score may contain measurement error since it is generated from a set of observable indicators that do not perfectly measure human recognition, but the score should contain less measurement error than the individual indicators do. (For more detailed treatment of the theory behind factor analysis, see Bollen 1989 and Loehlin 1998.)

To obtain a value for the full index, the factor scores for each of the domains are then weighted and aggregated to generate a measure of total human recognition received. This measure can be used in regression analysis or by programs to assess recognition levels and track changes over time. For study of human recognition receipt in a particular domain, the individual domain scores can also be used without aggregating them in the full index. The Kenya application in this paper is an example of using just one domain, and the India application is an example of using all three domains.

Factor analysis can be used when there is correlation among the measures, but it does require certain assumptions about lack of correlation among factors for models with multiple factors, and lack of correlation among error terms in the measurement models ($\delta_{ih}$ and $\delta_{ib}$ in the example above). For the data used here, these assumptions appear valid as discussed in Sections V and VI.

The biggest challenge to using factor analysis for the index is interpreting the factors. In cases where only pre-existing indicators that were not specifically designed to measure human recognition are used, such as the applications in this paper, identification of the factors can be open to interpretation, and the factor one researcher interprets to be human recognition may be interpreted by another researcher to be empowerment or another related concept. However inclusion of a range of indicators for human recognition – such as health care worker behavior, privacy at clinics, instances of humiliation and violence, and attitudes – helps to minimize this challenge. Interpretation of factors is less of a problem when indicators specifically designed to measure human recognition are used, such as self-reported levels of recognition received (see Castleman 2011c). Since such indicators are specifically designed to measure recognition, interpreting a factor to be recognition requires a weaker assumption. Factor analysis has been used with both types of data. See Kishor 2000 and Green and Weisskopf 1990 for examples of using factor analysis with pre-existing variables to measure women's empowerment[4] and industry characteristics respectively; see Phipps and Clark 1995 and Odimegwu 1999 for examples of using it with variables specifically designed to measure unobserved variables, attitude toward economics and attitude toward family planning respectively.

To demonstrate how the index can be applied in different contexts, the next two sections apply it to measurement of human recognition using distinct datasets from two countries with distinct cultures, namely India and Kenya. The same index is applied to both datasets, but indicators within the index differ based on context and availability of data. Index measures are used to test the hypothesis that human recognition is a significant determinant of health outcomes.

[4] Some variables in Kishor 2000 were designed specifically to measure empowerment.

## V. Empirical Application with India NFHS-3 Data

Data and Variables

The first application of the index uses recently available data from the third India National Family and Health Survey (NFHS-3) carried out in 2005-6. The survey is also referred to as the Demographic and Health Survey. NFHS-3 was initiated by the Government of India and conducted by the International Institute for Population Sciences and partner institutions with technical assistance from Macro International and funding from the U.S. Agency for International Development (USAID) and other donors. NFHS-3 surveyed 124,385 women aged 15-49 and 74,369 men aged 15-54 in all 29 states of India. Data were collected about demographic, health, nutrition, fertility, and socio-economic status. Information about domestic violence, sexual activity, behavior of health care workers, and household decision making was also collected from both female and male respondents.

In order to include data collected from women's husbands about the human recognition they provide to their wives, the analysis includes only data from women whose husbands were also surveyed. Identifiers of observations are used to link data from women's responses to data from their husband's responses. Since the indicators for which data were collected from men required a smaller sample size than the indicators for which data were collected from women (which included information about women's children), NFHS-3 collected data from fewer men (74,369) than women (124,385). In order to obtain state-level estimates of HIV prevalence from seven states identified by the Government of India as having higher HIV prevalence, all men present in surveyed households in these states were eligible for data collection, as were all women. In the remaining 22 states, only men from a randomly selected subsample of households were eligible for data collection, while all women were eligible in these states. Therefore, since these seven states are overrepresented in the men's data and this analysis only includes women whose husbands' data were collected, women from these seven states are overrepresented in the sample, namely Andhra Pradesh, Karnataka, Maharashtra, Manipur, Nagaland, Tamil Nadu, and Uttar Pradesh[5]. As these states are fairly diverse geographically and culturally, it is not expected that this overrepresentation significantly affects the results, but it is possible that results would differ with more even representation among states.

Furthermore, while NFHS-3 collected data from all women aged 15-49, only married women are included here in order to include husbands' responses. Results may differ for unmarried, widowed, or divorced women. Exclusion of non-married women also means that the women in this sample are on average older than the overall NFHS-3 sample.

Since body mass index (BMI) of pregnant and non-pregnant women cannot be meaningfully compared and the World Health Organization's BMI cutoffs for malnutrition do not apply to pregnant women, pregnant women are not included in the analysis. It is possible the levels of human recognition or the frequency or effects of human recognition transactions differ systematically between pregnant and non-pregnant women. For example, possibly women who receive less human recognition have less control over fertility and are more likely to experience frequent pregnancies, in which case the sample used here would have higher overall human recognition levels than the

[5] Uttar Pradesh does not have a high prevalence of HIV but data on HIV prevalence were collected.

population of women in India as a whole (including pregnant women). Or perhaps the health and nutritional status of pregnant women are more sensitive to human recognition transactions since pregnant women may require particular care from household members or health facilities. The impact of human recognition on nutritional status measured here would not capture such an effect because only non-pregnant women are included in the sample. Lastly, women in India may receive systematically higher or lower levels of human recognition during pregnancy – higher because they are given special attention, they have a higher position in the household, and their needs paid greater attention to; or lower because they are mistreated and their additional needs and constraints (e.g. avoiding strenuous housework) are not valued. If there are such systematic differences, the results presented here do not capture them because pregnant women are not included in the sample.

Furthermore, the NFHS-3 randomly selected a sub-sample of women to answer questions about domestic violence and only these women are included in the analysis. The total sample used in the analysis is 26,125 women. When anemia is included as a control variable, the sample decreases to 24,360 women because women in the state of Nagaland were not tested for anemia in NFHS-3.

From the NFHS-3 data, variables were identified that measure human recognition received by women in each of the three domains: household, community, and institutions. Table 2 lists the variables used and summarizes the status of each variable for the sample of respondents included in this analysis. Data were collected from women for all variables, except for the three variables marked as having been collected from men.

**Table 2: India NFHS-3: Variables Measuring Women's Human Recognition**

| Domain | Variable | Respondent Status |
|---|---|---|
| *Household* | Ever experienced physical violence by husband | Yes: 32% |
| | How often experienced physical violence by husband in past 12 months | Often/sometimes: 20% |
| | Ever experienced physical violence by other household member | Yes: 1% |
| | How often humiliated, threatened or insulted by husband | Often/sometimes in past 12 months: 10% |
| | Ever forced to have sexual relations with husband against will | Yes: 7% |
| | How often forced to have sexual relations with husband against will in past 12 months | Often/sometimes: 5% |
| | Number of reasons (out of 7)[6] the woman believes a husband is justified in beating his wife | At least 1 reason: 58%<br>All 7 reasons: 7% |
| | Not permitted by husband to meet female friends and/or family members | Not permitted: 16% |
| | Husband's view of who should make decisions related to wife's visits to her family members and related to spending of money the wife earns *(husband's response)* | Wife should decide both on own: 7%<br>Husband should decide both on own: 9% |
| | Number of reasons (out of 7) husband believes a husband is justified in beating his wife *(husband's response)* | At least 1 reason: 50%<br>All 7 reasons: 2% |
| | Number of reasons (out of 3)[7] husband believes a wife is justified in refusing to have sex *(husband's response)* | All 3 reasons: 74%<br>No reasons: 4% |
| *Community* | Ever experienced physical violence from other relative or boyfriend | Yes: 0.4% |
| | How often experienced physical violence from other relative or boyfriend in past 12 months | Often/sometimes: 0.1% |
| | First experience of forced sexual relations was by a family friend, other friend/acquaintance, non-household relative | Yes: 0.4% |
| *Organizations and Institutions* | Privacy respected during visit to health care facility during past 3 months | Not respected: 10% (of those visiting facility) |
| | Health care facility staff were responsive to problems and needs in visit during past 3 months | Not responsive: 4% (of those visiting facility) |
| | Spoken to nicely during contact with nurse or other local health worker in past 3 months | Somewhat/not nicely: 21% (of those contacted by worker) |
| | Ever experienced physical violence from teacher, employer, police, or soldier | Yes: 1% |
| | First experience of forced sexual relations was by teacher, employer, religious leader, police, or soldier | Yes: 0.02% |

[6] Reasons are: wife goes out without informing husband; wife neglects the children; wife argues with husband; wife refuses to have sex; wife burns the food; wife is unfaithful; wife is disrespectful to in-laws.
[7] Reasons are: husband has STD; husband has other women; wife is tired or not in the mood.

For most of these variables, the rationale for inclusion in the human recognition index is clear. Humiliation and other types of emotional violence is perhaps the variable that comes closest to directly measuring human recognition; to humiliate someone is to degrade and devalue her as a human being. Physical and sexual violence are manifestations of high magnitudes of negative recognition because domestic violence – and violence more generally – often involves objectification of the victim of violence. However, it is worth noting that a woman's level of recognition may itself affect her response to questions about whether she has been humiliated or forced to have sexual relations since it may influence her definition of what constitutes humiliation or force, i.e., her recognition level may influence where she draws the line between what she considers appropriate and inappropriate behavior by her husband or others.

Women's beliefs about permissibility of violence are included because the belief that violence by a husband is permissible is expected to be both an enabling factor and a result of receiving low human recognition. Martin et al. (2002) suggest that in India viewing domestic violence to be permissible is a contributing factor to the incidence of violence and subjugation. Conversely, continued receipt of low recognition may lead an individual to accept the belief that being beaten by one's husband is justified.

Whether women are permitted by their husbands to meet friends and family is included because it reflects an aspect of human recognition in the household – the extent to which respondents are acknowledged and recognized as autonomous individuals. This is an indicator of a type of empowerment that involves human recognition interactions.

The variables measuring husbands' views about the permissibility of violence against one's wife, under what circumstances wives are justified in refusing to have sex, and women's autonomy in decision-making reflect the human recognition *provided* to women by male household members. A husband's views about a wife's rights regarding such basic human issues as being beaten, refusing sexual relations, and visiting family members reflect the extent to which the husband acknowledges and values his wife's basic needs, rights, and preferences.

Variables about the behavior of health care staff towards women (privacy respected, responsive to needs and problems, and speaking nicely) measure the level of human recognition women receive from health care institutions. The variables on privacy and responsiveness were only collected from women who had visited a health care facility in the past three months and the variable on speaking nicely was only collected from women who had contact with a community health worker (Auxiliary Nurse-Midwife, Lady Health Supervisor, Anganwadi Worker, Accredited Social Health Activist, or Multi-Purpose Health Worker) in the past three months. Including only these women in the analysis would reduce the sample size from 26,125 to 1,834. In order to maintain a large sample size, respondents who did not have values for these variables (because they had not had contact with the health care staff in question in the past three months) were assigned neutral values of 0; respondents who reported receiving positive human recognition (privacy respected, staff were responsive to needs, and spoken to nicely) were assigned negative values (e.g. -1), and women who reported receiving negative human recognition were assigned positive values (e.g. 1). The reason for reversing the signs of the recognition level is that all of the human recognition variables are scaled such that higher values are assigned for lower levels of human recognition.

Both respondents who have and those who have not accessed health services in the past three months are included in the analysis, and only respondents who visited a health center or who were visited by a health worker in the past three months are assigned positive or negative values for these variables.  It is possible that women who have not had contact with the health system in the past three months have systematically higher or lower recognition levels than those that have.  For example, those not in contact with the health system may be healthier (and related to greater health also have higher levels of recognition); or they may receive lower levels of recognition from their household or from the health facilities at an earlier visit (which would not be captured in these variables), reducing the likelihood of their accessing health services.

The basic model examining the role human recognition plays in women's health status is:

$$BMI_i = \alpha + \beta_1 wealth_i + \beta_2 education_i + \beta_3 agei + \beta_4 anemia_i + \gamma recognition_i + e_i$$

Note that since the data are cross-sectional, the model tests whether individuals with greater levels of human recognition have higher BMIs, controlling for other factors, not whether a change in recognition received by a given individual is associated with a change in BMI.

The outcome variable is BMI, which is calculated by dividing an individual's weight in kilograms by the square of her height in meters.  BMI is a commonly used measure of adult nutritional status.  The World Health Organization (WHO) has established BMI cutoffs for non-pregnant adults for severe malnutrition (BMI < 16.0 kg/m$^2$), moderate malnutrition (16.0 $\geq$ BMI < 17.0), mild malnutrition (17.0 $\geq$ BMI < 18.5), normal nutritional status (18.5 $\geq$ BMI < 25.0), overweight (25.0 $\geq$ BMI < 30.0), and obese (BMI $\geq$ 30.0) (WHO 1999, WHO 1998).

Malnutrition has been identified as “the single leading global cause of health loss” (Ezzati et al. 2002) and is closely linked to morbidity and maternal mortality. Among women and children, being underweight is responsible for 2 million deaths each year and between 80 to 138 million disability-adjusted life years (DALYs), a commonly used measure of the burden of disease that reflects the years of life lost to death and disease (Black et al. 2008, Ezzati et al. 2002). Malnutrition also diminishes productivity by decreasing physical and mental capacity, and the reduced productivity limits economic development at both individual and national levels. (See Victora et al. 2008 for a recent synthesis of the long term impacts of malnutrition.) India has one of the highest rates of low BMI among women, and has been identified as one of the countries of critical concern in this area by public health experts (Black et al. 2008).

The mean BMI of the sample is 21.4 kg/m$^2$, with a standard deviation of 4.26, and 27% of women have BMIs that fall in the malnourished category, i.e. BMI < 18.5 kg/m$^2$. The sample of women has a higher BMI than the total sample of women surveyed in NFHS-3 (mean BMI 20.5 kg/m$^2$). This is likely because this sample only included women whose husbands also provided data so only married women were included. Married women tend to be older than the overall population of women aged 15-49, and older women tend to have higher BMIs than younger women do. For example, in this sample there is a positive correlation of 0.24 between age and BMI.

Four control variables are included in the model to control for socio-economic and health factors that affect nutritional status. The *wealth* variable uses a wealth index that classifies respondents into five categories of wealth, based on 13 household variables[8]. In the sample used in this analysis, 28% of respondents fall in the poor or poorest categories, 21% in middle category, and 51% in the rich or richest categories. There is also a *livingstandard* variable available that applies a standard of living index to classify respondents into three categories based on the facilities and possessions in the respondent's house; the index is comprised of 30 variables. In the sample used in this analysis, 20% of respondents are in the low category, 33% in the medium category, and 46% in the high category, with 0.5% of respondents not *de jure* residents of the household. Some variables are part of both the wealth index and the standard of living index, there is a very high correlation between the two (0.76), and substituting *livingstandard* for *wealth* in the regressions does not alter the results significantly. Given the strong overlap and correlation between the two variables, both are not included in the regression models.

*Age* is another control variable, and respondents' ages range from 15 to 49, with a mean age of 32 (sd = 7.48). *Education* measures the number of years of schooling the respondent has had; mean years of education for the sample is 5.4 (sd = 5.2). Anemia status is included to control for other health and nutritional factors. *Anemia* measures the level of anemia, with 2% of women severely anemic, 13% moderately anemic, 38% mildly anemic, and 48% not anemic.

The *recognition* variable measures human recognition received in household, community, and institution domains, using factor scores from variables in Table 2. As detailed in the results below, separate factor analyses are carried out for human

[8] The wealth index is composed of the following variables: drinking water source, non-drinking water source, toilet facility, household electrification, household possessions, type of cooking fuel, main floor material, main roof material, main wall material, type of windows, number of de jure members per sleeping room, house ownership, household member having a bank or post office account.

recognition variables in the household, community, and institution domains. Factor scores for human recognition in each domain are then generated. These comprise the $r_i^{ho}$, $r_i^{c}$, and $r_i^{in}$ variables in the index. Once these have been calculated, the next step is to identify weights (ω's) for the domain values to generate a total value for the overall human recognition index.

It is not possible to empirically test the values of the weights, and there is likely to be some measurement error for any set of weights chosen. Therefore, the regression models are run using different combinations of weights to check for robustness. Results are reported for a model in which the value of the recognition variable is $r_{i_r} = 0.5\, r_i^{ho} + 0.15\, r_i^{c} + 0.35\, r_i^{in}$. While other weights yielded similar results, this set of weights was selected because given the number and content of the variables used in each domain and based on the variation in these variables among respondents, it is expected that human recognition in the household will have the greatest impact, followed by recognition in institutions and the community respectively. Furthermore, it is expected that most respondents have more and closer interactions with household members than with individuals from their communities or institutions, suggesting that the household domain is likely to have the greatest impact on overall recognition levels.

To address possible endogeneity of *recognition* and *anemia* due to omitted variables bias and/or simultaneity, three instrumental variables are used initially: *caste*, *occupation*, and *religion*. The Hansen J test statistic used to test exogeneity of instruments indicates that *religion* is not a valid instrument because it is not exogenous to the models, so results are reported using only *caste* and *occupation* as instruments[9]. Since these are cross-sectional data, differencing is not possible, and instrumental variables are used to address endogeneity. *Caste* categorizes respondents' castes using the classifications common in India: 18% are scheduled (lowest) caste, 13% are scheduled (lowest) tribe, 37% are other backward caste, and 32% are other (higher) castes. *Occupation* refers to the occupations of the respondents: 56% are not working, 23% work as agricultural laborers/employees, and 21% work in other sectors. *Religion* refers to the religion of the respondents: 79% are Hindu, Sikh, or Jain; 10% are Muslim, 9% are Christian, and 2% are other religions. The rationale for using these variables as instruments is discussed along with sources of endogeneity in the section below on regression results.

All estimates are generated using Stata 10.

Factor Analysis Results

To obtain values for human recognition for each observation, factor analysis is carried out for each domain. The factor analysis for human recognition received in the household is based on the following model:

$$\textit{viol\_husb}_i = \lambda_{1vh}\textit{hhrecognition}_i + \lambda_{2vh}\mu_i + \lambda_{3vh}\nu_i + \delta_{ivh}$$
$$\textit{viol\_freq}_i = \lambda_{1vf}\textit{hhrecognition}_i + \lambda_{2vf}\mu_i + \lambda_{3vf}\nu_i + \delta_{ivf}$$
$$\textit{viol\_oth}_i = \lambda_{1ov}\textit{hhrecognition}_i + \lambda_{2ov}\mu_i + \lambda_{3ov}\nu_i + \delta_{iov}$$
$$\textit{emot\_abuse}_i = \lambda_{1ea}\textit{hhrecognition}_i + \lambda_{2ea}\mu_i + \lambda_{3ea}\nu_i + \delta_{iea}$$
$$\textit{force\_sex}_i = \lambda_{1fs}\textit{hhrecognition}_i + \lambda_{2fs}\mu_i + \lambda_{3fs}\nu_i + \delta_{ifs}$$
$$\textit{force\_sex\_freq}_i = \lambda_{1ff}\textit{hhrecognition}_i + \lambda_{2ff}\mu_i + \lambda_{3ff}\nu_i + \delta_{iff}$$
$$\textit{viol\_ok}_i = \lambda_{1vo}\textit{hhrecognition}_i + \lambda_{2vo}\mu_i + \lambda_{3vo}\nu_i + \delta_{ivo}$$
$$\textit{visit}_i = \lambda_{1v}\textit{hhrecognition}_i + \lambda_{2v}\mu_i + \lambda_{3v}\nu_i + \delta_{iv}$$

[9] All three instruments are tested with Stata's orthog command that uses a C test to check the difference in J statistics with and without the suspect instrument, and only *religion* is identified as problematic. Without *religion*, the Hansen J statistic indicates the other two instruments are exogenous.

$$husb\_decide_i = \lambda_{1hd}hhrecognition_i + \lambda_{2hd}\mu_i + \lambda_{3hd}\nu_i + \delta_{ihd}$$
$$husb\_beat\_ok_i = \lambda_{1hb}hhrecognition_i + \lambda_{2hb}\mu_i + \lambda_{3hb}\nu_i + \delta_{ihb}$$
$$husb\_sex_i = \lambda_{1hs}hhrecognition_i + \lambda_{2hs}\mu_i + \lambda_{3hs}\nu_i + \delta_{ihs} \qquad i = 1\ldots26{,}125$$

$Hhrecognition_i$ is the latent variable (factor) of human recognition that individual i receives in the household, the λs are the factor loadings, $\mu_i$ and $\nu_i$ are other latent variables (i.e. other factors) underlying the eleven measures, and $\delta_{ix}$ are unique factors (i.e., error terms) that affect the measures. The letter subscripts (vh…hs) refer to the eleven measures. Exploratory factor analysis assumes the unique factors, $\delta_{ivh}\ldots\delta_{ihs}$, are uncorrelated. This appears to be a reasonable assumption for these data because the common sources of variation in each of the measures are captured in the three factors, and the unique factors reflect other determinants, which are not likely to be strongly correlated across the 11 measures. For example, hostile or restrictive household environments faced by some women that contribute to very limited autonomy and also to regular incidences of violence would be captured in the $hhrecognition_i$ (she receives low levels of recognition) and in the other factors (which may be age, empowerment, or income), not in the unique factors.

Three factors ($hhrecognition_i$, $\mu_i$, and $\nu_i$) are included in this model because, applying the conventional cutoff of Eigenvalue $\geq 1$ as the threshold for keeping a factor, factor analysis yields three significant factors. Table 3 presents the results. The results are consistent with interpretation of Factor 1 to be human recognition received in the household. All the factor loadings for Factor 1 are positive, which is expected since all of the measured variables have higher values for receipt of lower levels of recognition. The common factor is interpreted to be *low* levels of human recognition.

**Table 3: Factor Analysis Results for Human Recognition Received in the Household: India NFHS-3**

| Variable | Factor 1 Loadings | Factor 2 Loadings | Factor 3 Loadings | Uniqueness |
|---|---|---|---|---|
| violence by husband | 0.7705 | 0.1808 | -0.3267 | 0.2669 |
| frequency of violence | 0.8363 | 0.1482 | -0.3260 | 0.1723 |
| violence by others in hhold | 0.1876 | -0.0256 | -0.1063 | 0.9529 |
| humiliated, emotional abuse | 0.6652 | 0.1109 | -0.3076 | 0.4506 |
| forced sexual relations | 0.7257 | -0.4369 | 0.4642 | 0.0670 |
| frequency of forced sex | 0.7306 | -0.4364 | 0.4571 | 0.0669 |
| reasons woman believes it's okay for a husband to beat his wife | 0.1616 | 0.5189 | 0.2501 | 0.6421 |
| permitted to visit friends, family | 0.3695 | 0.1588 | -0.2058 | 0.7959 |
| husband's view of decision-making by wife | 0.0841 | 0.3549 | 0.3694 | 0.7305 |
| reasons husband believes it's okay to beat one's wife | 0.1640 | 0.5938 | 0.3519 | 0.4967 |
| reasons husband believes a wife can refuse sex | 0.0903 | 0.3755 | 0.2767 | 0.7743 |
| *Eigenvalue* | *3.04* | *1.36* | *1.19* | |

When iterated principal factor analysis is used instead, re-estimating the factor loadings (estimates of the $\lambda$ parameters) using only the significant factors as discussed above, then one significant factor remains (results not shown)[10]. The loadings of that factor for all eleven measures are consistent with interpretation of the factor to be human recognition received in the household.

The uniqueness given for each measured variable equals unity minus the communality value for the variable. Communality measures how much of the variation in a measured variable is correlated with variation in the other measured variables, which is what the factor captures. So uniqueness reflects how much variation in the measured variable is not explained by the factors. Uniqueness is quite high for violence by others

[10] For all three domains, using iterated principal factors leads to somewhat different coefficient estimates in the subsequent regressions but does not change the signs or statistical significance of any of the results.

in the household, which suggests that these three factors do not explain variation in this variable very well. This may be because all of the other variables relate to recognition received from the woman's husband and this relates to recognition from other household members. Indeed, the simple correlation between violence by one's husband and violence by others in the household is only 0.10. Also, the percentage of women in the sample reporting violence by others in the household is quite low (1%).

The factor loadings are combined with the values of the eleven measures to produce a factor score for human recognition for each of the 26,125 women. These factor scores, normalized to have approximately mean 0 and standard deviation 1, are estimates of $r_i^{ho}$ in the index for each observation. They can be used in empirical models as a measure of human recognition received in the household, or can be combined in the index with factor scores from the other domains to generate an estimate of total recognition for each observation. Summary statistics for the factor score *hhrecognition* are: mean = -0.003745, sd = 0.9990.

The factor analysis for human recognition received in the community is based on the following model:

$$force\_sex\_comm_{\mathrm{i}} = \lambda_{1sc} comm\_recognition_{\mathrm{i}} + \delta_{isc}$$
$$viol\_comm_{\mathrm{i}} = \lambda_{1vc} comm\_recognition_{\mathrm{i}} + \delta_{ivc}$$
$$viol\_comm\_freq_{\mathrm{i}} = \lambda_{1cf} comm\_recognition_{\mathrm{i}} + \delta_{icf} \qquad \mathrm{i} = 1, 2\ldots..26{,}125$$

In this case only one factor is included because only one factor achieves an Eiegenvalue $\geq$ 1. Factor analysis results are given in Table 4.

**Table 4: Factor Analysis Results for Human Recognition Received in the Community: India NFHS-3**

| Variable | Factor 1 Loadings | Uniqueness |
|---|---|---|
| violence by community members | 0.8839 | 0.2167 |
| frequency of violence by community members | 0.88863 | 0.2145 |
| first experience of forced sexual relations was by community member | 0.1465 | 0.9785 |
| *Eigenvalue* | *1.59* | |

The results are consistent with the factor being human recognition received in the community, with positive factor scores for all three measured variables. Uniqueness is high for the variable measuring whether the first experience of forced sexual relations was by a community member because it is not highly correlated with other forms of violence (the correlation between it and the violence by community members is .04 and correlation with the frequency of violence is .06).

Factor loadings are combined with measured values of the three variables to obtain factor scores for each observation that are estimates of $r_i^c$ in the index. Summary statistics for the factor score *comm_recognition* are: mean = -0.002265, sd = 0.9740.

The factor analysis for human recognition received from institutions is based on the following model:

$$privacy_i = \lambda_{1p} inst\_recognition_i + \lambda_{2p}\mu_i + \delta_{ip}$$
$$responsive_i = \lambda_{1r} inst\_recognition_i + \lambda_{2r}\mu_i + \delta_{ir}$$
$$nicely_i = \lambda_{1n} inst\_recognition_i + \lambda_{2n}\mu_i + \delta_{in}$$
$$viol\_inst_i = \lambda_{1vi} inst\_recognition_i + \lambda_{2vi}\mu_i + \delta_{ivi}$$
$$force\_sex\_inst_i = \lambda_{1si} inst\_recognition_i + \lambda_{2si}\mu_i + \delta_{isi} \qquad i = 1, 2\ldots..26{,}125$$

In this case two factors are included because two factors achieve an Eiegenvalue ≥ 1. Factor analysis results are given in Table 5.

**Table 5: Factor Analysis Results for Human Recognition Received from Institutions: India NFHS-3**

| Variable | Factor 1 Loadings | Factor 2 Loadings | Uniqueness |
|---|---|---|---|
| privacy respected at health facility | 0.8747 | -0.0147 | 0.2348 |
| health care worker responsive to problems/needs | 0.8861 | -0.0140 | 0.2147 |
| spoken to nicely by nurse or local health worker | 0.3012 | 0.1379 | 0.8903 |
| violence by teacher, employer, police | -0.0337 | 0.7836 | 0.3848 |
| first experience of forced sexual relations was by teacher, employer, police, religious leader | -0.0168 | -0.6068 | 0.6315 |
| *Eigenvalue* | *1.64* | *1.00* | |

In this case, the factor loadings for Factor 1 indicate that the factor is related to human recognition received from institutions but is dominated by recognition received from the health care system. Variations in recognition from the health care system across individuals is not correlated with variations in violence experienced by teachers, employers, and police, which is why the factor loadings on the violence variables are negative and of very small magnitude. Uniqueness is not too high for the two violence variables because Factor 2 does explain some of the variation in these variables. When an Eigenvalue cutoff of 1.1 is used instead of 1.0, only Factor 1 is included and the uniquenesses for the two violence indicators increase to 0.9989 and 0.9997, while the uniquenesses for the three health care system variables remain approximately the same (results not shown) [11]. Factor 1 still appears to be human recognition but in this case variation in one set of measured variables (behavior of health workers) is not correlated with variation in the other set of variables (violence).

These results point to a potential limitation of using factor analysis to combine human recognition indicators for distinct organizations or institutions. Recognition levels provided by different institutions to a given individual may not be strongly correlated, e.g., in this case recognition provided by health care workers and recognition from community members in the form of violence. In such a case, factor analysis will not be as effective at measuring the common factor of human recognition. Other reasons for the low correlation among the two sets of variables may be that the health care variables are only non-zero for those clients who had recent contact with health care services, and because only a small proportion of respondents had experienced the types of violence in the violence variables (1% and 0.02%).

Uniqueness is high for the variable on having been spoken to nicely by local health workers, which may be because the first two questions on treatment by the health care system are asked of respondents who had visited a health facility in the past three months while the third one is asked of those who have been visited by a local nurse or health worker in the past three months. Neutral values are assigned to *privacy* and *responsive* for the women who did not visit a facility in the past three months and to *nicely* for those who have not seen a local health worker. Therefore, the variation in privacy and responsive are likely to be more highly correlated than either is with *nicely*. Supporting this explanation, the correlations between *nicely* and both *privacy* and

[11] All results that are not shown are available with the author.

*responsive* are considerably lower than the correlation between *responsive* and *privacy*.

Based on these results, the factor loadings for Factor 1 can be combined with measured values of the variables to obtain factor scores for each observation, which become estimates of $r_i^{in}$ in the index. Summary statistics for the factor score *inst_recognition* are: mean = -0.0008029, sd = 1.000.

The three factor scores, the estimates of $r_i^h$, $r_i^c$, and $r_i^{in}$, can be weighted and combined to obtain a value for total recognition received, $r_{i_r}$. As discussed above, it is not possible to empirically test what the correct weights are, so there is likely to be some measurement error with whichever weights are used. Regression models are run using a few different combinations of weights, and while the coefficient estimates on *recognition* in the models differ depending on the weights used, the signs and significance levels of the results do not differ. Results are reported below using $\omega_h = 0.5$, $\omega_c = 0.15$, and $\omega_{in} = 0.35$, that is $recognition_i = r_{i_r} = 0.5 r_i^h + 0.15 r_i^c + 0.35 r_i^{in}$. As discussed above, results are reported using these weights because it is expected that the household domain contributes the most to overall human recognition levels, followed by the institutions and community domains respectively.

Regression Results

The following model is estimated:

$$BMI_i = \alpha + \beta_1 wealth_i + \beta_2 education_i + \beta_3 agei + \beta_4 anemia_i + \gamma recognition_i + e_i$$

where *recognition* uses the factor scores and weights given above. In the model, the inverses of the factor scores are used, i.e. $recognition_i = -(0.5 r_i^h + 0.15 r_i^c + 0.35 r_i^{in})$. Because the recognition variables are scaled such that higher values signify lower human recognition levels (e.g. having experienced violence, privacy not respected at health facilities), using the inverse of the factor scores means that now higher values of *recognition* signify higher levels of recognition.

Initially the model is estimated using OLS. Results are reported in Table 6. The model is significant ($R^2$ = .24, Pr>F < .0001) and the coefficient on *recognition* is positive and significant as predicted.

**Table 6: Results of BMI Model Using India NFHS-3 Data: OLS**

| Dependent Variable | Estimation Method | Intercept | Wealth | Education | Age |
|---|---|---|---|---|---|
| BMI | OLS (heterosk.-consistent) | **12.12 (0.138)** | **1.066 (0.022)** | **0.081 (0.006)** | **0.115 (0.003)** |
| | | Anemia | Recognition | | n, $R^2$, Pr>F |
| | | **0.483 (0.031)** | **0.129 (0.036)** | | n = 24,360<br>$R^2$ = 0.24<br>Pr>F < .0001 |

Parameter estimates in bold indicate significance at the .05 level. Standard errors are given in parentheses.

However, a Hausman specification test indicates that *recognition* and/or *anemia* are endogenous (Pr > $\chi^2$ = .001).[12] Other explanatory variables are treated as exogenous in the model; for example, a Hausman specification test does not reject exogeneity of *wealth* (Pr > $\chi^2$ = .133).

There are three possible reasons why the human recognition factor score would be endogenous. The first is a simultaneity problem whereby not only is human recognition a determinant of nutritional status, but nutritional status is also a determinant of the level of human recognition one receives in the household. For example, it could be that not only

[12] The Hausman specification test tests the equivalence of coefficient estimates from the two-stage least squares estimation (which will be consistent with or without exogeneity of explanatory variables) and coefficient estimates from the OLS estimation (which will be efficient if explanatory variables are exogenous but inconsistent if they are endogenous).

does experiencing domestic violence increase the likelihood of having a worse nutritional status through poorer food intake, reduced access to health care, etc., but worse nutritional status may also increase the likelihood of experiencing domestic violence. Undernutrition can cause low energy levels, listlessness, and symptoms associated with depression. Women experiencing these symptoms may be more likely to be beaten by their husbands because of reduced capacity to complete household duties, melancholy attitudes, or other factors. Similar logic could hold for other manifestations of low human recognition such as emotional abuse.

The second possible reason for endogeneity is an omitted variable bias. There may be other variables not included in the model that are determinants of both BMI and human recognition. Unobserved characteristics of some women or their households, such as substance abuse or HIV infection within the household, could lead both to poorer nutritional status of women and to lower levels of human recognition. This would make the recognition variable correlated with the error terms in the BMI specification because variation in the omitted variable would systematically affect both the factor scores for human recognition and the unobserved determinants of BMI.

The third possible reason is measurement error caused by the weights assigned to the three domains in the index, which do not precisely reflect the relative weight each domain has in total received recognition.

The endogeneity of *anemia* is likely a problem of omitted variables. Some variables, such as dietary practices, that are not included in the model affect both BMI and anemia.

To address the endogeneity, two-stage least squares (2SLS) is used to re-estimate the specification using *caste* and *occupation* as external instruments. Intuitively, it is expected that these two variables are correlated with the combined factor scores for human recognition. Caste can influence the recognition that women receive in the household as autonomous individuals, e.g. participation in decisions about visiting family or spending money they earn as norms can differ across castes. Caste can also affect the level of human recognition received from the community and institutions, as lower caste individuals may be more likely to be treated disrespectfully in some settings (Mendelsohn and Vicziany 1998). Occupation is used as an instrument because women who work and earn money may be treated with different levels of recognition, e.g. greater participation and control in decision-making or less likely to experience domestic violence (Schuler et al. 1996), though there has also been evidence that earning income increases women's risk of domestic violence in some settings (Naved and Persson 2005).

*Caste* and *occupation* are also expected to be correlated with *anemia*. Caste affects dietary habits, including consumption of meat and other animal-source foods, which is likely to be one pathway by which it affects anemia. Caste also affects socio-economic status, which can influence anemia through factors other than wealth, which is controlled for here. In the full NFHS-2 sample, anemia rates were 51% among high caste women, 58% among scheduled (lowest) caste women, and 69% among scheduled tribe women, and the differences are greater when only severe anemia is considered (IIPS 2007). Occupation can also affect access to food and dietary habits that contribute to anemia, and in the reverse direction anemia may affect the capacity to perform occupations requiring physical labor.

Evidence that these two variables are correlated with the two endogenous variables, *recognition* and *anemia*, can be found in the highly significant Anderson canonical correlation likelihood ratio statistic ($\chi^2$ p < .0001).

Further evidence is found in the highly significant coefficients on both the instrumental variables when *recognition* is regressed on these variables, the control variables, and other determinants of the human recognition measures. Results of this regression with *recognition* as the dependent variable are given in Table 7. The overall regression is significant, and all coefficients are significant, though the $R^2$ is low. Additional variables may need to be added to fully examine the determinants of human recognition; the purpose of the model here is to examine the relevance of the variables selected as instruments in the BMI model. *Religion* is also initially included as an instrument, but as mentioned above, the Hansen J test statistic indicates it is correlated with the error and therefore not a valid instrument.

**Table 7: Determinants of Human Recognition and Relevance of Instruments**

| Dependent Variable | Estimation Method | Intercept | Wealth | Education | Age | Caste |
|---|---|---|---|---|---|---|
| recognition | OLS (heterosk.-consistent) | **-0.3134**<br>(0.0264) | **0.0778**<br>(0.0037) | **0.0189**<br>(0.0010) | **-0.0038**<br>(0.0005) | **0.1644**<br>(0.0039) |
| | | Occupation | Religion | | | n, $R^2$, F |
| | | **-0.0325**<br>(0.0048) | **0.0228**<br>(0.0077) | | | n=26,125<br>$R^2$=0.09<br>F = 441 |

Parameter estimates in bold indicate significance at the .05 level. Standard errors are given in parentheses.

Similarly, *caste* and *occupation* are also significant determinants of *anemia* when a similar model is run for *anemia* (results not shown).

For *caste* and *occupation* to be valid as instruments, they should not be correlated with the error terms in the BMI specification. There is no simultaneity bias between any of these variables and BMI. Since the instruments are relatively fixed variables for a given individual, BMI is not likely to be a determinant of these variables. If one's BMI changes, it will not affect one's caste, and only in exceptional cases would it be expected to change one's occupation. Omitted factors such as substance abuse, HIV infection, or even attitudes of respondents or their husbands may affect BMI and may affect human recognition levels, but are unlikely to affect caste. There could, however, be some effect on occupation, which would reduce its validity as an instrument.

However, the causation may also work in the opposite direction, and some omitted determinants of BMI could be affected by the variables selected as instruments, which could pose problems with using them as instruments. That is, an individual's caste, occupation, or religion could affect BMI through pathways other than human recognition. Some of these pathways may involve variables not included in the model, which are therefore part of the error term. For example, caste may influence dietary intake of animal-source foods, which in turn may affect BMI; in India Muslims and Christians and certain Hindu castes eat meat while other Hindu castes do not, and certain castes are more likely to raise dairy cows and therefore consume more dairy products. If the instruments significantly affect BMI through pathways such as these, it reduces the validity of the instruments.

In support of the validity of these instruments, the Hansen J test statistic is insignificant ($\chi^2$ p = .233 when *religion* is removed as an instrument), providing evidence that both *caste* and *occupation* are exogenous to the model and uncorrelated with the error term.

Given that these are cross-sectional data, instrumental variables appear to be the best available option for dealing with endogeneity, and these variables seem to be the strongest instruments available in the dataset. Therefore, for this application of the index the model is estimated with two-stage least squares using these two external instruments. When panel data are available, differencing can be used to address endogeneity, and leading variables can be used as instruments (Castleman 2011c).

Estimation results from the BMI model using two-stage least squares with these two instruments are given in Table 8.

**Table 8: Results of BMI Model Using India NFHS-3 Data: 2SLS**

| Dependent Variable | Estimation Method | Intercept | Wealth | Education | Age |
|---|---|---|---|---|---|
| BMI | 2SLS (heterosk.-consistent) | 10.49<br>(8.39) | **0.525**<br>**(0.195)** | -0.041<br>(0.035) | **0.136**<br>**(0.010)** |
| | | Anemia | Recognition | | n, Pr>F |
| | | 1.522<br>(2.822) | **5.968**<br>**(1.191)** | | n = 24,360<br>Pr>F < .0001 |
| Endogenous variables: Recognition, Anemia<br>External instruments: Caste, Occupation | | | | | |

Parameter estimates in bold indicate significance at the .05 level. Standard errors are given in parentheses.

An F test indicates that the specification has significant explanatory power at the < .0001 level. The coefficient on *recognition* is positive and significant at the < .001 level, indicating that having higher levels of human recognition is a significant determinant of better nutritional status, controlling for the other variables. The magnitude of this effect is quite large. The standard deviation of the full recognition variable is 0.64, so having a 1 standard deviation higher recognition level corresponds to having a BMI that is 3.8 $kg/m^2$ higher, controlling for the other variables. Average height of respondents is 1.5 meters so this corresponds to an 8.6 kg. greater weight

Higher levels of wealth and greater age are also significantly associated with higher BMI. When the standard of living index is used in place of the wealth index, its coefficient is positive and highly significant (results not shown), but when both variables are included only *wealth* is significant (results not shown). This is likely because the wealth and standard of living indicators share several of the same components, as discussed earlier. The *anemia* variable is included to control for other measures of health and nutritional status. While the coefficient on anemia is positive and highly significant in the OLS model, once the endogeneity of *anemia* is addressed with instruments, its coefficient is no longer significant. As mentioned above, applying different weights to the factor scores and substituting the resulting values of *recognition* does not change the results significantly.

The significance of the coefficient on *recognition* suggests that human recognition is a determinant of BMI in this population, even after controlling for this other measure of nutritional and health status. Further analysis can help draw stronger conclusions about the robustness of this result and the pathways through which it operates.

## VI. Empirical Application with Kenya DHS Data

### Data and Variables

The second application of the index uses data from the Kenya Demographic and Health Survey (DHS) 2003. The Kenya DHS 2003 was conducted by the Kenya Central Bureau of Statistics in collaboration with the Kenya Ministry of Health with technical support from ORC Macro and funding from USAID and other donors. The survey collected demographic, health, nutrition, fertility, and socio-economic data in 2003 from

8,195 women aged 15-49 in 8,561 households from all eight provinces of Kenya. Data were also collected about domestic violence, women's decision-making status in the household, female genital cutting, and women's attitudes about domestic violence.

The sample used in this analysis is smaller than the 8,195 women surveyed because data for all of the variables used are not available for all women surveyed. In particular, 4,312 women were randomly selected from the total sample to answer questions related to domestic violence. Since BMI is used as the dependent variable, women for whom BMI was not recorded or for whom unrealistically high values (>50 $kg/m^2$) are recorded were not included in the analysis. The lowest recorded BMI was 12.44 $kg/m^2$, which is not unrealistically low in this population so no observations were dropped due to low BMI. Also, since BMI of pregnant and non-pregnant women cannot be meaningfully compared, pregnant women are not included in the analysis. The same implications of only including non-pregnant women in the sample that were discussed above for the India data also apply here.

Since whether one's daughter underwent female genital cutting is one of the variables used, only women with daughters who answered this question (4,594 women) were included. The final sample is 2,556 women. (For the model that includes antenatal care as a control variable, only women who responded to questions about antenatal care were included, and the sample for this model is 1,811 women).

A comparison with the full survey sample suggests that the sample of women used in this paper is somewhat older than the full sample (mean = 32.2 years, sd = 7.9, compared to mean = 28.1 years, sd = 9.3 for the full sample). This appears to be due to the exclusion of pregnant women and women without daughters.

From these data, variables were identified that measure the human recognition women receive in the household. This dataset does not have sufficient data on human recognition received in the community or institutions, and the analysis demonstrates how data from just one domain can be used. Focusing on human recognition within individual domains instead of the full index can be valuable in its own right, for example to understand how intra-household interactions affect health outcomes, a relationship this paper examines. Similarly, program managers may want to understand how program interventions affect human recognition in the particular domain they are working in, e.g. in households, the community, or in institutions such as schools. The household is hypothesized to be a primary source of recognition for women in developing countries, and given the critical decisions and processes that occur at the household level (e.g. control over resources, fertility, and food allocation), it is hypothesized that human recognition in the household domain is relevant to other material outcomes such as health.

Table 9 lists the variables used and summarizes the status of each variable among respondents in the sample. Data for all variables were collected from the women themselves. Units for all of the variables are such that higher values signify lower levels of human recognition received by the respondent. In the regression analysis, the inverse of the factor score is used so that higher values signify higher levels of recognition.

The rationale for the variables related to physical, emotional, and sexual violence, permissibility of violence, and autonomy are the same as described above for the India NFHS-3 data.

**Table 9: Kenya DHS 2003: Variables Measuring Human Recognition Women Receive in the Household**

| Domain | Variables | Respondent Status |
|---|---|---|
| *Household* | Ever experienced physical violence by husband | Yes: 42% |
| | Ever experienced physical violence by other household members | Yes: 21% |
| | Ever humiliated in public or threatened by husband | Yes: 26% |
| | Ever forced to have sexual relations with husband against will | Yes: 15% |
| | Number of reasons (out of 5) the woman believes a husband is justified for beating his wife[13] | At least 1 reason: 72%<br>All reasons: 11% |
| | Daughter underwent female genital cutting (FGC) | Yes: 11% |
| | Who makes decisions about the woman visiting her family | Entirely made by others: 39% |
| | Who makes decisions about the woman's health care | Entirely made by others: 45% |

The variable on daughter's female genital cutting, *FGC*, is an example of a variable reflecting human recognition that is heavily influenced by cultural factors and occurs in the household domain, as the decision is generally made by household members. Data are also available on whether the respondent herself underwent FGC, but since that decision occurred in a different household (the respondent's parental household, not the respondent's conjugal household), the daughter's FGC status is used here measure recognition received in the household and to be consistent with the other measures used. The decision to practice FGC on a daughter reflects the level of recognition females in the household receive, in particular the relative value placed on the daughter's pain, health risks, psychological effects, loss of sexual feeling, and other negative impacts the practice causes, compared to the relative value placed on cultural beliefs or pressures to perform the practice. (See El-Defrawi et al. 2001 and WHO 2006

[13] Reasons: going out without telling husband; neglecting children; arguing with husband; refusing sex; burning the food.

on the impacts of FGC.) It is assumed that this recognition in part reflects the level of recognition females in the household receive, including the respondent. Respondents' FGC and their daughters' are positively correlated (corr = .41). When the respondent's own FGC is used instead of the daughter's, the factor loading on *FGC* is lower, but the sign and significance level of subsequent regression results using the factor do not differ significantly (results not shown). The lower factor loading may be because one's own FGC reflects recognition in the respondent's parental household rather than recognition in her conjugal household as the other variables in the factor analysis do.

The basic model is:

$$BMI_i = \alpha + \beta_1 income_i + \beta_2 education_i + \beta_3 floor_i + \beta_4 age_i + \gamma hhrecognition_i + e_i.$$

The specification differs somewhat from the specification in the India application because different data are available. The *hhrcecognition* variable is the level of human recognition received in the household, using factor analysis scores for the indicators in Table 9. As discussed above, recognition in other domains are not included because sufficient data for these domains are not available in the Kenya DHS 2003. As with the India data, since the data are cross-sectional, the model tests whether individuals with higher levels of human recognition have higher BMIs, controlling for other factors, not whether a change in recognition received by a given individual is associated with increased BMI.

The dependent variable for the multivariate analysis is the woman's BMI, as it was in the India analysis. In the sample, the mean BMI is 23.01, with a standard deviation of 4.47, and 11.8% of the women have a BMI < 18.5 kg/$m^2$, which the WHO classifies as malnourished.

In order to control for other factors affecting nutritional status, the model includes data on socio-economic factors. Unlike the India data, the Kenya DHS do not include a wealth index or standard of living index but do include income quintile data. *Income* measures the income quintile of the respondent (mean 3.05, sd 1.47). *Education* measures the respondent's educational attainment. In the sample 19.8% have no education, 54.9% have primary education, 20.1% have secondary education, and 5.2% have higher education. *Age* measures the age in years of the respondent (mean 32.2, sd 7.89). *Floor* measures the main floor materials in the respondent's house and classifies it such that higher values represent more expensive and durable materials. *Floor* is used as a measure of housing quality, which is one component of standard of living. Results do not differ significantly when housing quality is measured using a dummy variable for electricity instead of the *floor* variable. The Kenya DHS 2003 does not include data on anemia status of women so it is not possible to include an anemia variable as is done in the India application.

In order to test the extent to which human recognition's effect on nutritional status occurs through access to health care, a dummy variable for whether the respondent received antenatal care is added to the model. In Kenya and elsewhere it is recommended that all pregnant women undergo periodic antenatal care checkups. *ANC* equals 1 if the client received at least one antenatal checkup during her first pregnancy and 0 otherwise. In the sample 87.6% of women received at least one checkup and 12.4% did not. Data are not available on this indicator for many respondents so the model that includes *ANC* has a smaller sample size.

The *ANC* variable reflects access to health care services but it may also reflect the health status of the respondent, with healthier women less likely to have checkups (though it is recommended that all pregnant women have them). These two factors are expected to have opposing effects on nutritional status with greater access to health care (higher ANC) contributing to higher BMI and poorer health (higher ANC) contributing to lower BMI. The significant positive coefficient on the *ANC* variable in the BMI specification suggests that the access component is dominant. One way greater levels of human recognition in the household may lead to improved health status is through enabling greater access to health care services. The variable *ANC* is used to check to what extent this is the case by controlling for access to services.

To address possible endogeneity of *hhrecognition* due to omitted variable bias and/or simultaneity, three variables related to culture and religion are used as instruments. *Ethnicity* is an indicator of the respondent's ethnicity or tribe. *Religion* is an indicator of the respondent's religion and equals 1 if Christian and 2 if Muslim. In the sample, 86% of women are Christian and 14% are Muslim. *FGCarea* is a dummy variable that equals 1 if the respondent lives in a community where FGC is practiced and equals 0 otherwise. In the sample 38.5% of respondents live in an area where FGC is practiced. The rationale for using these variables as instruments is discussed along with sources of endogeneity in the section below on regression results.

All estimates are generated using Stata 10.

## Factor Analysis Results

The factor analysis process used is similar to the India application. However, some additional steps are taken to demonstrate the need and role for factor analysis and to understand the nature of the factor other than human recognition that emerges from the factor analysis. As a preliminary step to demonstrate the need for factor analysis, the model is first estimated using the eight individual variables from Table 8. The model is specified as follows:

$$BMI_i = \alpha + \beta_1 income_i + \beta_2 educ_i + \beta_3 floor_i + \beta_4 age_i + \gamma_1 physviol_i + \gamma_2 othviol_i + \gamma_3 emotabuse_i + \gamma_4 forcedsex_i + \gamma_5 violOK_i + \gamma_6 FGC_i + \gamma_7 decisionfamily_i + \gamma_{8i} decisionhealth_i + e_i$$

Estimation results for this regression are not robust (results not shown). Two specific problems are discussed here. First, a Hausman-Wu test indicates that some of the explanatory variables are endogenous. But since a number of the variables representing human recognition are closely related and correlated with each other, it is difficult to determine which of the variables are endogenous and which are not. Using *ethnicity*, *FGCarea*, and *religion* as external instruments, estimation results will differ depending on which explanatory variables are treated as endogenous. If it is assumed that all eight measures of recognition are endogenous, there are not enough valid external instruments available in the data to estimate coefficients for all the variables in the same model.

Secondly, there is fairly high correlation among some of the recognition variables, and including so many closely related variables in the regression reduces the significance of any particular variable and clouds the interpretation of human recognition's effect on the outcome variable. If what is common among these variables is the level of human recognition received in the household, then when all of the variables are included in the regression, this common factor (or the variation in the common factor across observations) may be controlled for by the other variables, and the significance of individual coefficients reflects the effects of *other* aspects of these variables. Yet what we are trying to measure is the effect of the common factor – human recognition – on the outcome variable.

What is needed is a method to draw out the common factor of human recognition from the eight variables and regress the outcome variable on the common factor. This is what factor analysis does.

Exploratory factor analysis is used to identify how many factors explain the eight measures and to obtain a measure of the household human recognition factor. Because factor analysis is based on identifying common variation among the measured variables, an inherent assumption in using it here is that there is common variation among the different variables across observations. This seems to be a reasonable assumption, and the statistical correlations among the variables confirm this.

The factor analysis is based on the following model:

$$\begin{aligned}
\textit{physical violence}_{\mathrm{i}} &= \lambda_{1p}\textit{hhrecognition}_{\mathrm{i}} + \lambda_{2p}\mu_{\mathrm{i}} + \delta_{ip} \\
\textit{others' violence}_{\mathrm{i}} &= \lambda_{1o}\textit{hhrecognition}_{\mathrm{i}} + \lambda_{2o}\mu_{\mathrm{i}} + \delta_{io} \\
\textit{emotional abuse}_{\mathrm{i}} &= \lambda_{1e}\textit{hhrecognition}_{\mathrm{i}} + \lambda_{2e}\mu_{\mathrm{i}} + \delta_{ie} \\
\textit{forced sex}_{\mathrm{i}} &= \lambda_{1s}\textit{hhrecognition}_{\mathrm{i}} + \lambda_{2s}\mu_{\mathrm{i}} + \delta_{is} \\
\textit{violence OK}_{\mathrm{i}} &= \lambda_{1v}\textit{hhrecognition}_{\mathrm{i}} + \lambda_{2v}\mu_{\mathrm{i}} + \delta_{iv} \\
\textit{FGC}_{\mathrm{i}} &= \lambda_{1f}\textit{hhrecognition}_{\mathrm{i}} + \lambda_{2f}\mu_{\mathrm{i}} + \delta_{if} \\
\textit{decision family}_{\mathrm{i}} &= \lambda_{1d}\textit{hhrecognition}_{\mathrm{i}} + \lambda_{2d}\mu_{\mathrm{i}} + \delta_{id} \\
\textit{decision health}_{\mathrm{i}} &= \lambda_{1h}\textit{hhrecognition}_{\mathrm{i}} + \lambda_{2h}\mu_{\mathrm{i}} + \delta_{ih} \qquad \mathrm{i} = 1\ldots2{,}556
\end{aligned}$$

The $\textit{hhrecognition}_{\mathrm{i}}$ variable is the latent variable of the human recognition that individual i receives in the household, the $\lambda$'s are the factor loadings, $\mu_{\mathrm{i}}$ is another latent variable (i.e. another factor) underlying the eight measures, and the $\delta_{ix}$'s are unique factors that affect the individual measures. The letter subscripts (p…h) refer to the eight measured variables. Exploratory factor analysis assumes the unique factors, $\delta_{ip}\ldots\delta_{ih}$, are uncorrelated. This appears to be a reasonable assumption for these data for the same reasons discussed above for the India data.

Two common factors (*hhrecognition*$_i$ and $\mu_i$) are included in this model because applying the conventional cutoff of Eigenvalue $\geq$ 1 as the threshold for keeping a factor, the factor analysis results indicate there are two significant factors (Table 10). Iterated principal factor analysis is used, and the factor loadings (estimates of the $\lambda$ parameters) are re-estimated using only these two factors to obtain more accurate estimates of the factor loadings. Table 10 presents the results.

**Table 10: Factor Analysis Results: Kenya DHS**

| Variable | Factor 1 Loadings | Factor 2 Loadings | Uniqueness |
|---|---|---|---|
| physical violence | 0.713 | -0.126 | 0.475 |
| others' violence | 0.055 | -0.008 | 0.990 |
| emotional abuse | 0.619 | -0.150 | 0.595 |
| forced sex | 0.446 | -0.147 | 0.779 |
| violence OK | 0.192 | 0.160 | 0.937 |
| FGC | 0.098 | 0.044 | 0.988 |
| decision family | 0.141 | 0.625 | 0.590 |
| decision health | 0.183 | 0.681 | 0.502 |
| *Eigenvalue* | *1.194* | *0.942* | |

The results are consistent with interpretation of Factor 1 to be human recognition received in the household. All factor loadings for Factor 1 are positive, which is expected since all of the measured variables have higher values for receipt of lower levels of recognition. The common factor is interpreted to be low levels of human recognition.

The uniquenesses are quite high for some of the variables, in particular *FGC*, *others' violence*, and *violence OK*. This suggests that human recognition in the household and Factor 2 do not explain variation in these variables very well. While this may indicate that these measures do not strongly reflect human recognition, there may be other explanations as well. Communality (and uniqueness) is determined by how much a measure's variation is correlated with variation in the other measures. All three of these measures involve human recognition from other domains: *FGC* from community; *others' violence* from community and institutions (e.g. school)[14]; and *violence OK* possibly from community and from one's upbringing in one's parents' household. This interpretation of the high levels of uniqueness of these measures suggests that these measures may belong in other domains of the human recognition composite index, not in the household domain. Alternatively, it may be that these three measures are quite independent from the other measures used so variation across individuals in these measures is not correlated with variation in the other measures. Within a given domain, recognition received from different individuals may not be strongly correlated.

While it is not immediately clear what Factor 2 refers to, the factor loadings suggest this factor may be related to age and experience. Older women[15] are more likely to have experienced the various forms of violence and abuse measured, if for no other reason than because they have had more years of life and marriage for it to occur[16]. The negative factor loadings on all the violence and abuse measures are consistent with an interpretation of Factor 2 to be youth. Furthermore, older women are more likely to participate in decisions related to their health and visits to family; in sub-Saharan Africa

---

[14] 94% of women in the sample who experienced violence from individuals other than their husbands experienced it from individuals outside the households they and their husbands inhabit.

[15] Note that women in the sample range from age 15-49 so "older women" refers to the higher side of this range. Also, note that since all women in the sample had daughters to have values for *FGC*, never married women (who may have more freedom) are not included in the sample.

[16] Another reason older women may be more likely to have had these experiences is because views and treatment of women in Kenya have improved over the past two decades so younger women are less likely to have had these experiences. However, data on most of these variables have only recently begun to be collected so empirical evidence supporting this interpretation is difficult to find.

young women often have less control over household decisions than older women do, especially decisions related to traveling out of the community. The positive factor loadings on these measures are consistent with this interpretation, including less decision-making power for younger women. This interpretation is also consistent with the significant negative correlations between age and the two decision variables: -0.41 and -0.33.

It is not intuitively clear whether younger or older women are more likely to think being beaten by one's husband is justified, but applying this interpretation of the factor suggests that younger women are more likely to believe this. The positive factor loading on *FGC* is the only one not consistent with this interpretation since one would expect the daughters of older women to be more likely to have undergone FGC since FGC incidence has been declining over time in recent years; data indicate that in Kenya FGC has decreased by nearly 50% over the past two decades (UNICEF 2005). However, the factor loading on *FGC* for Factor 2 does have a small magnitude (.044) and there is a very high uniqueness (0.988), indicating that the factors do not explain the variation in *FGC* very well.

To check this interpretation of Factor 2, age is regressed on the eight measures using OLS. Results are given in Table 11. For measures with factor loadings greater than 0.1, all of the coefficient signs are consistent with the factor loadings. Coefficients on *others' violence* and *FGC* are not consistent with the factor loadings, but both of these have very small factor loadings and high uniquenesses in the factor analysis.

**Table 11: Regressing Age on Eight Human Recognition Variables**

| Dependent Variable | Estimation Method | Intercept | Physical violence | Others' violence | Emotional abuse | Forced Sex |
|---|---|---|---|---|---|---|
| Age | OLS | **33.67** (0.285) | -0.020 (0.338) | **-1.97** (0.354) | **1.37** (0.372) | 0.375 (0.431) |
| | | | | | | |
| | | Violence OK | FGC | Decision family | Decision health | n, $R^2$, F |
| | | **-2.18** (0.084) | **5.89** (0.456) | **-2.44** (0.332) | **-1.53** (0.326) | n=2,686 $R^2$=0.12 F=43.99 |

Parameter estimates in bold indicate significance at the .05 level.

The purpose of factor analysis is to identify common factors that are latent, unobservable variables; if a variable is observable, it can be used directly in empirical estimation. The fact that data on age are available for this sample suggests that factor analysis is clearly unnecessary for age. The above exercise is presented only to more fully interpret the factor analysis results. The main objective of the factor analysis here is to measure human recognition in the household, and the results support interpretation of Factor 1 to be human recognition.

The factor loadings are combined with the values for the eight measures to produce a factor score for human recognition for each observation. This factor score, normalized to have approximately mean 0 and standard deviation 1, can be used in empirical estimation as a measure of human recognition received in the household. Summary statistics for the factor score *hhrecognition* are: mean = 3.50e-09, sd = 0.824.

Regression Results

The factor scores are used to estimate the following model:

$$BMI_i = \alpha + \beta_1 income_i + \beta_2 education_i + \beta_3 floor_i + \beta_4 age_i + \gamma hhrecognition_i + e_i$$

When estimated with OLS, the specification is significant and the coefficient on *hhrecognition* is positive and significant. Results are reported in Table 12.

**Table 12: Results of Nutritional Status Model using Kenya DHS Data: OLS**

| Dependent Variable | Estimation Method | Intercept | Income | Education | Floor |
|---|---|---|---|---|---|
| BMI | OLS (heterosk.-consistent) | **15.32**<br>**(0.350)** | **0.874**<br>**(0.083)** | **0.892**<br>**(0.127)** | **0.295**<br>**(0.131)** |
| | | | | | |
| | | Age | Household recognition | | n, $R^2$, F |
| | | **0.110**<br>**(0.010)** | **0.179**<br>**(0.092)** | | n=2,628<br>$R^2$=0.23<br>Pr>F < .0001 |

Parameter estimates in bold indicate significance at the .05 level. Standard errors are given in parentheses.

However, a Hausman specification test indicates that *hhrecognition* is endogenous ($Pr > \chi^2 = .058$). This is consistent with the earlier finding that some of the individual variables are endogenous in this specification. A Hausman specification test does not reject exogeneity of *income* ($Pr > \chi^2 = .93$) and it is treated as exogenous in the estimation.

Two possible sources of endogeneity – simultaneity and omitted variable bias – are at play. Measurement error is less likely to be a problem than in the India application because weights are not assigned since only one domain is measured. To address the endogeneity, 2SLS is used to re-estimate the specification with *ethnicity*, *religion*, and *FGCarea* as external instruments. Intuitively, it is expected that these three variables are correlated with the endogenous variable, the factor score for household human recognition. *FGCarea* and *ethnicity* are both highly correlated with whether one's daughter experienced FGC. Ethnicity, religion and degree of traditionalism (which may be captured in *FGCarea*) are also likely to affect other variables included in the human recognition measure, such as women's control over decisions and possibly permissibility and pervasiveness of domestic violence.

The Anderson correlation likelihood ratio statistic for the instruments is highly significant ($\chi^2$ $p < .0001$), indicating that these instruments are correlated with the endogenous variable.

Further evidence that these three variables are correlated with the factor score is found in the highly significant coefficients on all three variables when the factor score is regressed on these variables, the control variables, and other determinants of the human recognition measures. Results of this regression are given in Table 13. The signs of all the explanatory variables are as expected, except possibly *age*. The overall regression is significant, and all coefficients are significant, though the $R^2$ is low.

**Table 13: Determinants of Human Recognition in the Household and Relevance of Instruments**

| Dependent Variable | Estimation Method | Intercept | Income | Education | Floor | Age | Ethnicity | Religion |
|---|---|---|---|---|---|---|---|---|
| hhrecognition | OLS (heterosk.-consistent) | **-1.082**<br>(0.120) | 0.023<br>(0.017) | **0.061**<br>(0.027) | **0.087**<br>(0.025) | **-0.005**<br>(0.002) | **0.174**<br>(0.029) | **0.119**<br>(0.052) |
| | | | | | | | | |
| | | FGC in community | Age at first marriage | | Age at first intercourse | | n, $R^2$, F | |
| | | **-0.242**<br>(0.042) | **0.26**<br>(0.004) | | **0.003**<br>(0.0005) | | n=2,556<br>$R^2$=0.08<br>Pr>F < .0001 | |

Parameter estimates in bold indicate significance at the .05 level. Standard errors are given in parentheses.

In order to be valid instruments, the three variables also need to be uncorrelated with the errors in the BMI specification. There is no simultaneity bias between any of these variables and BMI. Since they are fixed variables for a given individual, BMI cannot be a determinant of these variables; if one's BMI changes, it will not affect one's ethnicity, religion, or residence location. Omitted variables such as substance abuse, HIV infection, or even a family member's or one's own attitudes may affect BMI and may affect human recognition levels, but are very unlikely to affect religion or ethnicity. There could be some effect on residence location but these factors are not expected to significantly affect whether one resides in a location where FGC is commonly practiced.

However, the causation may work in the opposite direction, and an individual's culture or location may affect BMI through pathways other than human recognition. If some of these pathways involve variables that are not included in the model and are therefore part of the error term, this would reduce the validity of the instruments. For example, ethnicity, religion, or residence location may affect dietary practices, which in turn could affect BMI.

The Hansen J test statistic for these instruments is insignificant ($\chi^2$ p = .733), providing evidence that these three variables are not correlated with the error terms and are therefore valid instruments.

As with the India data, given that these are cross-sectional data, instrumental variables appear to be the best available option for dealing with endogeneity, and these variables seem to be the best instruments available in the dataset. Therefore, for this application of the index, the model is estimated with 2SLS using these three external instruments, with the caveats about the instruments discussed above. Endogeneity can be better addressed with panel data (Castleman 2011c).

Estimation results from the above model using two-stage least squares with these three instruments are given in Table 14. Results indicate that higher levels of recognition are associated with higher BMI (better nutritional status)[17] among Kenyan women, as are higher levels of income, education, and greater age.

[17] In developed country settings, a higher BMI is associated with being *less* healthy due to overweight and obesity. This is not the case in the sample, as only 205 (8%) of women in the sample have BMI $\geq$ 30, the cutoff for obesity, and results do not change significantly when these women are dropped from the analysis.

**Table 14: Results of Nutritional Status Model using Kenya DHS Data: 2SLS**

| Dependent Variable | Estimation Method | Intercept | Income | Education | Floor |
|---|---|---|---|---|---|
| BMI | 2SLS (heterosk.-consistent) | **15.54**<br>**(0.389)** | **0.871**<br>**(0.089)** | **0.777**<br>**(0.135)** | 0.123<br>(0.155) |
| | | Age | Household recognition | | n, $R^2$, F |
| | | **0.117**<br>**(0.010)** | **1.65**<br>**(0.572)** | | n=2,556<br>$R^2$=0.16<br>$Pr>F < .0001$ |
| Endogenous variable: Hhrecognition<br>External instruments: Ethnicity, Religion, FGC practiced in community | | | | | |

Parameter estimates in bold indicate significance at the .05 level. Standard errors are given in parentheses.

An F test indicates that the specification has significant explanatory power at the .01 level. Coefficients on all the control explanatory variables are significant, except for *floor*. Coefficient signs are as expected: higher income, greater education, and greater age are associated with higher BMI. The coefficient on *hhrecognition* is positive and significant at the .01 level (p = .004), indicating that receipt of higher levels of human recognition in the household is associated with higher BMI. The standard deviation of household recognition is 0.82 so controlling for the other variables, women with a 1 standard deviation higher level of household recognition have, on average, a 1.35 kg/m$^2$ higher BMI. With average respondent height of 1.59 meters, this corresponds to weighing 3.4 kg more.

These results offer evidence that the human recognition a woman receives in the household domain is a significant determinant of her nutritional status, even after controlling for age and socio-economic characteristics. There are a number of possible pathways through which this relationship between human recognition in the household and nutritional status may occur. For example, the level of human recognition received may be associated with food intake, especially if there are cultural norms for women to eat last and least or to reduce intake by more than other household members during periods of food shortages.

Another possible pathway – one that can be tested with these data – is that women who receive lower levels of human recognition avail health services less than those receiving higher levels of recognition. This may be due to restrictions on a woman's autonomy and travel, her weakened capacity to demand care, or general disregard for a woman's well-being by her husband or other household members controlling resources or decision-making. Amartya Sen has examined how women's lack of capabilities and freedom lead to diminished access to vital services such as health care (Sen 1999). Women less able to avail health services are likely to receive less information about good nutrition and/or to less quickly treat infections and other health problems, all of which may lead to worse nutritional status.

The extent to which the effect of human recognition on nutritional status occurs through access to health services is tested by including a measure of access[18] to health care – receipt of antenatal care during pregnancy – in the BMI specification. The 2SLS specification is modified to include a dummy variable *ANC*. Estimation results are given in Table 15.

[18] I use the term "access" here in a broad sense. Women may have physical access to health care but not avail it due to household constraints, which is interpreted as lacking access in the way the term is used here.

**Table 15: BMI Regression using Human Recognition Factor Score and ANC**

| Dependent Variable | Estimation Method | Intercept | Income | Education | Floor |
|---|---|---|---|---|---|
| BMI | 2SLS (heterosk.-consistent) | **15.46**<br>**(0.493)** | 0.655<br>(0.093) | **0.898**<br>**(0.153)** | 0.268<br>(0.174) |
| | | Age | ANC | Household recognition | n, $R^2$, F |
| | | **0.105**<br>**(0.013)** | **0.609**<br>**(0.236)** | ***0.934****<br>***(0.525)*** | n=1,811<br>$R^2$=0.18<br>Pr>F < .0001 |
| Endogenous variable: Household recognition<br>External instruments: Ethnicity, Religion, FGC practiced in community | | | | | |

Parameter estimates in bold indicate significance at the .05 level and estimates in bold italics indicate significance at the .1 level. *p=0.075

The signs and significance of coefficients on the control variables are unchanged, and the coefficient on *ANC* is positive and significant (p=.01), indicating that women who availed ANC have higher BMI and are healthier. As discussed earlier, the positive sign and significance of the coefficient on *ANC* indicates that the part of *ANC* explained by access to health care dominates the part of *ANC* explained by being less healthy and therefore in greater need of antenatal care. The sign on *hhrecognition* is still positive but its significance has decreased from p = .004 to p = .075. Having a 1 standard deviation higher level of household recognition now corresponds to a 1.9 kg. greater weight. When a variable that reflects access to health care services is controlled for, the significance of human recognition's effect on nutritional status decreases. This suggests that one pathway by which human recognition in the household affects nutritional status is through access to health care, with higher levels of recognition leading to greater access to health care which in turn supports better nutritional status.

The fact that the coefficient on human recognition is still positive and marginally significant suggests that in addition to this effect through access to health care, recognition also affects nutritional status through other pathways. Possible other pathways may include human recognition's psychic effects on health, food consumption, and habits; recognition's effects on vulnerability to HIV based on evidence linking violence against women and vulnerability to HIV (Dunkle et al. 2004), which often damages nutritional status; and recognition's effects on household practices affecting nutritional status such as daily workload or hygiene and sanitation. Further empirical work is needed to be able to draw stronger conclusions about human recognition's effect on health and nutritional status through access to services and other pathways.

## VII. Discussion and Areas for Further Research

This paper developed and applied a methodology for measuring the level of human recognition that individuals receive. Based on a framework that organizes the sources of human recognition into the various domains of an individual's life, an index was developed that measures the human recognition received in each of the different domains and combines the domain-specific measures into a single overall measure. The index was applied to two different datasets to demonstrate how it can be used in different settings for measuring human recognition levels and estimating empirical models of the role recognition plays in other development outcomes.

This work is intended to lay the groundwork for further empirical study of human recognition and to support efforts to measure other intangible dimensions of development that could apply similar frameworks and indices. In addition to supporting research, the framework and index can also be used by programs to measure and monitor human recognition levels among targeted populations. Such applications can assist programs in understanding recognition status and needs to inform program design and track progress and results related to human recognition and its effects on other outcomes. While identifying accurate domain weights and addressing endogeneity in regression models pose challenges for empirical research, these issues may not be of as great concern to programs. By applying the index with available data or adding a few indicators to existing monitoring and evaluation systems, programs can measure human recognition as they measure other characteristics of targeted populations and other program impacts.

The empirical work presented here demonstrates how exploratory factor analysis can be used to measure human recognition from specific indicators in different domains of an individual's life. Factor analysis has been used by others to measure related concepts such as women's empowerment and is well-suited to measuring unobserved variables that occur in multiple areas of one's life. The results of the factor analysis were consistent with predictions about the relationship between the observed measures and human recognition. By producing a single measure for human recognition received in a given domain, factor analysis avoids methodological problems associated with using several individual variables for human recognition. One limitation of using factor analysis is that it requires interpretation of the factors, as there is no way to empirically test what the factor signifies. In these applications factor analysis results are consistent with interpreting the primary factors to be human recognition, but no empirical test exists that can provide statistical evidence about such an interpretation. Stronger assumptions are required for this interpretation for the survey data used in this paper than for data specially designed to measure human recognition.

It may be valuable to validate the measurement approach by comparing index measures to direct and objective measures of human recognition, if data on such measures can be obtained. However, finding standardized objective measures is challenging for latent, intangible variables such as human recognition. Possibly data on self-reported provision and receipt of recognition and data collected by trained specialists observing interactions could be used to validate an index of indicators.

Further work is also needed on identifying the domain weights in the index. While the index addresses major identified challenges to measuring human recognition, without a sound method of assigning weights, combining domain-specific measures into a single measure is subject to a certain degree of arbitrariness. This limitation was addressed in the empirical work by using various combinations of weights to check the robustness of results.

The two applications of the measurement method used in this paper demonstrate how human recognition can be measured for different populations using cross-sectional survey data. The measurement method can also be applied using panel data (Castleman 2011c). The applications in this paper also showed how the measurement method can be applied using one or multiple domains. Application of the same index in multiple countries raises the possibility of comparing human recognition status across countries. But caution is needed in such exercises because using different indicators to form the index in different countries limits the validity of such comparisons. Cross-country comparisons of the direction and rate of change in recognition using the index may be more defensible.

Results from the empirical applications indicate that human recognition plays a significant role in the nutritional status of women, offering evidence to support the hypothesis that human recognition is a significant determinant of health outcomes. Women in these populations who receive higher levels of human recognition have higher BMIs. If additional empirical work further supports this finding, then given the significant role nutritional status plays in health, mortality, and productivity, human recognition issues may need to be explicitly considered in the design and implementation of development activities, especially those aiming to improve health outcomes. A possible starting point would be to assess which program interventions and implementation approaches are most effective at improving human recognition and which, if any, worsen it. Another paper examines the human recognition impact of two particular interventions: provision of supplementary food to malnourished adult HIV patients and medical treatment of HIV (Castleman 2011c).

Results from the Kenya DHS application offer some initial evidence that one mechanism through which human recognition affects health status is access to health care services. Further study is needed, but if this finding is confirmed, it may have implications for the design of interventions aimed at improving women's access to health care in Kenya and other developing countries.

The primary objective of the empirical applications presented here was to demonstrate how the index for measuring human recognition can be applied with cross-sectional data. Greater understanding of the determinants and effects of human recognition requires more in-depth empirical study. Nevertheless, the findings from this paper suggest that human recognition is a significant determinant of nutritional status in the populations studied, and that despite a number of challenges, measurement of human recognition is feasible.

**References**

Baliamoune-Lutz, Mina and Stefan H. Lutz. "The Contribution of Income, Social Capital, and Institutions to Human Well-Being in Africa". ZEI Working Paper B07. Center for European Integration Stuides, Bonn, Germany, February 15, 2004.

Black, Robert E, Lindsay H Allen, Zulfiqar A Bhutta, Laura E. Caulfield, Mercedes de onis, Majid Ezzati, Colin Mathers, Juan Rivera. "Maternal and Child Undernutrition: Global and Regional Exposures and Health Consequences." *The Lancet*. January 17, 2008.

Bollen, Kenneth A. *Structural Equations with Latent Variables*. New York: John Wiley & Sons, Inc., 1989.

Castleman, Tony. "Human Recognition and Its Role in Economic Development: A Descriptive Review." 2011. The George Washington University. Institute for International Economic Policy Working Paper 2011-08.

Castleman, Tony. "Human Recognition in Economic Development: A Theoretical Model." 2011a. The George Washington University. Institute for International Economic Policy Working Paper 2011-09.

Castleman, Tony. "Human Recognition among HIV-Infected Adults: Empirical Evidence from a Randomized Controlled Trial in Kenya." 2011c. The George Washington University. Institute for International Economic Policy Working Paper 2011-11.

Central Bureau of Statistics (CBS) Kenya, Ministry of Health (MOH) Kenya, and ORC Macro. *Kenya Demographic and Health Survey 2003*. Calverton, Maryland: CBS, MOH, and ORC Macro, 2004.

Diener Ed and Robert Biswas-Diener. "Psychological Empowerment and Subjective Well-Being." *Measuring Empowerment: Cross-Disciplinary Perspectives*. Deepa Narayan, ed. Washington DC: The World Bank, 2005.

Dunkle, Kristin L., Rachel K. Jewkes, Heather C. Brown, Gledna E. Gray, James A. McIntyre, and Sioban D. Harlow. "Gender-Based Violence, Relationship Power, and Risk of HIV Infection in Women Attending Antenatal Clinics in South Africa". *The Lancet*. Vol. 363: May 1, 2004.

Durlauf, Steven N. "On the Empirics of Social Capital." Mimeo, University of Wisconsin, April 2001.

El-Defrawi, Mohammed H., Galal Lotfy, Khadiga Dandash, Amany H. Refaat, and Moustafa Eyada. "Female Genital Mutilation and its Psychosexual Impact". *Journal of Sex and Marital Therapy*. Volume 27, Number 5: 465-473. October 2001.

Erulkar, Annabel S. "The Experience of Sexual Coercion among Young People in Kenya". *International Family Planning Perspectives*. Vol. 30, No. 4: December 2004.

Ezzati, Majid, Alan D. Lopez, Anthony Rodgers, Stephen Vander Hoorn, Christopher J. L. Murray, and the Comparative Risk Assessment Collaborating Group. "Selected Major Risk Factors and Global and Regional Burden of Disease." *The Lancet*. October 30, 2002.

Fukuyama, F. *The Great Disruption*. New York: Simon and Schuster, 1999.

Green, Francis and Thomas E. Weisskopf. "The Worker Discipline Effect: A Disaggregative Analysis". *The Review of Economics and Statistics*. 1990. Vol. 72, No. 2: 241-249.

Grootaert, Christian. "Social Capital: The Missing Link?" *Social Capital and Participation in Everyday Life*. Dekker, P. and E. Uslaner, eds. Routledge, 2001.

International Institute for Population Sciences (IIPS) and Macro International. 2007. *National Family Health Survey (NFHS-3), 2005-06: India: Volume I and Volume II.* Mumbai: IIPS.

Kabeer, Naila. "Reflections on the Measurement of Women's Empowerment". *Discussing Women's Empowerment – Theory and Practice*. Sida Studies No. 3. Novum Grafiska AB: Stockholm, 2001.

Khwaja, Asim Ijaz. "Measuring Empowerment at the Community Level: An Economist's Perspective." *Measuring Empowerment: Cross-Disciplinary Perspectives*. Deepa Narayan, ed. Washington DC: The World Bank, 2005.

Kishor, Sunita. "Empowerment of Women in Egypt and Links to the Survival and Health of their Infants". *Women's Empowerment and Demographic Processes: Moving Beyond Cairo*. Harriet Presser and Gita Sen, eds. New York: Oxford University Press, 2000.

Knack, S. and P. Keefer. "Does Social Capital Have an Economic Impact? A Cross-Country Investigation". *Quarterly Journal of Economics*. 1997. Vol. 112, No. 4: 1252-1288.

Lane, Robert. *The Market Experience*. Cambridge: Cambridge University Press, 1991.

Loehlin, John. *Latent Variable Models: An introduction to factor, path, and structural analysis.* New Jersey: Lawrence Erlbaum Associates, 1998.

Malhotra, Anju and Sidney Ruth Schuler. "Women's Empowerment as a Variable in International Development." *Measuring Empowerment: Cross-Disciplinary Perspectives*. Deepa Narayan, ed. Washington DC: The World Bank, 2005.

Malhotrra, Anju, Sidney Ruth Schuler, and Carol Boender. "Measuring Women's Empowerment as a Variable in International Development". Commissioned by the World Bank, 2002.

Martin, Sandra L., Kathryn E. Moracco, Julian Garro, Amy Ong Tsui, Lawrence L. Kupper, Jennifer L. Chase, and Jacquelyn C. Campbell. "Domestic Violence Across Generations: Finding from Northern India". *International Journal of Epidemiology*. 2002. 31: 560-572.

Mason, Karen Oppenheim. "Measuring Women's Empowerment: Learning from Cross-National Research." *Measuring Empowerment: Cross-Disciplinary Perspectives*. Deepa Narayan, ed. Washington DC: The World Bank, 2005.

Mendelsohn, Oliver and Marika Vicziany. *The Untouchables: Subordination, Poverty and the State in Modern India.* New York: Cambridge University Press, 1998.

Moffit, Robert A. "Policy Interventions, Low-Level Equilibria and Social Interactions". *Social Dynamics*. Steven N. Durlauf and H. Peyton Young, eds. Washington DC: Brookings Institute Press, 2001.

Narayan, Deepa and L. Pritchett. "Cents and Sociability: Household Income and Social Capital in Rural Tanzania." *Economic Development and Social Change*. 1999. Vol. 47, No. 4: 871-897.

Narayan, Deepa. "Conceptual Framework and Methodological Challenges." *Measuring Empowerment: Cross-Disciplinary Perspectives*. Deepa Narayan, ed. Washington DC: The World Bank, 2005.

Naved, Ruchira Tabassum and Lars Ake Persson. "Factors Associated with Spousal Physical Violence against Women in Bangladesh". *Studies in Family Planning*. Vol. 36, Issue 4. Population Council Inc., December 2005.

Odimegwu, Clifford O. "Family Planning Attitudes and Use in Nigeria: A Factor Analysis". *International Family Planning Perspectives*. 1999. Vol. 25, No. 2: 86-91.

Phipps, Barbara J. and James E. Clark. **"**Attitudes toward Economics: Uni- or Multidimensional?" *The Journal of Economic Education*. 1993. Vol. 24, No. 3: 195-212.

Pillareisetti, J. Ram and Mark McGillivray. "Human Development and Gender Empowerment: Methodological and Measurement Issues". *Development Policy Review*. Vol. 16: 1998.

Schuler, Sidney Ruth, Syed M. Hashemi, Ann P. Riley, and Shireen Akhter. "Credit Programs, Patriarchy and Men's Violence against Women in Rural Bangladesh". *Social Science & Medicine*. Vol. 43, Issue 12, December 1996.

Sen, Amartya. *Development as Freedom*. New York: Anchor Books (Random House), 1999.

Stone, Wendy. "Measuring Social Capital: Towards a Theoretically Informed Measurement Framework for Researching Family and Community Life". Australian Institute of Family Studies. Research Paper No. 24. Melbourne, February 2001.

UNDP. "Technical Note 1: Calculating the Human Development Indices. *Human Development Report 2004*. 2004.

UNICEF. "Kenya FGM/C Country Profile: Technical Note." February 2005.

Victora, Cesar G, Linda Adair, Caroline Fall, Pedro C Hallal, Reynaldo Martorell, Linda Richter, Harshpal Singh Sachdev. "Maternal and Child Undernutrition: Consequences for Adult Health and Human Capital." *The Lancet*. January 17, 2008.

WHO Study Group on Female Genital Mutilation and Obstetric Outcome. "Female Genital Mutilation and Obstetric Outcome: WHO Collaborative Prospective Study in Six African Countries". *Lancet* 2006. 367:1835–41.

Williams, Jill. "Measuring Gender and Women's Empowerment Using Confirmatory Factor Analysis". Institute of Behavior Science Working Paper. April 2005.

World Health Organization. *Management of severe malnutrition: A manual for physicians and other senior health workers*. Geneva, Switzerland: World Health Organization, 1999.

World Health Organization. *Obesity, Preventing and Managing the Global Epidemic: Report of a WHO Consultation on Obesity*. Geneva, Switzerland: World Health Organization, 1998.