The Role of Ordinal Data in Economics

by Chrysanthi Hatzimasoura

B.Sc. in Statistics and Economics, July 2007, Birkbeck College, University of London
M.A. in Economics, January 2012, The George Washington University

A Dissertation submitted to

The Faculty of
The Columbian College of Arts and Sciences
of The George Washington University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

August 31, 2015

Dissertation directed by

James E. Foster
Professor of Economics and International Affairs

The Columbian College of Arts and Sciences of The George Washington University certifies that Chrysanthi Hatzimasoura has passed the Final Examination for the degree of Doctor of Philosophy as of April 22, 2015. This is the final and approved form of the dissertation.

The Role of Ordinal Data in Economics

Chrysanthi Hatzimasoura

Dissertation Research Committee:

James E. Foster, Professor of Economics and International Affairs, Dissertation Director

Stephen C. Smith, Professor of Economics and International Affairs, Committee Member

Paul E. Carrillo, Associate Professor of Economics and International Affairs, Committee Member

Dedication

To Athona and Philip

Acknowledgments

I do not know where to begin when it comes to thanking people who have hand-held me and, often unbeknownst to them, led to my PhD. So I choose a somewhat chronological order to thank each and every one.

I thank my parents for providing me with everything possible while going through financial hardship. I also thank Takis and Dorina for all their support and unconditional parental love throughout the last 15 years. My sister Elizabeth, a true scientist, who has always been present, forgiving, loving, patient, and supportive. Even after I shred her diary to pieces when we were teenagers. She has never been upset at me for more than 5 minutes. I love her more than past the stars. My brother Jonathan who walked me down the aisle and teaches me to stay true to myself. He is the most genuine soul I know. My beloved sister Barbara, one of the sharpest minds in my life, for keeping me on my toes every time I have an out-of-touch academic view of the world. My one and only auntie Anne who planted the academic seed in me and sent me undergraduate psychology textbooks to read when I was 10 and living in Spetses. My sister Rena for taking a chance and agreeing to go through the casino night-shift madness with me while completing our undergraduate degrees in London – and who I forced to eat Selfridges chocolate-dusted almonds every day. And Giorgos for sitting next to me in all our classes at school and supporting me through illness with humor, laughter, food, lots of presents, optimism, and no judgement. Ever.

I thank my primary school teacher Alexis who told me I will be a great scientist one day based on no evidence whatsoever and for shaping my confidence and aspirations from a young age. Kenneth Van Barthold whose teachings and mentoring in music follow me to this day. My high-school teacher Miroula for instilling the bug in me to explore solutions that lead to social change. I thank Professor Smith for being instrumental in my acceptance to the PhD program and encouraging me to pursue non-conventional pathways in research in economics. I thank Professor Carrillo and Professor Williams for their helpful comments. I thank Dr. Luis-Felipe Lopez-Calva for his comments and for kindly sharing data with me. And last, but most certainly not least, I thank Professor Foster for all his kindness, guidance, and support throughout my PhD research – which included some very difficult times from which I would not have recovered without his help. And for shaping the way I view research in economics. His voice will forever follow me every time I approach a subject.

Abstract of Dissertation

The Role of Ordinal Data in Economics

This thesis is a contribution to the understanding of the use and role of ordinal information in economics. It raises questions about, and in certain instances provides assurance for, current ordinal data practices in the measurement and evaluation of impact. The goals of this thesis are to i) formally explore and organize certain robustness issues that are inherent in analysis with ordinal data ii) introduce methodologies that address these issues and can be applied to specific ordinal data environments and iii) present related applications.

I begin with an exploration of the use of ordinal data in poverty measurement and develop ordinal counterparts to existing cardinal poverty measures. This class of measures is then extended to the multidimensional case and is adjusted in order to be used for analysis that goes beyond poverty measurement, in particular for the context of randomized evaluation of impact on multidimensional ordinal variables. These multidimensional measures are then demonstrated with a randomized controlled trial in Nicaragua where impact on ordinal psychosocial variables, pertaining to aspirations, is evaluated. Finally, the dissertation ends with a pilot project that studies aspirations in Washington DC's underserved youth and is designed in such a way that it takes advantage of the additional insights that my proposed measures offer.

Table of Contents

Table of Figures

Table of Tables

Chapter 1: Introduction

Human capital can be broadly defined as the stock of one's education, health, and psychosocial capital, or, more generally, capital that resides within an individual. Economists regard human capital to be an important component of the economic lives and choices of individuals. Notably, it is often measured with ordinal data. This type of data does not allow for the same kind of analysis that cardinal data allow for. Psychologists and other social scientists have extensively used ordinal data in their studies for the past 50 years. But it is relatively recent that applied economists include ordinal variables, which are not measured on the real line, in their studies. Findings that rely on ordinal data are robust when the research methodologies are based on rank-correlation. However, most empirical studies in economics typically use regression-based analysis. Such analysis may not be robust when using ordinal data as if they are cardinal since regression findings rely on a mean-based aggregation of ordinal information. Mean-based analysis assigns meaning to the distances between ordered numbers on an ordinal scale.

In the field of Econometrics, there exist variations of the probit and logit estimation models that are suitable for establishing causality when variables are recorded on multiple ordered categories. This thesis however does not pertain to the estimation of ordered thresholds for an underlying latent variable. Rather, it is concerned with the *aggregation* of information captured by responses on ordinal scales and the use of such aggregate measures in an evaluation setting. It is well known that the simplest form of regression analysis that estimates the average treatment effect in randomized evaluations can be reduced to the difference in conditional expectations: between those who are treated and those who are not treated. Estimates of the average treatment effect that rely on ordinal data may not be robust since conditional expectations are mean-based aggregations of ordinal information.[1]

Ordinal indicators may be developed so that there is some form of assurance that the distances between ordered categories are well defined, an example being the Likert scale which is used for psychometric indicators. As was mentioned before, the distances are not a concern for rank-correlation based analysis. But researchers may care about distances in scales not only for the purposes of aggregation, which takes place after the data has been collected, but also for collecting responses that match the perception of distances by respondents[2]. Often however, particularly in economics and education, ordinal scales do not possess any logical assurance that the distances between chosen ordered categories for an indicator are well defined. This holds true even when each ordered state on a scale is assigned a specific description. Consider, for example, an empowerment indicator that captures whether a woman is free to leave her house. This indicator may have three ordered categories indicating the extent to which she is allowed leave the house: The first category may be described as "not at all", the second category as "accompanied by her husband", and the third category as "alone". Though these categories are logically ordered on an ascending empowerment continuum, the distances between the ordered states are not well defined.

For the case that distances cannot be defended as being well defined, I propose robust evaluation methods that ultimately aggregate individual ordinal responses of a chosen population into an estimated aggregate index.[3] To do this, I borrow insights from the poverty and inequality measurement literature. This literature conveniently lends itself for this purpose due to its axiomatic structure. It explicitly studies whether properties

[1] This concern only applies to cases of more than 2 ordered categories, not 0-1 binary information.
[2] For example, respondents may perceive upper ordered categories as being closer to each other than lower categories. The goal is to ensure that the ordinal scale reflects how respondents distinguish between ordered categories.
[3] The chosen population could be a conditional population in regression analysis.

of measures are satisfied based on how the available information, whether cardinal or ordinal, is aggregated into a single index. Meaningfully aggregating data should lead to robust measures that maintain the property of reflecting changes in information that the researcher considers to be important. Treating the numerical representation of an ordinal scale as cardinal is not a robust practice unless one is assured that the distances between the ordered categories are well defined. I propose ordinal measures that rely on the ordering and the number of categories in an ordinal scale.

The simplest way with which ordinal data can be aggregated into a robust measure is by dichotomizing ordinal information into two subsets of the ordered categories and then assigning a value of zero and one to each of them. In the previous empowerment example, a researcher may assign a value of zero to the first category "not at all" and a value of one to the second and third ordered categories "accompanied by her husband" and "alone". This binary approach is often employed in practice and is useful. It ensures that a mean-based aggregation leads to evaluation rankings that are robust. However, the researcher may want the estimates to be sensitive to variation within the upper two categories so that the information that was collected on leaving the house alone or accompanied is also accounted for. For the case that the researcher wishes for the aggregate measure to reflect the variation of responses within all ordered categories, so that it is not silent in an evaluation, the binary approach is not sufficient.

With ordinal scales it is not always clear how average scores can represent a tangible concept and whether their implicit measurement properties will affect their use in evaluations. Can an increase in the -non cardinal and hence non-continuous- score be meaningfully inferred as an outcome variable impact? When a resulting index is an average, what is the interpretation that this average receives? In addition to these inference issues, it is also important to understand whether the measures can be used robustly in regressions. When used in regressions, what measurement properties ensure robustness in the evaluation rankings? To address all the above, I present five chapters, including the introduction which is the first chapter.

The second chapter, introduces a class of unidimensional poverty measures that are suitable for reflecting any change in observations from one ordered category to another. Several of the most widely used measures in theoretical and empirical work in development economics presuppose a dimension of well-being that, like income, is cardinally measurable. Responding to recent interest in dimensions of well-being where outcome variables are recorded on an ordinal scale, this paper introduces a general methodology for constructing indices for ordinal variables that are robust to any arbitrary choice of scale. It will also be shown how this methodology may be applied to construct an ordinal analogue of the popular FGT class of poverty measures. The resulting ordinal measures retain the simplicity of the classical FGT poverty measures and some of their desirable features, including additive decomposability. Additionally, they satisfy ordinal invariance.

In the third chapter, I explore the current use of ordinal data in evaluations. I use randomized evaluation as a benchmark framework. Scales may be developed so that the aggregation of individual responses within various indicators results in a score that is intended for factor analysis and other similar methodologies. This kind of analysis does not lead to ambiguous evaluations since it is driven by an underlying rank-correlation approach. Economists' recent interest in psychosocial capital has led to increased popularity of these scales in empirical studies, particularly evaluations. In applied studies, the aggregation of the numerical values of ordered responses into single scores may appear as outcome variables in regressions. I re-examine the average treatment effect through a measurement perspective, thus adding to our understanding of the role of measurement in various types of evaluations using the randomized evaluation framework as a benchmark. The findings apply to evaluations that are randomized, retrospective, or assessments. I simulate randomized evaluations and run robustness checks for a variety of ordinal scales. Evaluation reversals occur with non-negligible incidence unless we are assured that the distances between ordered states are equal or otherwise well defined. I then propose a class of multidimensional measures for ordinal outcome variables that are useful for evaluations.

In the fourth chapter, I present an applied theoretical framework of a randomized controlled trial in Nicaragua. I evaluate the impact of an entrepreneurial package intervention using the methodology introduced in the previous chapter. Self-reported decision-making constructs are measured using dimensions which are recorded on ordinal scales, and examined in order to evaluate the impact on those treated. The applied framework adds to our understanding of the role of women's aspirations in the formation of their relative power in the household -which I represent by their Pareto weight in the efficient household model. In the function of bargaining empowerment the Pareto weight is not simply a function of aspirational levels. Rather, I introduce it as a function of the aspirations gap. The efficient household model assumes that existing preferences should be reflected in the Pareto weights. But by distinguishing between the beneficiary *meeting* her aspirations and simply *having* pre-existing aspirations, I challenge this assumption. I find that after holding aspirational preferences constant, relaxing income generating constraints for women leads to higher intrahousehold bargaining power levels or, equivalently, it leads to intrahousehold bargaining empowerment for women. Interestingly, these results hold true even though overall household income does not increase, which is in line with existing findings on the ability to earn independent income and bargaining empowerment.

In the fifth chapter, I propose a pilot study which targets adolescent populations in some of Washington DC's most underserved communities. I design and propose to implement a pilot attitudinal experiment in order to test the hypotheses derived from an applied framework. I model underrepresentation as a poverty trap that is driven by extrinsically updated subjective probabilities. I hypothesize that subtle social signals, controlling for information, can lead to attitudinal reversals. The pilot will add to our understanding of the role of many factors that contribute to the creation of aspirations -and subsequently economic decision making such as staying in high school. Subjective probabilities are recorded on ordinal scales, which I will then analyze using my proposed methodology in the third chapter.

Chapter 2: Poverty Measurement with Ordinal Data[4]

2.1 Introduction

The Foster, Greer, and Thorbecke (1984) class of measures nests several of the most widely used indices in theoretical and empirical work on economic poverty. Use of this general class of indices, however, presupposes a dimension of well-being that, like income, is cardinally measurable. Responding to recent interest in dimensions of well-being where achievements are recorded on an ordinal scale, this paper introduces a general methodology for constructing ordinal indices of poverty and, in particular, shows how this methodology may be applied to construct an ordinal analogue of the popular FGT class of indices. The resulting ordinal FGT indices retain the simplicity of the classical FGT indices and also many of their desirable features, including additive decomposability. To illustrate their use, we apply the ordinal FGT indices to self-reported data on health status in Canada and the United States.

In the thirty years since it was first introduced, the FGT family of measures has become the most widely used class in empirical work on the measurement of poverty. The attractiveness of the FGT measures stems largely from their simple structure, their ease of interpretation, and their sound axiomatic properties. Being defined by two parameters, namely the poverty line ℓ and a scalar measure of poverty aversion α, each member of the FGT class is easily computed as an average of the power function defined by α whose argument is the normalized income shortfall from ℓ. Specific members include the well-known poverty gap, squared poverty gap, and headcount ratio (i.e., the proportion of the population identified as poor).

[4] Co-authored with Chris Bennett

Use of the general FGT class of measures presupposes a dimension of well-being that, like income, is cardinally measurable. Recently, however, considerable interest has emerged in measures of aggregate deprivation in dimensions of well-being other than income, and, in particular, in dimensions of well-being—for example, health, education, empowerment, and social inclusion, etc.—which are often recorded on an ordinal scale.[5] For example, individual level health data is often available in the form of self-assessments in which survey participants are asked to characterize their health status as either *poor*, *fair*, *good*, *very good*, or *excellent*. While it is common to assign numerical levels such as 1, 2, 3, 4, and 5 to the individual responses and even for such a scale to capture a sense of intensity, except for the headcount ratio, none of the FGT measures are meaningful when applied to such data. As a consequence, "a crucial emerging issue is how to measure poverty when data do not have the characteristics of income, which is typically taken to be cardinal and comparable across persons ... Must we retreat to the headcount ratio [with ordinal data], or can we continue to evaluate the depth or distribution of deprivations—key benefits provided by the higher order FGT measures when the variable is cardinal?" Foster, Greer, and Thorbecke (2010).

In this paper, we introduce a computational counterpart to the classical FGT class of measures for use with variables measured on an ordinal scale. Our particular construction gives rise to an ordinal counterpart to the FGT class of measures which (i) has sound axiomatic properties; (ii) retains many of the attractive properties of the classical FGT measures (including, for example, additive decomposability); and yet (iii) is without the obvious shortcomings inherent in the application of conventional poverty measures to ordinal data.

[5] Problems surrounding the measurement of poverty with ordinal data are raised, for example, in Foster, Greer, and Thorbecke (2010) and Alkire and Foster (2011a,b). Allison and Foster (2004) were the first to stress the problems raised by ordinal data in the related context of inequality measurement. See Zheng (2008), Abul Naga and Yalcin (2004), and Madden (2010) for more on the use of ordinal data in this latter context.

A particularly important concern that arises when applying measures to ordinal data has to do with the arbitrary nature of the numerical values representing the various levels of achievement. Specifically, because no individual is made better or worse off by an order-preserving transformation applied to the levels of achievement and to the poverty line, any suitable poverty measure (or at least its induced poverty ordering) should be invariant to *any* such transformation.[6] However, it is easy to construct examples in which the ordering of two distributions by virtually any existing measure is reversed when the levels of achievement and poverty line are subjected to a positive monotonic transformation.

The ordinal FGT measures developed in this paper not only avoid this obvious shortcoming, but they also satisfy a number of other attractive properties. In the literature on income poverty, for example, there are a number of core properties generally regarded that any poverty measure should satisfy. A further contribution of the present paper is our development of ordinal analogues of certain FGT properties, and our verification of the axioms that are satisfied by our proposed measures.

We briefly review the FGT class of poverty measures. In the next section, we also present an equivalent reformulation of this class which gives rise to an interesting interpretation of the FGT measures. We exploit this novel reformulation to construct an ordinal counterpart to the classical FGT measures. We then develop a formal axiomatic framework for the measurement of poverty with ordinal variables and identify the axioms that are satisfied by our proposed class of measures. In the final section, we illustrate the application of our poverty measures to self-reported health data from the

[6] This point is emphasized in Alkire and Foster (2011b, p. 306) where they state that "the key requirement for ordinal data is that if the cut-off and variables are changed by a monotonic transformation, the level of poverty must remain unchanged, and the same people must be identified as poor." Also, Allison and Foster (2004) discuss this problem in the context of measuring inequality with self-reported health data.

United States and Canada. When applied to this dataset, our measures suggest that there is unambiguously greater ill-health in the United States than in Canada for the bottom 20% of their income distributions.

2.2 The FGT Class of Poverty Measures

Foster, Greer, and Thorbecke (1984), hereafter FGT, introduced the class of poverty measures[7]

$$\pi_{\alpha,\ell}(Y) = \mathbf{E}_Y[[g(Y;\ell)]^{\alpha} 1(Y \leq \ell)] \quad \alpha, \ell \in \mathbb{R}_+ \tag{2-1}$$

where Y is income, ℓ is the poverty line, $g(y;\ell) = (\ell - y)/\ell$ is the normalized income shortfall, and $1(\cdot)$ is the indicator function which is equal to one if the argument is true and equal to zero otherwise.[8] The parameter α which appears in (2-1) controls the degree to which shortfalls are penalized and is therefore often interpreted as an indicator of "poverty aversion." The resulting parametric class nests several of the most widely used measures in both theoretical and empirical studies of poverty measurement. When $\alpha = 0$, for example, the measure reduces to

$$\pi_{0,\ell}(Y) = \mathbf{P}[Y \leq \ell]$$

which is the proportion of the population in poverty, commonly known as the *headcount ratio*. When $\alpha = 1$, the FGT class gives rise to the *poverty gap* measure which has a simple interpretation as the average normalized shortfall. The *squared poverty gap* measure, which reports the average of the *squared* normalized shortfalls, also emerges from the FGT class upon setting $\alpha = 2$. The popularity of these measures stems not only from their ease of interpretation—being based on powers of normalized

[7] The FGT class of measures is commonly denoted by P_α, where, for example, P_0 is the headcount ratio, P_1 is the poverty gap measure, and P_2 is the squared gap measure. Here, we depart from this conventional notation because of our later usage of probability measures and, hence, our desire to prevent any possible confusion between these objects.

[8] Here, as throughout, we adopt the strong definition of the poor (Donaldson and Weymark 1986) under which the poor consists of all individuals with achievement level less than or equal to the poverty line.

shortfalls—but also from their sound axiomatic properties which include the attractive additive decomposability and subgroup consistency properties. For a detailed account of the FGT class of measures, including a discussion of the axiomatic foundations and related literature, the reader is referred to Foster, Greer, and Thorbecke (2010).

The classical formulation of the FGT class of measures in terms of "shortfalls" relies critically on a well-defined notion of distance between an individual's level of income and the poverty cut-off. While this formulation is quite natural and appealing in the context of income poverty where the distance between an individual's level of income and the poverty cut-off is not only meaningful but also quite easy to interpret, it does not carry over in any natural way to situations in which the variable of interest has only ordinal significance. Our objective in this section is to develop an alternative formulation of the FGT class of measures which not only gives rise to a novel interpretation of them as the expected outcomes of particular income lotteries, but which also allows us to adapt the FGT class quite naturally to situations where the underlying variables are ordinal.

To begin, let F denote the distribution of income and Y denote a random draw from F. Also, let U_l be the random variable which is uniformly distributed on $[0,\ell]$ but which is stochastically independent of Y. We then define the survival function of U_ℓ as $G_\ell(y) = \mathbf{P}[U_\ell \geq y]$, and for $\alpha > 0$ we denote by $U_{\alpha,\ell}$ the random variable with survival function $G_\ell^{(\alpha)}(y) \equiv [G_\ell(y)]^\alpha$. Lastly, in the case $\alpha = 0$, we define $U_{0,\ell}$ to be a random variable which has probability mass one at ℓ. With $U_{\alpha,\ell}$ in hand, we introduce the binary random variable

$$\mathrm{I}_{\alpha,\ell} = 1(Y \le U_{\alpha,\ell})1(Y \le \ell). \tag{2-2}$$

Thus, $\mathrm{I}_{\alpha,\ell}$ is equal to one if, and only if, an individual's income generated from F is less than the poverty line *and* less than the realization of the random draw $U_{\alpha,\ell}$ from incomes below the poverty line.

Of particular interest for the measurement of poverty is the expected value of this variable, namely

$$\begin{aligned} \mathbf{E}_{Y,U_{\alpha,\ell}}[1(Y \le U_{\alpha,\ell})1(Y \le \ell)] &= \mathbf{E}_{Y,U_{\alpha,\ell}}[\mathrm{I}_{\alpha,\ell}] \\ &= \mathbf{P}[\mathrm{I}_{\alpha,\ell} = 1], \end{aligned} \tag{2-3}$$

which is the *probability* that an individual would be better off with the realization of $U_{\alpha,\ell}$ rather than the realization of Y. In order to connect this probabilistic statement to classical approaches to poverty measurement, consider taking expectations of (2-2) first with respect to $U_{\alpha,\ell}$. Doing so yields

$$\mathbf{E}_{Y,U_{\alpha,\ell}}[1(Y \le U_{\alpha,\ell})1(Y \le \ell)] = \mathbf{E}_Y\left[(\frac{\ell - Y}{\ell})^{\alpha} 1(Y \le \ell)\right], \tag{2-4}$$

which is the FGT measure $\pi_{\alpha,\ell}(Y)$. Consequently, we see that each member of the FGT class of measures when applied to income has a dual interpretation: first, as the average of (a functional of) normalized income shortfalls; and second, as the probability of accepting the income which is drawn at random from the poor incomes according to $G_{\ell}^{(\alpha)}$.

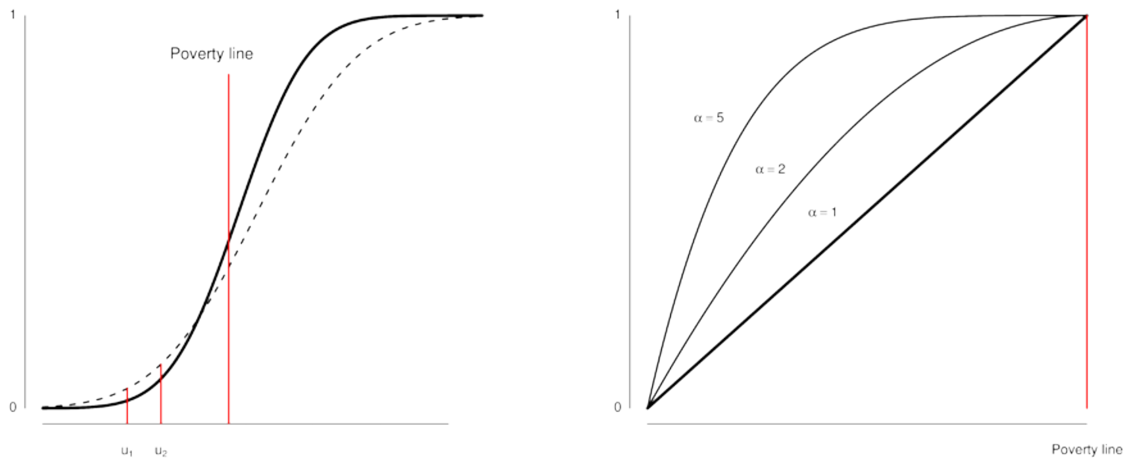
When we reverse the order of expectations we obtain an alternative formulation of the FGT class of measures as

$$\begin{aligned}\mathbf{E}_{Y,U_{\alpha,\ell}}[1(Y \le U_{\alpha,\ell})1(Y \le \ell)] &= \mathbf{E}_{U_{\alpha,\ell}}[F_{Y|Y\le\ell}(U_{\alpha,\ell})] \times \mathbf{P}[Y \le \ell] \\ &= \mathbf{E}_{U_{\alpha,\ell}}[F(U_{\alpha})]\end{aligned} \qquad (2\text{-}5)$$

for $\alpha, \ell \in \mathbb{R}_+$. Equation (2-5) shows that each member of the FGT class is also equivalent to an expected headcount, where the expectation is with respect to a random variable from the set $\{U_{\alpha,\ell}, \alpha \ge 0\}$. That is, imagine using the realization of $U_{\alpha,\ell}$ as the poverty line in the computation of the headcount. The last line of equation (2-5) shows that the FGT measure with parameters (α, ℓ) is in fact equivalent to the average of the headcounts corresponding to the various realizations of $U_{\alpha,\ell}$.

This latter formulation of the FGT class of measures gives rise to an interesting reinterpretation in the spirit of Harsanyi (1953) and Rawls (1999), where one's value judgement on the distribution of income corresponds to a hypothetical situation in which an individual has complete ignorance of her own relative position in society and instead faces an income lottery. When $\alpha = 1$, for example, the income lottery is such that the individual has the same chance of obtaining the lowest income as any other level of income up to the poverty line. The expectation in the last line of (2-5) may therefore be interpreted as the percentile of the income distribution an individual, say, would expect to find herself in if faced with the prospect of drawing her income at random from a $U[0,\ell]$ distribution.

Figure 2-1: FGT measures as expected outcomes of an income lottery



(a) Comparing the values of two hypothetical CDFs at random draws u_1 and u_2]

(b) Distributions generating the u_i for various levels of α]

When $\alpha = 2$, the measure may again be interpreted as an expected percentile albeit with the expectation taken with respect to the distribution $1 - G_\ell^{(2)}(y)$. Because the distribution $1 - G_\ell^{(2)}(y)$ as depicted in Figure 2-1(a) is everywhere above the uniform CDF (i.e., is stochastically dominated by the uniform CDF), there is a lower chance of observing a draw close to the poverty line (i.e., the probability of observing u_2 rather than u_1 in Figure 2-1(a) has fallen) and hence the corresponding percentiles contribute less to the expected value. This interpretation is also equivalent to regarding $\pi_{2,\ell}$ as a weighted average of the headcount ratios evaluated at all incomes in the interval $[0, \ell]$, but with weights tilted towards the poorest of the poor. Similar interpretations may be obtained for $\alpha > 2$.

In general, rather than framing the FGT measures in terms of average (normalized) shortfalls, the modified formulation obtained in the last line of equation (2-5) frames each FGT measure as an expected value of headcount ratios. Because shortfalls are no longer interpretable in the case of ordinal variables, whereas headcount ratios and even averages of headcount ratios are, the latter interpretation suggests that the FGT class of measures can be adapted so as to apply to ordinal variables.

2.3 Extending the FGT class of measures to ordinal variables

In this section we seek to extend the FGT class for use with measures of achievement that are recorded on an ordinal scale. We assume that the number of ordered categories of achievement, denoted by n, is exogenously determined, and we let $S = \{y_1, y_2, \ldots, y_n\}$ denote a numerical representation where $y_{i+1} > y_i$.[9] Let F record the distribution of achievements for a given population, and denote by Y a random draw from F. Thus, $F(y_1)$ is the proportion of individuals in the population with level of achievement y_1, whereas $F(y_k) - F(y_{k-1})$ is the proportion of individuals in the population with level of achievement y_k. We allow for the possibility that F apportions zero probability to one or more of the n levels of achievement.

Given a pre-specified poverty line $\ell \in \{y_1, \ldots, y_n\}$, we then define $U_{S,\ell}$ to be a random variable which is uniformly distributed on $S \cap [0, \ell]$. Also, we define the survival function $G_{S,\ell}(u) = P[U_{S,\ell} \geq u]$, and let $U_{\alpha,S,\ell}$ denote a random variable with survival function $G_{S,\ell}^{(\alpha)}(y) \equiv [G_{S,\ell}(y)]^{\alpha}$ for $\alpha > 0$. In the case $\alpha = 0$, we define $U_{0,S,\ell}$ to be a random variable which has probability mass one at the poverty line ℓ. With $U_{\alpha,S,\ell}$ in hand, define the random variable

---

[9] The numerical representation is arbitrary apart from the restriction that $0 \leq y_1 < \cdots < y_n < \infty$.

$$\mathrm{I}_{\alpha,\ell} = 1(Y \le U_{\alpha,S,\ell})1(Y \le \ell). \tag{2-6}$$

The random variable $\mathrm{I}_{\alpha,\ell}$ is the analogue of the binary random variable previously defined in (2-2) and has a similar interpretation. Specifically, the variable $I_{\alpha,\ell}$ assumes the value 1 if, and only if, the realization of Y according to F falls at or below the realization of $U_{\alpha,S,\ell}$ according to $G_{S,\ell}^{(\alpha)}$. The expected value of $\mathrm{I}_{\alpha,\ell}$ is again the probability that an individual would be better off with the realization of $U_{\alpha,S,\ell}$ as their level of well-being than the realization of Y. Expressed in terms of an average of headcounts, we have

$$\begin{aligned} \pi_{\alpha,\ell}(Y) &= \mathbf{E}_{Y,U_{\alpha,S,\ell}}[1(Y \le U_{\alpha,S,\ell})1(Y \le \ell)] \\ &= \mathbf{E}_{U_{\alpha,\ell}}[F_{Y|Y\le\ell}(U_{\alpha,S,\ell})] \times \mathbf{P}[Y \le \ell] \\ &= \mathbf{E}_{U_{\alpha,\ell}}[F(U_{\alpha,S,\ell})]. \end{aligned} \tag{2-7}$$

Because the distribution of $U_{\alpha,S,\ell}$ has its support restricted to the set $S \cap [0,\ell]$, the classical formulation of the FGT class of measures in terms of shortfalls is no longer recovered from (2-7) by simply reversing the order of integration. Notice, however, that if Y is indeed a cardinal random variable with support on the positive reals, then $S \cap [0,\ell] = [0,\ell]$, in which case $U_{\alpha,S,\ell}$ is as defined in (2-2), and the classical FGT class *is* recovered by reversing the order of integration. In this sense, the formulation of $\pi_{\alpha,\ell}$ in (2-7) may be seen to nest the classical FGT measures.

Irrespective of whether the underlying variable is ordinal or cardinal, the modified formulation in (2-7) retains the interpretation as an average headcount (or expected percentile) and also the dual interpretation as an expected poverty status. For example, when $\alpha = 1$, one's value judgement corresponds to the expected percentile in the distribution of levels of achievement in society when an individual has complete ignorance of his own relative position in society and instead faces a lottery with the same chance of obtaining the lowest level of achievement as any other possible level of achievement up to the poverty line.

It is worth emphasizing that the key distinction between (2-7) and the classical formulation of the FGT measures—cf., equation (2-5)—lies in the expectation being over a potentially finite number of states $y_1, \cdots, y_k$ where y_k, is the most preferred state for which someone is still identified as poor, rather than necessarily being over the continuum $[0, \ell]$. It is also important to emphasize that this distinction is not merely technical: when Y is a cardinal variable, such as income, distances between y_1 and y_k, or between y_1 and any $y \in [y_1, y_k]$ are well defined; in contrast, when Y is ordinal, distances are no longer meaningful and, moreover, "levels" of achievement outside of the set $\{y_1, y_2, \ldots, y_n\}$ are not well-defined.

In addition to retaining precisely the same interpretation, the ordinal class of FGT measures as defined in (2-7) also retain the computational simplicity of the classical FGT class of measures. To see this, suppose that Y is an ordinal random variable with distribution $(p_1, y_1; \ldots; p_n, y_n)$. Also, for concreteness, suppose that $\ell = y_k$.[10] Then, enumerating the values of the survival function for any fixed $\alpha > 0$, we obtain

$$[G_{S,\ell}(y_{k-j})]^\alpha = (\frac{j+1}{k})^\alpha, \quad j = 0, \ldots, k-1.$$

Note that

[10] Incidentally, because the support of the random variable $U_{\alpha,S,\ell}$ is $\{y_1, \ldots, y_n\} \cap [0, \ell]$, the computational formula is the same for any value of the poverty line in the interval $[y_k, y_{k+1})$.

$$[G_{S,\ell}(y_1)]^{\alpha} = 1.$$

The probability mass at y_j is thus given by

$$\mathbf{P}[U_{\alpha,S,\ell} = y_j] = \left(\frac{k-j+1}{k}\right)^{\alpha} - \left(\frac{k-j}{k}\right)^{\alpha} \tag{2-8}$$

for $j \in \{1,\ldots,k\}$. With the probability mass function (2-8) in hand, we immediately obtain the computational formula for the ordinal measure $\pi_{\alpha,\ell}$ for any $\alpha > 0$ as

$$\begin{aligned}\pi_{\alpha,\ell}(Y) &= \sum_{j=1}^{k} F(y_j)\left[\left(\frac{k-j+1}{k}\right)^{\alpha} - \left(\frac{k-j}{k}\right)^{\alpha}\right] \\ &= \sum_{j=1}^{k} p_j \left(\frac{k-j+1}{k}\right)^{\alpha}\end{aligned} \tag{2-9}$$

where $F(y_j) = \sum_{i \le j} p_i$, and the last line follows as a consequence of Abel's partial summation formula Apostol (1974).

2.4 Properties of the Ordinal FGT Class of Measures

A number of intuitively appealing desiderata (axioms) have been put forth in the literature on income poverty; see, for example, Zheng (1997). While some of these axioms carry over without modification to the present context (e.g., the focus, symmetry, and replication invariance axioms), many do not. In fact, a number of axioms which are widely considered to be intuitively appealing in the context of income poverty become rather unappealing when the underlying measure of achievement has only ordinal significance. In this section, we collect—and suitably reformulate when necessary—a set of core axioms for poverty measurement with ordinal variables. We also identify the particular axioms that are satisfied by various members of our proposed parametric class of ordinal FGT measures.

A particularly important issue that arises in the context of measurement with ordinal data has to do with the arbitrary nature of the numerical representation assigned to the levels of achievement. Specifically, because no individual is made better or worse off by an order-preserving transformation applied to the levels of achievement and to the poverty line, any suitable poverty measure (or at least its induced poverty ordering) should be invariant to *any* such transformation. This leads to the following *invariance* axiom:

Axiom 2.1 (Ordinal Invariance): Suppose that ℓ is given and that Y has distribution $(p_1, y_1; p_2, y_2; \ldots; p_n, y_n)$. If $g : \mathbf{R} \to \mathbf{R}$ is monotone increasing, then

$$\pi_{\alpha,\ell}(Y) = \pi_{\alpha,g(\ell)}(g(Y))$$

for every $\ell \in \{y_1, \ldots, y_n\}$ and all $\alpha > 0$.

Ensuring that the ordinal FGT measures satisfy Ordinal Invariance is particularly important in allowing us to apply the measure to ordinal variables as it implies that only changes in individuals' ranks in society can produce changes in the measured level of poverty. To see this more clearly, let Y be an ordinal random variable with distribution $(p_1, y_1; \ldots; p_n, y_n)$. Also, for concreteness, consider the case $\alpha = 1$ so that $U_{1,S,\ell}$ is uniformly distributed on $S \cap [0, \ell]$ (again independent of Y), and suppose that $\ell = y_k$. Then,

$$
\begin{aligned}
\pi_{1,\ell}(Y) &= \mathbf{E}_{Y,U_{1,S,\ell}}[1(Y \le U_{1,S,\ell})1(Y \le \ell)] \\
&= \mathbf{E}_{U_{1,S,\ell}}[F_{Y|Y\le\ell}(U_{1,S,\ell})] \times \mathbf{P}[Y \le \ell] \\
&= \frac{\frac{1}{k}\sum_{j=1}^{k}(\sum_{i=1}^{j} p_i)}{\sum_{j=1}^{k} p_j} \times \mathbf{P}[Y \le \ell] \\
&= \frac{p_1}{k} + \frac{(p_1+p_2)}{k} + \cdots + \frac{(p_1+p_2+\ldots+p_k)}{k}
\end{aligned} \tag{2-10}
$$

Now suppose that Y is subjected to a positive monotonic transformation which gives rise to $\tilde{Y} \equiv g(Y)$ with distribution $(p_1, \tilde{y}_1; \ldots; p_n, \tilde{y}_n)$. The corresponding random variate $U_{1,\tilde{S},\tilde{\ell}}$ is then, by definition, uniformly distributed on $[\tilde{y}_1, \tilde{y}_2, \cdots, \tilde{y}_k]$. We then have

$$
\begin{aligned}
\pi_{1,m(\ell)}(\tilde{Y}) &= \mathbf{E}_{\tilde{Y},U_{1,\tilde{S},\tilde{\ell}}}[1(\tilde{Y} \le U_{1,\tilde{S},\tilde{\ell}})1(\tilde{Y} \le g(\ell))] \\
&= \mathbf{E}_{U_{\tilde{Y},1}}[F_{\tilde{Y}|\tilde{Y}\le\ell}(U_{1,\tilde{S},\tilde{\ell}})] \times \mathbf{P}[g(Y) \le g(\ell)] \\
&= \frac{\frac{1}{k}\sum_{j=1}^{k}(\sum_{i=1}^{j} p_i)}{\sum_{j=1}^{k} p_j} \times \mathbf{P}[Y \le \ell] \\
&= \frac{p_1}{k} + \frac{(p_1+p_2)}{k} + \ldots + \frac{(p_1+p_2+\ldots+p_k)}{k}
\end{aligned} \tag{2-11}
$$

so that the modified FGT measure is unchanged. Invariance of the general class of ordinal FGT measures to positive monotonic transformations may be verified with only a slight modification of the argument presented above and, hence, the proof is omitted.

If one accepts Ordinal Invariance as a basic axiom in the context of poverty measurement with ordinal data, then outside of our ordinal FGT measures there is currently no option other than to employ the headcount ratio as it is the only measure in the FGT class that does not violate these elementary conditions. However, the well-known shortcomings associated with the headcount ratio as a measure of poverty make the use of this measure in the context of ordinal data most unfortunate. Indeed, it was Sen's influential 1976 paper in which he pointed to the failure of the headcount ratio to satisfy a basic monotonicity axiom and a transfer axiom that gave impetus to the now vast literature on poverty measurement. Fortunately, the class of ordinal FGT measures can avoid these shortcomings. To be more precise, let $y_1,\ldots,y_k$ denote the states of poverty where $y_j \succ y_i$ for any $j > i$ with "$\succ$" denoting the strict preference relation. The monotonicity axiom in the case of income poverty relates to the response of a poverty measure to an increase or decrease in a single individual's level of income. When translated into statements about the distribution, such an increase or decrease amounts to a transfer of probability mass. Consequently, we formulate the ordinal analogue of Sen's monotonicity axiom as follows:

Axiom 2.2 (Monotonicity): Let ℓ be given and suppose that $(p_1, p_2,\ldots, p_k)$ with $\sum_{i\le k} p_i \le 1$ denotes the distribution of Y in each of the corresponding k states of poverty. If $\tilde{Y}$ has the same distribution as Y with the exception that $\tilde{p}_j = p_j - \varepsilon$ for some $j \in \{1,\ldots,k\}$ and $\tilde{p}_i = p_i + \varepsilon$ for some $i < j$, then $\pi_{\alpha,\ell}(\tilde{Y}) > \pi_{\alpha,\ell}(Y)$.[11]

Similarly, transfer in the context of income poverty relates to the response of a poverty measure to the reallocation of income from one individual to another. When translated into statements about the distribution, a transfer shifts the rank of both individuals and, hence, amounts to two separate transfers of probability mass.

[11] Implicit in our statement of the monotonicity axiom is that $0 < \varepsilon \le p_j$. Similar restrictions on the transfer of probability mass is maintained in our statement of transfer axiom.

Axiom 2.3 (Transfer): Let ℓ be given and suppose that $(p_1, p_2, \ldots, p_k)$ with $\sum_{i \le k} p_i \le 1$ denotes the distribution of Y in each of the corresponding k states of poverty. If $\tilde{Y}$ is generated from the distribution of Y by transfers of probability mass $\tilde{p}_i = p_i - \varepsilon$ offset by $\tilde{p}_{i-m} = p_{i-m} + \varepsilon$, and $\tilde{p}_j = p_j - \varepsilon$ offset by $\tilde{p}_{j+m} = p_{j+m} + \varepsilon$, then $\pi_{\alpha,\ell}(\tilde{Y}) > \pi_{\alpha,\ell}(Y)$ whenever $1 \le i - m < m \le k$ and $i \le j$ with $i, j, m \in S \cap [0, \ell]$.

We now show that the ordinal FGT class of measures satisfies Monotonicity when $\alpha > 0$ and Transfer when $\alpha > 1$. To do this we use the formula for $\pi_{\alpha,\ell}$ in the first line of (3 1), and suppose that $\tilde{Y}$ with distribution $(q_1, y_1; \ldots; q_n, y_n)$ is obtained from $(p_1, y_1; \ldots; p_n, y_n)$ by shifting probability mass from p_i to p_{i-m}. More precisely, suppose that $q_i = p_i - \varepsilon$ and $q_{i-m} = p_{i-m} + \varepsilon$ for some $\varepsilon > 0$ with $m + 1 \le i \le k$. The difference in the levels of poverty according to the measure $\pi_{\alpha,\ell}$ is then given by

$$\begin{aligned} \pi_{\alpha,\ell}(\tilde{Y}) - \pi_{\alpha,\ell}(Y) &= \sum_{j=1}^{k} \left(\sum_{r=1}^{j} (q_r - p_r) \right) \left[\left(\frac{k-j+1}{k} \right)^{\alpha} - \left(\frac{k-j}{k} \right)^{\alpha} \right] \\ &= \sum_{j=1}^{k} \varepsilon 1(i-1 \le j \le i) \left[\left(\frac{k-j+1}{k} \right)^{\alpha} - \left(\frac{k-j}{k} \right)^{\alpha} \right] \end{aligned} \qquad (2\text{-}12)$$

That Monotonicity is satisfied for $\alpha > 0$ is thus verified upon noting that the second line is strictly positive for any $\alpha > 0$. The transfer axiom is closely related to the monotonicity axiom, albeit it involves two transfers of probability mass in equal amounts of, say, ε. If $\tilde{Y}$ denotes the outcome of such a transfer, then the effect on poverty as measured by $\pi_{\alpha,\ell}$ is of the form

$$\pi_{\alpha,\ell}(\tilde{Y})-\pi_{\alpha,\ell}(Y)= \sum_{r=1}^{k}\varepsilon[1(i\le r\le i+m)-1(j-m\le r\le j)] \times\left[\left(\frac{k-r+1}{k}\right)^{\alpha}-\left(\frac{k-r}{k}\right)^{\alpha}\right], \tag{2-13}$$

where $i, j,$ and m are integers satisfying $i+m\le j$. Clearly, such transfers have no net effect on poverty when $\alpha=1$, and are poverty decreasing only when $\alpha>1$. It follows that the class of poverty measures $\pi_{\alpha,\ell}$ satisfy Transfer whenever α exceeds 1.

The last set of axioms we wish to consider are Additive Decomposability and Subgroup Consistency. Both are satisfied by the classical FGT measures and are widely considered to be important since they "allow poverty to be evaluated across population subgroups in a coherent way", (Foster et al 2010, p. 492). These axioms are stated formally below:

Axiom 2.4 (Additive Decomposability): Suppose that Y_1 and Y_2 have distributions of achievement P_1 and P_2, and suppose that Y has distribution of achievement $P=\lambda P_1+(1-\lambda)P_2$. Then,

$$\pi_{\alpha,\ell}(Y)=\lambda\pi_{\alpha,\ell}(Y_1)+(1-\lambda)\pi_{\alpha,\ell}(Y_2).$$

Axiom 2.5 (Subgroup Consistency): Suppose that Y_1 and Y_2 have distributions of achievement P_1 and P_2, and suppose that Y has distribution of achievement $P=\lambda P_1+(1-\lambda)P_2$. Also, let X_1, X_2, and X have distributions Q_1, Q_2, and

$Q = \lambda Q_1 + (1-\lambda) Q_2$, respectively. Then,

$$\pi_{\alpha,\ell}(Y) < \pi_{\alpha,\ell}(X)$$

whenever $\pi_{\alpha,\ell}(Y_1) < \pi_{\alpha,\ell}(X_1)$ and $\pi_{\alpha,\ell}(Y_2) = \pi_{\alpha,\ell}(X_2)$

Verifying that the ordinal FGT measures satisfy Additive Decomposability and Subgroup Consistency is straightforward in light of the linearity of the expectations operator. The details are therefore omitted.

2.5 A Useful Computational Representation

With ordinal data, there are K ordered *categories* or *states* of achievement represented numerically by an ordered set $\mathbb{Y} = \{y_1, y_2, \ldots, y_K\}$ in such a way that $y_i > y_j$ if, and only if, state i is preferred to state j.[12] The observed levels of achievement in a population of size N are recorded in $\mathbf{y} \in \mathbb{Y}^N$ and individuals within this population are identified as "poor" if they fall into one of the k worst states, or equivalently if their *level* of achievement falls at or below y_k, where $y_k < y_K$.

To demonstrate the computational connection to the classical FGT indices, let $G_{\mathbb{Y}}$ denote the cumulative distribution function that assigns equal probability to the potential achievement levels in $\mathbb{Y}$. The cumulative distribution $G_{\mathbb{Y}}(\cdot)$ is a convenient mathematical device that maps a given ordered response $\mathbf{y}_i \in \mathbb{Y}$ to its corresponding (normalized) achievement rank $G_{\mathbb{Y}}(\mathbf{y}_i) \in \{\frac{1}{K}, \frac{2}{K}, \ldots, 1\}$. Thus, for example, $G_{\mathbb{Y}}(y_K) = 1$ becomes the highest achievement rank and $G_{\mathbb{Y}}(y_j) = j/K$ becomes the

[12] The number of states can be countably infinite. Our focus on the case where the number of states is finite is, however, without loss of generality.

ordered response rank of an individual in the j^{th} response. With the distribution of (normalized) achievement ranks and poverty rank cut-off defined as

$$\mathbf{x} \equiv (G_{\mathbb{Y}}(\mathbf{y}_1), G_{\mathbb{Y}}(\mathbf{y}_2), \ldots, G_{\mathbb{Y}}(\mathbf{y}_N)) \in [0,1]^N,$$

and $z \equiv G_{\mathbb{Y}}(y_{k+1})$, respectively, the proposed indices $\pi_\alpha(\mathbf{y}, y_k)$, $\alpha > 0$ —which operate on the levels—can be presented as

$$\Pi_\alpha(\mathbf{x}; z) = \frac{1}{N}\sum_{j=1}^{N}\left(\frac{z - \mathbf{x}_j}{z - G_{\mathbb{Y}}(y_1)}\right)^{\alpha} \mathbf{1}(\mathbf{x}_j < z), \quad \alpha > 0 \qquad (2\text{-}14)$$

which operate on the (normalized) ranks, with $\mathbf{1}(\cdot)$ denoting the indicator function.

The representation of our proposed measures in (2-14), which is formulated in terms of (normalized) ranks, is identical to the computational formula for the classical FGT class of indices with the exception that the normalizing constant is $z - G_{\mathbb{Y}}(y_1)$ rather than z.

It is worth noting that (2-14) is not a unique representation of our proposed measures. Intuitively, the proposed measures are an α-weighted average of the observed headcounts that correspond to all ordered categories at and below the chosen category that serves as the poverty line. The (2-14) representation of the measures in terms of (normalized) ranks is particularly useful as it assigns a value to every individual observation. This is a computational convenience should the researcher wish to use these measures in regression analysis which I explore in chapter 3 of the thesis.

2.6 Empirical Illustration

We now illustrate the ordinal FGT measures using self-reported health statuses in Canada and the United States from the Joint Canada/United States Survey of Health

(JCUSH). In these surveys, approximately 3,500 Canadian and 5,200 U.S. residents rated their individual health as either *poor, fair, good, very good, or excellent*. Due to the complex sampling design and over-sampling of certain populations, sampling weights have been appended to the survey data by the Centers for Disease Control and Prevention and Statistics Canada to render the samples representative of their respective populations. We use these sampling weights in our subsequent analysis.

Table 2-1: Ordinal FGT

Headcount (α=0) and Ordinal FGT measure (α=1) decomposed by income quintiles. Estimates based on 2,960 and 3,815 Canadian and U.S. respondents from the 2003 Joint Canada/United States Survey of Health.

<table>
<tr><th rowspan="2">Measure</th><th rowspan="2">Country</th><th rowspan="2">Income Quintile</th><th colspan="5">Cut-Off</th></tr>
<tr><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th></tr>
<tr><td rowspan="12">Headcount, $\alpha = 0$</td><td rowspan="6">USA</td><td>All</td><td>0.037</td><td>0.136</td><td>0.398</td><td>0.732</td><td>1.000</td></tr>
<tr><td>1</td><td>0.087</td><td>0.243</td><td>0.532</td><td>0.790</td><td>1.000</td></tr>
<tr><td>2</td><td>0.052</td><td>0.202</td><td>0.497</td><td>0.783</td><td>1.000</td></tr>
<tr><td>3</td><td>0.014</td><td>0.096</td><td>0.398</td><td>0.750</td><td>1.000</td></tr>
<tr><td>4</td><td>0.012</td><td>0.057</td><td>0.288</td><td>0.684</td><td>1.000</td></tr>
<tr><td>5</td><td>0.008</td><td>0.056</td><td>0.247</td><td>0.644</td><td>1.000</td></tr>
<tr><td rowspan="6">CAN</td><td>All</td><td>0.032</td><td>0.111</td><td>0.384</td><td>0.757</td><td>1.000</td></tr>
<tr><td>1</td><td>0.051</td><td>0.154</td><td>0.495</td><td>0.820</td><td>1.000</td></tr>
<tr><td>2</td><td>0.054</td><td>0.170</td><td>0.467</td><td>0.810</td><td>1.000</td></tr>
<tr><td>3</td><td>0.032</td><td>0.104</td><td>0.362</td><td>0.756</td><td>1.000</td></tr>
<tr><td>4</td><td>0.013</td><td>0.084</td><td>0.337</td><td>0.746</td><td>1.000</td></tr>
<tr><td>5</td><td>0.005</td><td>0.042</td><td>0.249</td><td>0.651</td><td>1.000</td></tr>
<tr><td rowspan="12">Ordinal FGT, $\alpha = 1$</td><td rowspan="6">USA</td><td>All</td><td>0.037</td><td>0.087</td><td>0.191</td><td>0.326</td><td>0.461</td></tr>
<tr><td>1</td><td>0.087</td><td>0.165</td><td>0.287</td><td>0.413</td><td>0.530</td></tr>
<tr><td>2</td><td>0.052</td><td>0.127</td><td>0.251</td><td>0.384</td><td>0.507</td></tr>
<tr><td>3</td><td>0.014</td><td>0.055</td><td>0.170</td><td>0.315</td><td>0.452</td></tr>
<tr><td>4</td><td>0.012</td><td>0.034</td><td>0.119</td><td>0.260</td><td>0.409</td></tr>
<tr><td>5</td><td>0.008</td><td>0.032</td><td>0.104</td><td>0.238</td><td>0.391</td></tr>
<tr><td rowspan="6">CAN</td><td>All</td><td>0.032</td><td>0.072</td><td>0.176</td><td>0.321</td><td>0.456</td></tr>
<tr><td>1</td><td>0.051</td><td>0.102</td><td>0.233</td><td>0.380</td><td>0.504</td></tr>
<tr><td>2</td><td>0.054</td><td>0.112</td><td>0.230</td><td>0.375</td><td>0.500</td></tr>
<tr><td>3</td><td>0.032</td><td>0.068</td><td>0.166</td><td>0.314</td><td>0.451</td></tr>
<tr><td>4</td><td>0.013</td><td>0.049</td><td>0.145</td><td>0.295</td><td>0.436</td></tr>
<tr><td>5</td><td>0.005</td><td>0.023</td><td>0.098</td><td>0.236</td><td>0.389</td></tr>
</table>

We apply the ordinal FGT measures to examine health deprivation or health poverty, as well as to examine health poverty when the population is decomposed by income quintiles. We begin by considering the headcount ratios ($\alpha = 0$) in each country and at various cut-offs.[13] As can be seen in Table 1, more U.S. residents as a proportion of the population report their health as being less than or equal to *poor*, *fair*, or *good*, than do in Canada. On the other hand, Canadians are less likely than U.S residents to rate their health status as *excellent*.

For $\alpha = 1$, the ordinal FGT measures suggest that health status in the U.S. is worse than in Canada for every cut-off. Perhaps more interestingly, the decomposition by income quintiles demonstrates that the greatest contribution to the disparity between the two countries occurs at the lowest income quintile. In other words, the disparity in health statuses between the two countries is greatest at the bottom income quintile where the self-reported health statuses of income poor U.S. residents are being compared to self-reported health statuses of income poor Canadians. The $\alpha = 1$ case provides us with more insight into the distribution of the poor than the headcount ratios do by themselves. Such insight may be helpful to policymakers when designing and targeting their health-care policies.

These data can also be used to illustrate the simple interpretation of the ordinal FGT measures provided above. For example, if we focus on the first income quintile and a cut-off of 2, we observe that the FGT measures when $\alpha = 1$ are 0.165 in the U.S. and 0.102 in Canada. Consequently, we have that 165 out of every 1,000 U.S. residents would prefer [14] an equiprobable lottery from the two lowest states of health rather than

[13] A similar analysis using the headcount ratio to look at poverty with self-reported health data was performed in Allison and Foster (2004).
[14] Strictly speaking, this is not a decision-making setting. The agent wouldn't necessarily "prefer" one lottery to another. The description is meant to capture a probabilistic interpretation where she is more likely to be facing better odds for ending up in a higher rank in the two-state lottery than the random draw from the overall distribution.

draw their health status from the actual distribution of health in society. In contrast, only 102 out of every 1,000 Canadian residents would prefer the equiprobable two-state lottery over the random draw from the societal distribution in Canada.

2.7 Concluding Remarks

In this paper, we have considered the issue of poverty measurement in the context of ordinal data. In particular, we have developed a new class of ordinal measures that retains many of the attractive properties of the FGT class of measures (including, for example, additive decomposability) and yet is without the obvious shortcomings inherent in the application of conventional poverty measures to ordinal data. Additionally, we have established a set of dominance conditions, which enable us to give welfare interpretations to our measures, and also allow us to obtain poverty rankings that are robust to the choice of poverty lines.

Applying the ordinal FGT measures to ordinal data on health statuses in Canada and the United States, we find that U.S. residents are health poor relative to Canadians at lower cut-offs when $\alpha = 0$. When α is increased to 1, however, the ranking is robust to the choice of poverty line since U.S. residents are found to be health poor relative to Canadians at all cut-offs. Interestingly, this latter finding suggests that the proportion of residents who would prefer an equiprobable lottery from the five states of health rather than draw their health status from their actual distribution of health in society is greater in the United States than in Canada. A decomposition by income quintiles also demonstrates that the greatest contribution to the disparity in health statuses between the two countries is greatest at the bottom income quintile where the self-reported health statuses of income poor U.S. residents are being compared to self-reported health statuses of income poor Canadians. Overall, our analysis suggests that the ordinal FGT measures for $\alpha \geq 1$ provide us with considerably more insight into the distribution of the poor than the headcount ratios do by themselves.

Chapter 3: What's the Deal with Ordinal Outcome Variables in Evaluations?

3.1 Introduction

Ordinal scales have allowed researchers to include ordinal dimensions in their studies and treat them as an important component of the economic lives and choices of individuals. There has been a growing and long standing interest in evaluations, often designed by economists, that are largely inclusive of non-material aspects of wellbeing. Yet due the qualitative nature of the ordinal outcome variables of interest, -such as empowerment, health, aspirations etc. - they are not measured in dollar units. Rather, they are measured by observations on ordinal scales which do not offer the ease of analysis that cardinal variables do. They are not continuous or differentiable.

Significant challenges exist in the social sciences, education, and economics when dealing with ordinal data. Though we can say that someone is happier one day than another we cannot claim that someone is twice as happy at one point in time than another nor assign a meaningful number to a magnitude of happiness. Yet we can rank one day's happiness to another, for example: "today I am happier than yesterday". Much of the discussion on ordinal data has focused on the interpersonal comparability of respondents, and the validity, reliability and explanatory power of ordinal indicators that measure various constructs (see Strauss, and Thomas (1996) for a comprehensive discussion of the measurement and mismeasurement of social indicators in applied economics). Aside from these issues however, there is an additional concern that arises a posteriori -once the ordinal data have been collected and the aforementioned, commonly discussed, issues have been considered. This thesis focuses on the challenges involved in evaluations, either in the form of regression analysis or an assessment, when outcome variables are ordinal. (see Hansen, Lemke, and Sorensen

(2014) for a discussion of ordinal indicators in educational assessments of personnel evaluations).

The plethora of ordinal psychometric scales in the psychological measurement literature suggests that "soft" dimensions are measurable and it is now standard practice to include some of them in questionnaires and surveys in applied economics. Psychometric scales, in particular, were originally developed in order to be analyzed with rank-correlation-based methodologies, such as factor analysis and principal components analysis. For example, Kahneman, and Krueger (2006) study the measurement of subjective wellbeing and keep their analysis within a rank-correlation setting. Following this trend in economics RCTs, and other evaluations, encompass broader efforts to incorporate non-material variables, such as psychosocial variables, where researchers develop ordered responses for certain indicators they consider to be important to their study. Examples include empowerment indicators, self-reported health, aspirations, and observational data that may be recorded by an assessor. The absence of assurance that distances between ordered categories of responses are well defined may raise concerns in an evaluation setting when the analysis goes beyond the rank-correlation environment and requires a mean-based approach.

An additional area of interest is that of educational assessments. Teacher evaluations, for example, are averages of normalized ordinal scales yet it is not straightforward how to interpret many of these aggregate measures. Hansen, Lemke, and Sorensen (2014) discuss the interpretational difficulties associated with the widely used "profile approach" and they stress that the aggregate score is not easily interpreted and understood by policy makers. Feature normalization is a scale invariant normalization, however the meaning of the rescaled variable is not well defined. Subjective criteria are put forward by policy makers and researchers when choosing an aggregation method that ultimately leads to a single score which then facilitates school rankings and assessments.

In RCTs there are many instances in which a collection of ordinal items in an evaluation is intended to measure a single construct. Aggregating the observations or responses to ordinal items into a single score, so that it is sensitive to variation within all ordered categories, may be desirable in an evaluation setting. For example, observational data for multiple items can be recorded on ordered states such as *unsatisfactory*, *emerging*, *proficient*, *accomplished*, *distinguished*. These need to be aggregated for multiple indicators into a single score in order to evaluate impact on these outcome variables yet the distances between the ordered categories are not well defined. Researchers collecting observational data may not agree on whether the ordered states *accomplished* and *distinguished* are closer than the ordered states *unsatisfactory* and *emerging*.

For the case that researchers wish to include aggregations of multiple ordinal indicators in regressions as outcome variables and for the case that aggregate indicators are used to rank various entities, such as schools, relying on mean-based aggregates, I explore the problems that may arise and then propose one possible way to address them. I examine the average treatment effect, the simplest of evaluations, through a measurement perspective. It is well known that the average treatment effect essentially reduces to the comparison of two means which are estimated by conditional expectations (see Wooldridge (2002)). The comparison of means is also the tool of analysis in educational assessments. Meanwhile, it has been well established in the poverty and inequality measurement literature, in Allison, and Foster (2004), that the comparison of means when using ordinal data can lead to reversals in the mean-rankings. I borrow insights from the poverty and inequality measurement literature, which is a literature that follows an axiomatic approach, in order to understand how the properties of measures can determine whether evaluations are robust to reversals when outcome variables are ordinal.

The binary approach, which is often employed in order to avoid the ambiguity that arises with ordinal data, addresses these concerns and ensures that the scores reside in probability space. This approach has historically been popular among applied economists and ignores the variability within information that may have been recorded on a scale that has more than two categories. In these cases, the aggregation method becomes particularly important.

A key feature of the methodology I propose is that the standard practice of dichotomizing variables on a scale that contains more than two ordered categories is no longer necessary. The measures have an intuitive interpretation and are suitable for reporting the evaluation of impact in understandable terms. The aforementioned concerns may have also become a reason for ordered responses beyond two categories to be avoided all together in questionnaires. This is an additional cost to be considered when developing items that could be best captured with multiple ordered categories.

Responding to these concerns, I propose a class of multidimensional ordinal measures that are particularly suitable for application in the context of evaluation and assessments and can be used as outcome variables in regressions. Some of the measures in this class are sensitive to movement across the full set of information that has been recorded by multiple ordered categories. A particular measure from this class, for example, has an intuitive interpretation where the aggregate score is an average of average cumulative frequencies and has a likelihood interpretation.[15]

3.2 Evaluation reversals

I simulate evaluations in order to explore the magnitude of the issue at hand. Though a demonstration of an evaluation reversal could be performed on an existing RCT or assessment drawn from applied economics or education, this would not allow us to appreciate how big of a problem this can be overall. For meta-analytic purposes it is useful for researchers to have an estimate of the probability with which such reversals could occur. In a similar spirit Bertrand, Duflo, and Mullainathan (2004) use simulations in order to present an appreciation of serial correlation issues in difference-in-difference estimations.

I check the incidence of reversals from simulations that occur with randomly generated non-linear transformations of an ordinal scale[16]. I simulate randomized controlled trials and use randomly generated distributions of ordinal observations. I hypothesize that evaluation reversals occur with a probability that is equal to zero. Simulation results suggest that the incidence of reversals is non-negligible. After simulating $120,000$ distributions the incidence of reversals averaged between 4 and 13 per cent of the time. The choice of a numerical representation of a scale can often be arbitrary as is frequently demonstrated by questionnaire items in evaluations in applied economics and education. The results of the robustness checks that I performed in Matlab (see Appendix A for details) suggested that a non-negligible per cent of evaluations could be

[15] Deciding on the ordered states of an indicator is beyond the scope of this paper. Social scientists devote much of their resources to this particular concern. For example, “no wall at all" is worse than a “dirty wall". But what if the category of “no wall" was not ever included? This is a new category below the existing lowest one and could indeed be filled with observations in theory. But for schools in the US for example, we can probably rest assured that students are taught inside buildings with walls.

[16] Even when cardinal variables are transformed, it is well known by economists that this may lead to evaluation reversals too. When a dollar valued variable is logged, this can lead to a reversal and a change in the statistical significance of results. The difference between doing this for ordinal variables versus cardinal variables is that the original scale is defined accurately in the cardinal case. When transforming ordinal scales we do not know what we are transforming in the first place.

reversed with a simple renumbering of the ordinal scale on which observations are recorded. The results of this exercise are presented in Appendix A.

The underlying reason for this problem has not been well documented, understood and presented in the literature. Economists have historically been reluctant to treat ordinal data as cardinal. In order to understand the prohibitive constraint that ordinal data imply, it is important to make the distinction between regression and correlation. Rank-correlation is a measure of co-movement, whilst regression assumes a dependent variable to be a function of an independent variable and can be reduced to the comparison of estimated means (see Wooldridge (2002)). The implications of misunderstanding this concept are crucial when using ordinal scales. Psychometric scales satisfy ordinal invariance when used in the, intended, rank-correlation analysis environment. Psychometric scales, which can be defended as having distances that are well defined, also satisfy ordinal invariance when used in regression analysis. However, ordinal invariance is not satisfied when we are agnostic of the distances on ordinal scales, particularly when ordered categories were developed for the purposes of an evaluation and when data is observational in assessments.

3.3 Capturing Variability Within an Ordinal Dimension

In this section I present an alternative computational exposition of the class of measures proposed in chapter 2 and then propose a multidimensional variation of these measures that is suitable for regression analysis, and effectively goes beyond poverty measurement. The proposed class is inspired from, and is an extension of, the Alkire, and Foster (2011a) multidimensional class of poverty measures.

Let Y be a random variable that is ordinal and has distribution $(p_1, y_1; \ldots; p_\kappa, y_\kappa)$

where each y_j can fall in the ordered ranks $j = 1, 2, ..., \kappa$, and κ is the number of ordered categories. I present an extension of the Bennett, and Hatzimasoura (2011) unidimensional class of measures as an α weighted average of cumulative frequencies of observations in each category j, assuming the highest category as the poverty line:[17]

$$\pi_\alpha(Y) = \sum_{j=1}^{\kappa} F(y_j)\left[\left(\frac{\kappa - j + 1}{\kappa}\right)^\alpha - \left(\frac{\kappa - j}{\kappa}\right)^\alpha\right] \tag{3-1}$$

For $\alpha > 0$. As was presented in the previous chapter, in the case $\alpha = 0$ the probability mass is one at the poverty line. Parameter α is an aversion parameter which penalizes observations in lower ordered categories in a computationally similar manner to the poverty aversion parameter α of the FGT poverty measures. Intuitively, the proposed unidimensional measure is an α weighted average of cumulative frequencies –or equivalently headcounts. In poverty measurement, a cut-off is a choice of a threshold below which an individual is recorded as poor -or deprived in the multidimensional poverty measurement setting. In an evaluation context the researcher may choose to set a cut-off at the highest category in order to allow for variation within all ordered categories of each item.[18] In the Alkire-Foster multidimensional setting, a poverty cut-off is the number of deprived dimensions for an individual which renders her to be poor. In my presentation of the multidimensional ordinal measures the aggregation takes place over the matrix which has not identified anyone as poor.[19]

---

[17] As in chapter 2, I assume the strong definition of the poor as in Donaldson and Weymark (1986).

[18] Note that the Alkire-Foster methodology also allows for the possibility of setting the deprivation cut-off at the highest ordered category.

[19] Should a researcher choose to include a cut-off below which responses are considered to reflect a deprivation, and a cut-off for the number of deprived dimensions that identify individuals as being poor, the class of measures introduced in this paper lead to the Alkire-Foster measures where each ordinal dimension is measured by the ordinal FGTs introduced in chapter 2. These measures will be multidimensional poverty indices that are sensitive to ordinal depth. These measures can also be calculated if one was to simply use the

Assigning the poverty line at the highest ordered category implies that the number of the poor coincides with the number of all subjects under study. The proposed class of measures accounts for variation within ordered categories depending on the value assigned to α (on a scale of 1 to 5 for example, variability within observations recorded in all five categories will be accounted for when $\alpha > 0$). In the next section I introduce a class of multidimensional measures where the use of α accounts for variation within multiple ordinal dimensions.

3.4 A Proposal for a Multidimensional Class of Measures

Justifying the use of proposed measures often involves relating them to existing measures. For example, the proposed ordinal measures in chapter 2 resembled the already well understood and extensively used FGT measures.[20] Similarly, the proposed multidimensional ordinal measures of this chapter are an extension of the Alkire-Foster class of multidimensional measures. Furthermore, the proposed measures can perhaps be easily used and interpreted by applied economists since the measures have intuitive interpretations and are easy to compute.

It is usually the case that questionnaires include items which contain multiple indicators intended for measuring a single construct. They can then be aggregated in order to provide a "score" for the construct of interest. Examples of such indicators are included in the subsequent chapter where I demonstrate how intrinsic motivation, among other constructs, is measured based on multiple ordinal indicators. Consider for example a

---

classical FGT measures and assign the values $1, 2, 3, ..., \kappa$ to the ordered responses. Though one could obtain measures that are computationally identical to those proposed in this paper, the interpretation of the average of average headcounts would be lost.

[20] Interestingly, the technique that the proposed measures rely on, in order to capture variability within ordered responses beyond two categories, uses a similar normalization approach that underlies the extensively used logit models in applied econometrics. In logit models the independent variables are functions of cumulative frequencies of ordered responses. That is, each cumulative ordered state assumes that the previous one holds.

self-reported health item which may contain multiple indicators, each of which pertains to a separate dimension of overall health. A researcher may wish to aggregate the information into a single score, even though the researcher is agnostic regarding the distances between the ordered categories. An extension of the unidimensional ordinal FGT measures, introduced in the previous section, to the multidimensional case proves to be rather useful here. Another example is a collection of indicators for joint decision making in a household. Given that there are many areas within which household decisions are made, many indicators -or equivalently dimensions- contain information that is collectively useful when it is aggregated in order to compute a single aggregate score.

To extend to the multidimensional setting I apply (3-1) to each dimension. Let $\mathbf{Y}$ be the multidimensional extension of Y from (3-1) with distribution $(\mathbf{p}_1, \mathbf{y}_1; \ldots; \mathbf{p}_D, \mathbf{y}_D)$ where D is the number of dimensions and each y_{jd} can fall in the ordered ranks $j = 1, 2, ..., \kappa$ for a dimension d, and κ is the number of ordered categories as before. The general case for the multidimensional class of measures is:[21]

$$\pi_\alpha(Y) = 1/D \sum_{d=1}^{D} \sum_{j=1}^{\kappa} F(y_{jd}) w^\alpha \qquad \text{(3-2)}$$

where D is the number of dimensions (or equivalently items) and $w^\alpha = (\frac{\kappa - j + 1}{\kappa})^\alpha - (\frac{\kappa - j}{\kappa})^\alpha$. In the setup of the ordinal FGT measures introduced in chapter 2, for $\alpha = 0$ a probability mass of one was assumed at the poverty line and thus the measures coincided with the headcount ratio. Similarly, a probability mass of one is assumed at the deprivation cut-offs and the measures coincide with the

[21] This presentation assumes that each dimension has the same number of ordered categories. This is usually the case in questionnaires when multiple variables capture information about a single construct. However, exploring the possibility of a heterogeneous number of ordered categories could also be insightful.

Alkire-Foster M_0 measure. As the presentation has been setting the deprivation cut-off at the highest category the measure will be equal to 1 whenever $\alpha = 0$. This is because the measure would identify everyone as poor. For $\alpha = 1$ the measure is an average of average cumulative frequencies of observations in each ordered category -or alternatively, if measuring poverty it can be interpreted as an average of average headcounts. For the empirical illustrations in this chapter and chapter 4, I use the ordinal measures for $\alpha = 1$ and choose the highest category for the deprivation cut-offs.

For $\alpha = 1$ an alternative computational representation proves to be useful for regression analysis. Denote x_{id} as the individual observation of the ordinal outcome variable for dimension d, and define $\tilde{x}_{id} \equiv \frac{1}{\kappa}\sum_{j=1}^{\kappa} 1(x_{id} \leq j)$ which can then be used to recover the measure in (3-2):

$$\pi_\alpha(X) = \frac{1}{nD}\sum_{d=1}^{D}\sum_{i=1}^{n}\tilde{x}_{id} \qquad (3\text{-}3)$$

where X is the matrix for all observations, n is the number of individuals, and D is the number of all dimensions.

3.5 Weighting Within and Between Ordinal Dimensions

There is an implicit weighting of the ordered categories within the each ordinal dimension which is parametrically set by α. Intuitively, the proposed measure for $\alpha = 1$ conveys, on average, how many upward single category movements need to be made in order for all individuals to be reporting the highest possible category. When $a = 1$, a movement in observations in lower categories, for example a movement from the 1st to the 2nd category in a dimension, and an equal movement in higher categories, for example a movement from the 4th to the 5th, will increase the overall measure by the same amount.

In the case that $\alpha = 2$, the movement upwards from the $1st$ to the $2nd$ ordered category in a dimension will have a higher impact on the measure than an observation moving from the $4th$ to the $5th$ ordered category. This is because movements in observations towards lower categories “penalize” the overall measure by more than movements in observations towards higher categories.[22]

Deciding on the weighting of each ordinal indicator, which captures one of the many dimensions of a construct, is left to the discretion of the researcher. The reader is referred to Alkire, and Foster (2011a) and Foster, McGillivray, and Seth (2012) for an extensive discussion and analysis of dimensional weighting. Though the weighting of dimensions is a decision left to be made by the researcher, I also note an additional concern that is introduced when extending the Alkire-Foster methodology to being sensitive to variability within ordinal dimensions. In an evaluation setting, the proposed measures imply that an improvement in one ordinal dimension is considered to be equivalent to an equal improvement in another dimension. [23]

The researcher may consider equivalencies across ordinal dimensions to be inappropriate for the research at hand. There do however exist cases where such an equivalency may seem appropriate, such as the application I present in chapter 4 where I use the proposed methodology in order to evaluate the impact of an

---

[22] The measure can be set up in order to capture negative or positive variations depending on direction the construct being measured would suggest.
[23] This holds if we set α to be the same value for all dimensions. An exploration of heterogeneous weighting within and between dimensions could be an exercise for further research.

intervention on ordinal multidimensional psychosocial aggregate scores in a RCT setting.

3.6 Properties of the Measures as Outcome Variables

For $\alpha = 1$, (3-3) is a convenient representation of the proposed measure for a regression analysis context, since for each observation of the outcome variable we can use $\tilde{x}_{id} = \frac{1}{\kappa}\sum_{j=1}^{\kappa} 1(x_{id} \le j)$ from (3-3). Drawing from (3-3), X is the matrix of observations for all dimensions of the ordinal outcome variables of interest. In a randomized evaluation setting, the causal parameter β_{OLS} is estimated as the difference (Newey (2007)):

$$\beta_{OLS} = \hat{\mathbf{E}}_X(\pi_1(X) \mid T) - \hat{\mathbf{E}}_X(\pi_1(X) \mid C) \quad (3\text{-}4)$$

where T denotes the treatment group, C denotes the control group, and $\hat{\mathbf{E}}_X$ is the sample average over X.

I discuss the properties satisfied by the proposed ordinal multidimensional measures in (3-3). Each property is discussed in the context of randomized evaluation which can be extended to the general case of evaluations. Properties satisfied by $\pi_\alpha(X)$ include ordinal invariance, monotonicity, transfer, decomposability, subgroup consistency, non-triviality, symmetry.

Ordinal Invariance: If $g : \Re \to \Re$ is monotone increasing then $\pi_\alpha(X) = \pi_\alpha(g(X))$, and g is applied to each individual dimension.

In order to achieve evaluations that are robust to reversals which can arise due to a positive monotonic re-scaling of the ordered categories, this property ensures that such reversals will not occur no matter what the choice of scale is. This follows from the fact that the measures are built from the *proportions* of observations in each ordered category and are not a function of the numerical representation of the ordered categories which would pertain to analysis that treats ordinal data as cardinal. This property assures us that the estimated expectations are independent of the choice of the ordinal scale. Notably, the measures rely on the order and number of categories which avoids any ambiguity due to a choice of a particular scale.

Monotonicity: Let X be a matrix containing ordinal observations from a chosen collection of indicators that has a distribution matrix $(\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_D)$ [24] where each $\mathbf{p}_d$ for $d = 1, 2, ..., D$ is a distribution vector for an indicator in dimension d. If $\widetilde{X}$ has the same distribution matrix as X, denoted by $(\tilde{\mathbf{p}}_1, \tilde{\mathbf{p}}_2, ..., \tilde{\mathbf{p}}_D)$, but for some $\tilde{\mathbf{p}}_d$ probability mass is re-allocated so that $\widetilde{p}_k = p_k + \varepsilon$ and $\widetilde{p}_l = p_l - \varepsilon$ for some $k < l$, then $\pi_\alpha(\widetilde{X}) > \pi_\alpha(X)$ for $\alpha > 0$. [25]

This property assures us that the measures are sensitive to a particular form of variability which occurs when there is a net movement upwards or downwards for observations in the ordered categories.[26]

Transfer: Suppose $(\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_D)$ is the distribution matrix for X, a matrix containing ordinal observations from a chosen collection of indicators where each $\mathbf{p}_i$, for

[24] It is important to note that this paper does not consider the joint distribution of ordered responses across dimensions as is the case in Alkire-Foster.
[25] If a deprivation cut-off is chosen, then the Alkire-Foster monotonicity properties follow which are weak monotonicity and dimensional monotonicity.
[26] Net movement implies that it is not cancelled out by another equal movement. The exposition of this property is similar to that in Allison and Foster 2004.

$i = 1, 2, ..., D$, is a distribution vector for indicator d. If $\widetilde{X}$ is generated from X, so that is has a distribution matrix $(\tilde{\mathbf{p}}_1, \tilde{\mathbf{p}}_2, ..., \tilde{\mathbf{p}}_D)$, where each $\tilde{\mathbf{p}}_d$ is a vector $(\tilde{p}_{\mathbf{1}}, \tilde{p}_2, \ldots, \tilde{p}_\kappa)$ and probability mass is re-allocated so that $\widetilde{p}_k = p_k + \varepsilon$ is offset by $\widetilde{p}_{k-l} = p_{k-l} - \varepsilon$ and $\widetilde{p}_m = p_m - \varepsilon$ is offset by $\widetilde{p}_{m+l} = p_{m+l} + \varepsilon$ for some $k < m$, then $\pi_\alpha(\widetilde{X}) > \pi_\alpha(X)$ for $\alpha > 1$.[27]

In randomized evaluations, the transfer property allows the researcher to be in control of the weighting of movement within ordinal categories. For example, in statistics the standard deviation for a cardinal variable uses a weighting structure that increasingly "penalizes" deviations by squaring them. In the ordinal context, the value of α determines the degree to which observations in lower categories are "penalized". The parameter α may be used for additional robustness tests in order to explore whether a reversal occurs as α varies.[28]

Decomposability: Let λ reside between zero and one. Suppose that matrix X_1 contains self-reported responses from a collection of indicators for a treated subgroup of a population $\frac{n_1}{n}$ and X_2 is a matrix for the control subgroup with population $\frac{n_2}{n}$. Each matrix is has a distribution matrix P_{X_1} and P_{X_2} respectively. Also, suppose that matrix X has distribution matrix $P = \lambda(P_{X_1}) + (1-\lambda)P_{X_2}$,where $\lambda = \frac{n_1}{n}$. Then $\pi_\alpha(X) = \lambda\pi_\alpha(X_1) + (1-\lambda)\pi_\alpha(X_2)$.[29]

---

[28] It must be noted that the transfer axiom is fundamentally a cardinal property. Yet, because the measures satisfy ordinal invariance they do not treat the scales as if they are cardinal.
[29] In a poverty measurement setting, the variability within each dimension, post identification, will be

Rigorous analysis in an evaluation often requires an exploration of heterogeneous effects on outcome variables. This presupposes that the estimates that are to be used are decomposable. Since the overall estimate can be decomposed to the estimates for each subgroup, the proposed measures allow for heterogeneity analysis of the effects on the outcome variable. This allows one to infer about the impact on the measure of any subgroup that we choose. This property further leads us to subgroup consistency:

Subgroup consistency: In addition to the decomposability property, suppose that matrix Y_1 contains ordered observations from a collection of indicators for a particular the treated subgroup and Y_2 is a matrix for the control subgroup. Each matrix has a distribution matrix Q_{X_1} and Q_{X_2} respectively. Also, suppose that matrix Y has distribution matrix $P = \lambda P_{Y_1} + (1-\lambda)P_{Y_2}$. Then $\pi_\alpha(X) < \pi_\alpha(Y)$ whenever $\pi_\alpha(X_1) < \pi_\alpha(Y_1)$ and $\pi_\alpha(X_2) < \pi_\alpha(Y_2)$.

Non-triviality: Each measure of the proposed class has a maximum and a minimum. *Normalization* is not satisfied by this class of measures. The minimum is $\frac{1}{nkD}$ and the maximum is 1.

Symmetry and *Replication* are satisfied by virtue of the linear expectations operator.

3.7 Empirical Illustration

To demonstrate the proposed measures in a randomized evaluation setting I make use of an existing randomized controlled trial in Udaipur India for which the data is publicly

captured for the categories that are less than the category chosen to be the deprivation cut-off.

available at JPAL[30]. The experiment was designed in order to evaluate non-financial incentives compared to improving the health service delivery for immunizations for children between the ages 1 to 3 in Udaipur, India. The experiment is described in detail and evaluated in Banerjee, Duflo, Glennester, and Kothari (2010). The ordinal measures presented here use one dimension of self-reported health and are estimated in terms of achievement where a higher aggregate score indicates higher self-reported health. The measures in (3-3) are reported for $\alpha = 1$.

The randomization took place at the village level with 134 villages in total being involved in the experiment. 60 villages were in the treatment group and 74 villages were in the control group. In 30 out of the treated villages, 378 children were targeted for intervention A which offered more reliable health service delivery for immunizations. In the rest of the 30 treated villages, 382 children were offered metal plates and lentils for completed immunizations. The control group targeted 860 children who were offered no intervention.

Table 3-1: Ordinal Measures for Self-reported Health

	Control Group	Treatment A	Treatment B	Difference between control and treatment A	Difference between control and treatment B	Difference between treatment A and treatment B
Self-reported Health	0.455443	0.447472	0.432493	-0.00916	-0.01691	-0.0082364
Standard Deviation	0.01062	0.013482	0.016966	0.020733	0.019057	0.0238803

In this illustration I focus on self-reported health which was measured on a scale of 1 to 10. I use the proposed measures in order to evaluate the impact of each intervention

[30] Abdul Latif Jameel Poverty Action Lab, MIT.

on the self-reported health of parents for their children. I use the same specification as in Banerjee, Duflo, Glennester, and Kothari (2010) and make use of the full variability across all 10 categories used to capture self-reported health. Table 3-1 presents the ordinal measures for self-reported health. The control group appears to have higher self-reported health but is only marginally higher. It is also not statistically significant. Treatment A appears to have led to lower self-reported health, and treatment B seems to have led to even lower self-reported health.

For those who were fully immunized, the results are presented in Table 3-2. I am able to present the ordinal measures for self-reported health for this subgroup since the measures are decomposable. However, it is important to note that the sample size is even smaller once the population is conditioned on full immunization (only 253 children).

Table 3-2: Ordinal Measures for those Fully Immunized

	Control Group	Treatment A	Treatment B	Difference between control and treatment A	Difference between control and treatment B	Difference between treatment A and treatment B
Self-reported Health Completely Immunized	0.46875	0.421539	0.411029	-0.06392	-0.06049	0.020031
Standard deviation	0.04073	0.028277	0.021266	0.062558	0.049707	0.04906

According to the ordinal measures, conditioned on being fully immunized, those in treatment are reporting lower health statuses. And those in treatment B are reporting higher health statuses than those in treatment A. Those with no incentives and no improved service delivery tend to report higher self-reported health status for their children. An interpretation of such results is beyond the scope of the paper and this exercise simply serves as a demonstration for how the proposed measures may be used.

Below are the values for likelihoods at different cut-offs which are followed by the (3-3) ordinal measures. Had the researcher chosen the cut-off 8 for example, the program evaluation would be reversed compared to all other cut-offs. This would not be a problem if the researcher had good reason to choose a binary approach and such a cut-off. The proposed ordinal measures are suitable for the case that the researcher would have liked to report results that are sensitive to net movement within all ordered categories, whilst refraining from using the ordinal representation of the scale as if it is cardinal.

Table 3-3: Ordinal versus Binary Measures

Variable Description	Control	Treatment A	Treatment B
Likelihood at cut-off 1	0.007595	0.008427	0.002801
Standard errors	0.003419	0.004632	0.00274
Likelihood at cut-off 2	0.032911	0.02809	0.022409
Standard errors	0.006465	0.009005	0.007891
Likelihood at cut-off 3	0.096203	0.08427	0.061625
Standard errors	0.012616	0.017771	0.011427
Likelihood at cut-off 4	0.187342	0.191011	0.154062
Standard errors	0.017442	0.02472	0.023195
Likelihood at cut-off 5	0.326582	0.323034	0.277311
Standard errors	0.020329	0.029026	0.027354
Likelihood at cut-off 6	0.494937	0.463483	0.448179
Standard errors	0.024286	0.03185	0.03744
Likelihood at cut-off 7	0.646835	0.620787	0.607843
Standard errors	0.025415	0.02473	0.04438
Likelihood at cut-off 8	0.812658	0.83427	0.809524
Standard errors	0.018532	0.021787	0.039904
Likelihood at cut-off 9	0.949367	0.921348	0.941177

Standard errors	0.0101	0.018525	0.021427
Ordinal Measure	0.455443	0.447472	0.432493
Standard errors	0.01062	0.013482	0.016966

3.8 Concluding remarks

The inclusion of ordinal scales in evaluations has allowed researchers to evaluate impact on both unidimensional and multidimensional ordinal outcome variables such as social and economic attitudes, perceptions, self-reported health, and educational outcomes. However, in order to avoid ambiguous evaluations, aggregating information from observations on ordinal scales into an outcome variable requires particular care. In this chapter I explored some of the problems inherent in this practice and proposed one possible way to address the robustness issues introduced when such aggregations are used as outcome variables in evaluations. The proposed measures can be interpreted in probability terms. Rather than reporting a difference in cardinal units that is subject to a particular scale of choice, one can make a likelihood statement for the effect on observations in higher -or lower, depending on the direction of improvement- categories. The resulting likelihood is sensitive to a change within all ordered observations collected, and not within a subset of the observables which would have been the case had a binary approach been chosen.

Chapter 4: The Aspirations Gap and Women's Bargaining Empowerment: Evidence from a Randomized Controlled Trial in Nicaragua

4.1 Introduction

In this chapter I present an applied framework for a randomized controlled trial in Nicaragua and demonstrate the methodology introduced in the previous chapter. I evaluate the impact of productive transfers targeted at women in 24 communities in Pantasma, Nicaragua. I take advantage of the intervention's experimental design as well as the availability psychosocial ordinal outcome variables in order to explore the determinants of women's intrahousehold bargaining empowerment.

The literature on empowerment is somewhat inconsistent in its definition of "empowerment." One possible reason for this is because the context within which the construct of empowerment resides is not always clearly defined. This may not be so surprising, since disempowerment can manifest in many areas of one's life and theoretical discussions may lack behavioral and contextual justification. In the efficient household model literature, Pareto-weights have been referred to as the "power" of each individual household member, and they have also been referred to as "bargaining power." The determination of these weights is a process that is largely ignored in the empirical literature due to the lack of suitable data. I refer to the Pareto-weights, in absolute value, as bargaining power. When the intervention causes an increase in bargaining power, I refer to this dynamic process as bargaining empowerment.

In order to estimate the impact of the intervention on ordinal outcome variables, I use the methodology introduced in chapter 3 which allows me to make full use of the variation within the self-reported responses on ordinal scales. This ensures that the estimates of the treatment effect are robust to evaluation reversals as was discussed in chapter 3. Multidimensional ordinal outcome variables include empowerment, locus of control, self-esteem, intrinsic motivation and intrahousehold decision-making among others.

It is widely accepted through empirical justification that many of the assumptions of the efficient (unitary) household model do not hold in practice (See Bardhan, and Udry (1990) , Chiappori (1997), and Duflo, and Udry (2004)). Since the household is treated as a single decision maker, one key insight that the efficient household model ignores is the process through which intrahousehold bargaining equilibria are set. Adding to the variety of hypotheses of the literature that tests the efficient household model I study the role of aspirations in the intrahousehold bargaining process. One theoretical exploration of the process has been introduced by Basu (2006) where in a game theoretical framework one's decision will affect future decisions made thus bringing new insight to the importance of *process* in decision making. [31]

4.2 Introducing the Experiment

The program was implemented in rural communities in the municipality of Santa Maria de Pantasma, located in the department of Jinotega in the Northern region of Nicaragua. Santa Maria de Pantasma suffers from one of the highest incidences of poverty in Nicaragua. Twenty four communities were randomly assigned to treatment and control.[32] The intervention provided a package of productive transfers at low cost and technical assistance to selected poor women. These packages included business plan development, training and technical assistance, productive transfers and low cost credit (in cash and kind), training on gender issues awareness and the creation of a

[31] The applied framework I introduce relies on the design of the experiment.
[32] This pilot is run by the World Bank with principal investigators Renos Vakis and Patrick Premand. I became involved in the project after the follow-up survey was completed and was responsible for all program evaluation estimates.

beneficiary-led community bank (in cash and seeds).

The criteria for targeting a potential beneficiary in a household were to have participated in activities in the community and to be between the age of 16 and 60. Potential beneficiaries were invited to consider joining the program. In all eligible communities the potential beneficiaries were asked to enroll in an intent-to-participate list. They knew that they would only be able to participate in the program if their community was selected. Out of the 1956 households that contained potential beneficiaries in the 24 target communities, 877 (45%) enrolled and 980 (50%) did not enroll. The remaining 99 households (5%) were not eligible as they did not have a woman between the ages of 16 and 60 who could serve as a beneficiary.

Target communities where divided in 10 blocks of 2 or 3 neighboring communities. Each community (or pair of communities) was selected within each group (one allocated as treatment and the other as control) through a lottery. This block randomization procedure between neighboring communities maximized power given that the number of clusters to randomize was low (Bruhn and McKenzie, 2008). Block randomization led to the selection of 13 treatment communities and 11 control communities. Baseline census data was collected in all households in the eligible communities before the partner NGO began to implement the program.

The experiment was implemented through a collaboration of a local Nicaraguan NGO, FUMDEC, who anticipated that the intervention would lead to impacts consistent to what they referred to as the "pyramid" model. The NGO staff believed that economic empowerment constituted the basis of the pyramid which they viewed as a necessary condition in order for an increase in women's agency to occur. The productive transfers aimed to increase the beneficiaries' income generating capacity. The program aimed to enhance women's productivity in existing economic activities and foster entrepreneurial development. It promoted gender empowerment, and encouraged their participation in economic decisions within their households and their communities. The training and technical assistance components of the program focused on business plan development, business management, and financial literacy.

The follow-up survey took place between July and December 2011, approximately two years after the baseline survey, and approximately one year after the end of the program. The follow-up instrument aimed at capturing a broader set of outcomes, thus allowing a thorough evaluation of the program's impact, including economic activities and employment (such as involvement in non-subsistence agriculture, non-agricultural activities, micro-business, assets ownership, savings and credit), welfare outcomes such as consumption, poverty or income, and empowerment (such as intra-household decision-making, and aspirations).

Beneficiaries also participated in various workshops on social capital formation and leadership, gender awareness, power relations and conflict resolution. The gender awareness training also included awareness training for men. The beneficiaries received start-up capital in both cash and kind. Part of this capital was in the form of low cost credit to be repaid to community banks. Credit was offered at zero interest with the exception of a $200 micro business transfer of which half was in the form of credit at 3% monthly interest rate over the remaining balance. Additionally, seeds were repaid to the community seed bank at a rate of two seeds for each seed given. Beneficiaries also received training and technical assistance for business plan development.

The randomization's number of clusters was indeed small. However, balance between the control and treatment groups was attained. Overall, there are no notable differences between control and treatment communities. For the few significant differences, treatment communities appear slightly worse off than control communities in those particular dimensions. Consequently, the estimates of program impacts are lower-bound, conservative estimates.

Program impacts can be estimated using simple intent-to-treat procedures by comparing ex-post outcomes between beneficiaries in treatment communities (TSI) to those of women who selected into the program in control communities (CSI) who did not receive treatment due to the randomization. The average difference in outcomes between these two groups yields the average treatment effect on the treated, i.e. the direct effect of the intervention on similar households that expressed intention to participate in the program.

Conditional on the randomization achieving balance, program impacts could be estimated using simple intent-to-treat procedures. This entails comparing ex-post outcomes between beneficiaries in treatment communities to those of women who selected into the program in control communities. The ITT estimate captures how random assignment to the treatment group affects the outcome of interest and provides an estimate of the programs average effect on participants who had selected in:

$$Y = \alpha + \delta T_1 + X + C + \varepsilon \tag{4-1}$$

where Y is the outcome of interest, T_1 is an indicator function that takes on the value 1 if a household revealed an intention to participate in the program and was assigned to the treatment group, and 0 if it revealed an intention to participate in the program but was assigned to the control group. X is a fixed effect controlling for the variation within the blocks of communities on which randomization was based. The intent-to-treat estimate for program impact is given by δ. C is a set of controls which are drawn from the baseline study. Block fixed effects were used in order to maximize power. The block randomization of the design allows us to control for outcomes that are a function of pre-program characteristics that existed in the blocks – which could also affect the outcome of interest. In the estimations standard errors are clustered at the community level.

Program impacts are obtained by estimating the above specification using only the follow-up survey round. This allows the estimation of causal effects from items that were incorporated only in the follow-up survey. As mentioned above, the validity of the randomization design rendered single difference approaches to be appropriate, valid, and sufficient to infer a causal impact of the program. The attrition rate for beneficiaries was 5.2% which is balanced across the communities that received the treatment and those that did not. To ensure further robustness of the results, control variables were included from the baseline survey. These control variables were chosen when minor imbalances between treatment and control group at baseline existed. Though the power of the study was low, it was not prohibitively low. Power calculations estimated that the sample was large enough to identify moderate effects.

4.3 An Applied Framework

By design of the experiment, the beneficiaries self selected into the randomization process. Therefore, by construction this allows me to compare two similar groups of communities with the same aspirational/preference levels. Potential beneficiaries who selected into the program were better-off than those who selected out. Potential beneficiaries who selected in tended to be more empowered than potential beneficiaries who selected out. Admittedly, it was the most empowered women who self-selected into the program. The descriptive statistics at the baseline level suggested some significant differences between those who self-selected in the program and those who did not. It is worth noting that many interventions are demand driven yet the implications of such targeting could suggest that demand-driven targeting does not necessarily reach the poor.

I consider those selected in as not “failing to aspire” (the definition of aspirations failure will be presented in the last section along with a justification for the aforementioned assumption). I test for multiple equilibria, controlling for high aspirations, in which women may exercise different levels of intrahousehold bargaining power. I test whether it is income generating capacity constraints that are driving an aspirations-empowerment trap. I also test the hypothesis conditional on constant household income, in order to establish whether a simple re-allocation of income generating capacity can lead to bargaining empowerment. The results are somewhat suggestive that aspirations alone do not bring about change, validating the implementing NGO’s strong priors for a “pyramid model”. I have direct information for each household member’s share of non-food consumption and thus have available data on the intrahousehold allocation of consumption goods.

The study does not allow me to explore the process through which aspirational levels can be influenced. For those selected out, whose entrepreneurial aspirational levels are assumed to be low by experimental design, I can test for spillover effects. The findings suggest that bargaining power can only be enhanced once high aspirational levels are in place (the process of which this paper does not explore, but is nonetheless important).

In the efficient household (unitary) model the objective function that is maximized is the overall household welfare. I am not able to identify whether women's bargaining empowerment is a Pareto improvement for the whole household. More collaborative and participatory environments possibly yield higher household utility. Indeed, in my findings, participatory decision making increased which is suggestive of a more collaborative and participatory environment. I cannot, however, identify whether overall household welfare increased (men were not asked to report on any wellbeing scale).

I present the standard unitary household model where the household maximizes the following objective function where I consider the household as being comprised of the husband and wife, as in Duflo, and Udry (2004):

$$\mathbf{Max}[EU^M(c_M) + \lambda EU^F(c_F)] \tag{4-2}$$

where λ is the Pareto-weight representing each household member's "power", $EU^F(c_F)$ is the expected utility for the wife's consumption and $EU^M(c_M)$ is the expected utility of the husband's consumption.

Subject to:

$$p(c_F + c_M) + w(L_F + L_M) \leq F_F(L_F) + F_M(L_M) \tag{4-3}$$

where L_F is labor supplied by the lead female household member and L_M is labor supplied by their husbands. $F_F(L_F)$ is the wife's production function of labor and $F_M(L_M)$ is the husband's production function of labor, p and w are price and wage determined in a competitive market.

I assume the intervention enters the efficient allocation of resources only through its effect on the beneficiary's labor supply and then on the budget constraint. Another assumption is that I look at production on the farm and in the market collectively even though my evaluation of the intervention suggested that there was a re-allocation of employment from the farm to the market (changes were noticed for both beneficiaries and their partners).

According to the efficient household model, the change in the beneficiary's labor supply should not have an impact on the Pareto weight since her preferences remain constant (they have already selected in based on the experimental design). Relaxing the beneficiary's income-generating capacity should not have an effect on the re-allocation of consumption since her preferences are taken into account when bargaining takes place and each member falls back into the efficient context Rangel, and Thomas (2012).

In my setup I propose to explore the difference between *having aspirations* and *meeting aspirations*. The efficient household model assumes that this distinction should not matter. According to the unitary model λ is not a function of labor but is a result of the household members' preferences and other distributional factors which implies that:

$$\frac{\partial \lambda}{\partial L_F} = \frac{\partial \lambda}{\partial c_i} = 0. \tag{4-4}$$

where L_F represents employment opportunity for the female household head. The intervention targeted women who had selected in, denoted by TSI (Treated Selected In). Those in the treatment communities who indicated that they would like to receive the intervention packages are assumed to have equally high aspirations with those in the control communities who also stated that they would like to receive the packages. I consider their intent to participate in the intervention a high-aspirational attitude. I distinguish this attitude from the behavior of actually working towards meeting their entrepreneurial aspirations by being randomly selected to receive the packages. This leads to the following hypothesis which essentially makes the distinction between attitude and behavior:

$$\mathbf{E}(\lambda^T - \lambda^C \mid TSI) = 0 \tag{4-5}$$

4.4 The Hypothesized Aspirations Trap

Debraj Ray introduced the concept of "failure to aspire" in Ray (2006) where he defines one's cognitive neighborhood of acquaintances as an "aspirations window." An individual fails to aspire when examples do not exist within the decision maker's aspirations window. In this paper I deviate from his definition of aspirations failure while staying in line with Ray's definition which carefully associates this "failure" to the condition of poverty and not some inherent behavior of the poor.

In order to avoid "accusing" the poor of "failing" to behave in a manner that could potentially lift them out of poverty, I propose a definition for aspirations failure in the context of this study. I define "failure to aspire" as the extent to which opportunity sets are smaller under subjective beliefs than under objective frequencies. In other words, I consider a person as failing to aspire if they reject a *seemingly* available opportunity (in our case the available opportunity are the packages offered). This study does not explore the reasons or the process through which an offered opportunity is rejected. Rather, it separates people into two groups. Those who fail to aspire and those who do not. Those selected into the program are assumed to have higher entrepreneurial aspirations than those who selected out. Specifying the dimension of aspirations is rather straightforward as the packages offered were of entrepreneurial nature and offered the opportunity to build small businesses and enhance entrepreneurial skills.

Ray's "aspirations gap" measures a stated aspiration's distance from a realized aspirational target. I assume that the packages offered to the beneficiaries at least partially fill their entrepreneurial aspirations gap, $G^A = (\frac{A_H - A_R}{A_H})$ where A_H is the high level of entrepreneurial aspirations and A_R is the realized level of entrepreneurial aspirations. I do not have a measure of aspirational levels at baseline, so I build an index for intrinsic motivation based on the dimensions of the aspirations scale in self-determination theory in Deci, and Ryan (2000) and the ordinal multidimensional methodology introduced in chapter 3. I find no difference at follow up which further validates the conditionality of equal aspirational levels.

I begin by proposing that bargaining empowerment is a function of the aspirations gap. I assume that λ, bargaining power, is a decreasing function of the aspirations gap, G^A:

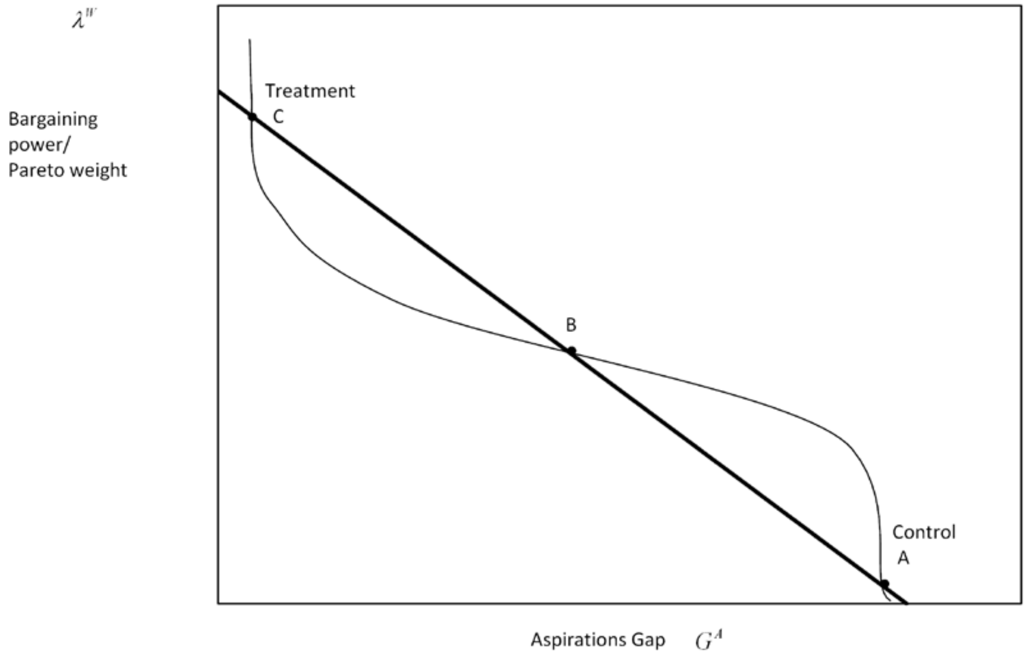
$$\lambda = \lambda(G^A) = \lambda\left(\frac{A_H - A_R}{A_H}\right) \tag{4-6}$$

where A_H is the high level of entrepreneurial aspirations and A_R is the realized level of entrepreneurial aspirations.

Intuitively, as the aspirations-gap closes the more the intrahousehold bargaining power of the beneficiary. Bargaining power is a function of *realized attitude/preference* (actual behavior) rather than *stated attitude/preference* (attitude alone), so in effect it is a function of utilized resources for implementing one's attitudes/aspirations.

One may for example have high aspirations and wish to have greater bargaining power in the household but not have the resources or the income generating capacity in order to support it. However, the closer one gets to implementing their aspirations, the more bargaining power they gain in the household. Given that the aspirational target was already high (women indeed did want to delve into entrepreneurial activity) they got closer to meeting their aspirations and this was subsequently reflected in their share of non-food consumption which I treat as a proxy for bargaining power.

Figure 4-1: Aspirations-Bargaining Empowerment Figure



At equilibrium C, the aspirations gap is smaller and is associated with high bargaining power in the household. At equilibrium A, women do have high entrepreneurial aspirations, but they do not have the resources to implement them and their entrepreneurial aspirations gap remains large, rendering their intrahousehold bargaining power to be low.

In the experimental setup, the stated preference (aspirational target) is kept constant and it is the distance to the target that drives the variability in the gap. I note that the process of determining λ is only investigated to the point that it is influenced by the aspirations gap G^{A}. The Pareto-weight (bargaining power) is a function of various distributional factors and preferences (observed and unobserved), Λ, and the aspirations Gap, G^{A}:

$$\lambda = \lambda(G^A, \Lambda). \tag{4-7}$$

All other factors contained in the vector Λ are not investigated here and are assumed to be constant thanks to the randomized experimental setup which supports this two-dimensional graphical representation.

The Hypothesis for Multiple Equilibria: I hypothesize an aspirations-gap trap in which there are two possible equilibria in a high-aspirations environment. Conditional on household income remaining the same, and treatment applied to those selected in -those with high aspirations- according to the efficient household model the change in bargaining power should be zero:

$$\mathbf{E}(\lambda^T - \lambda^C \mid TSI, \Delta\pi^*_{HH} = 0) = 0. \tag{4-8}$$

I find that overall household income did not change, even though income generated by the beneficiaries alone did indeed increase. In conjunction with the finding of bargaining empowerment for women, the findings are suggestive of the aforementioned aspirations-empowerment trap.

The presentation of the hypothetical aspirations trap in Figure 4-1 is reminiscent of the Basu (1999) child-labor trap model as it is presented in Todaro, and Smith (2014). I cannot identify a “trap” in the traditional sense, or trace the standard S-curve. I am however able to state that since we control for high entrepreneurial aspirational levels, based on the experimental design, we can explore the variability in bargaining power levels. Additionally, if the overall household income remains constant the findings become even more suggestive of the existence of multiple equilibria.

4.5 Empirical Results

As was discussed earlier on, this was a pilot experiment with a statistical power that was not prohibitively low. This suggested that in order for any impact to be captured, it had to be large enough. The reason for the low power was that randomization was performed at the community level. Randomizing at the community level protected the experiment from various sources of contamination which would have existed had the experiment been randomized at the household level. These communities are tightly knit, so there were plausible concerns for large spillover effects.

The results include fixed effects at the community level (in line with the design of the experiment) and errors are clustered at the community level. The specification follows Bruhn, and McKenzie (2008) so that inference accounts for the method of randomization. Based on the hypotheses derived from the framework, I find that after holding aspirational preferences constant, relaxing income generating constraints for women leads to higher intrahousehold bargaining power levels or, equivalently, it leads to intrahousehold bargaining empowerment for women Duflo (2012).

These results hold true even though overall household income did not increase. For non-food items, the survey provided information about the shares of members in the household. This information is treated as a proxy for within-household allocation of consumption and bargaining power. Table B-11 presents the results for the intra-household allocation of non-food items. Overall, beneficiaries' average share of non-food items has increased by 5.1%, while their partners' share decreased by 5.3%. The rest of the household members see no change in their share of non-food items. Since the overall value of non-food items is not higher in beneficiary households it is reasonable to expect that the beneficiaries' partners are experiencing a decrease in the consumption of non-food items. There is a change in the distribution of the same amount of non-food items among beneficiaries and their partners, with beneficiaries now enjoying a bigger share of the bundle of non-food items. Considering that the beneficiaries' incomes increased yet their husbands' did not, the increase in welfare may have been mostly enjoyed by the beneficiary in the form of improved bargaining power within the household.

4.5.1 Impact Evaluation for Ordinal Multidimensional Indices

Table B-14 presents program impacts on ordinal multidimensional aggregate indices, introduced in chapter 3, on intra-household decision-making. These correspond to the share of household decisions made solely by the beneficiary, made jointly by the husband and the women beneficiary, and made solely by the husband. In the literature there is no consensus as to whether a woman's sole decision-making or the couple's joint decision-making are decision processes that empower the woman. Hence, both are reported. In either case, the results show that the intervention appears to have had an impact on participatory decision-making. Beneficiaries are more likely to engage in joint decision-making in most dimensions (the survey asks about who makes decisions on various issues such as food, childcare, investments etc.). They are also simultaneously less likely to be the sole decision makers. Overall, 8.3% of the time, beneficiary women are more likely to be involved in joint decision making with their partners compared to those in the control group. Their partners are less likely to be sole decision makers 4.9% of the time.

Table B-12 presents program impact on an ordinal multidimensional relationship index which is based on a set of questions pertaining to the beneficiary's perception of her interactions with her husband. The aggregated index indicates that beneficiaries are 4.4% more likely to be engaged in more fulfilling and empowering relationships with their spouse. The impact is significant on all individual questions as well, including more sensitive questions such as whether the beneficiaries and her spouse kiss. Perhaps the participatory decision making process for the change in λ is suggestive of a pareto improvement. I am not able to identify, beyond income, whether overall household welfare stayed the same. It would be interesting to explore whether participatory environments have greater aggregate utility -we never asked the men if they are happier.

Table B-12 presents program impacts on a range of agency proxy indicators, including self-esteem, mood and intrinsic motivation. These ordinal multidimensional outcome variables were compiled using the methodology introduced in chapter 3. Following Deci, and Ryan (2000)'s theory of self-determination and the aspirations scale they have developed, the three dimensions of autonomy, competence and relatedness are the indicators that were measured in the follow-up survey. Overall, the results reveal very small impacts for the main beneficiary. For example, the self-esteem scale was heuristically validated in the field and only the dimensions that were relevant to the Nicaraguan context were kept. The beneficiaries report higher self-esteem levels but the result is not statistically significant. Similarly, the indicators that are included in the scale for mood are aggregated providing the total number of days spent on negative and positive realizations of ones emotional state. Beneficiaries report and extra day of feeling more positive and less negative, respectively. However these differences are not statistically significant. The only indicator in which beneficiaries scored significantly higher in this area is that of locus of control. However, the magnitude of this difference is 1.4% which is very small. As discussed, one explanation may be that the women that had selected into the program already had high levels of agency and as such, the intervention did not have a large effect on this due to this.

4.5.2 Further Impact Evaluation

Table B-3 presents program impacts on a range of employment outcomes for the main beneficiary. The first set of indicators relates to employment in the last 12 months: self-employment in agricultural activities, self-employment in non-agricultural activities, employment in agricultural wage jobs and employment in non-agricultural wage jobs. Results show very limited changes in overall employment portfolios for beneficiaries at the extensive margin. One notable exception is commerce, where women in treatment communities are almost six percentage points more likely to engage in this (an increase of more than 40% percent relative to the control group). There are, however, some differences along the intensive margin, and in the composition of employment. In addition, the program led to some impact on small household businesses. Beneficiaries are 5.4 percent more likely to sell non home-made items and 5.6 percent more likely to own a business. Of those who report that they have a business, 8.5 percent are more likely to report they expect to have their business in the future. Overall, these results are consistent with results above on changes in employment activities along the intensive margins.

Table B-1 presents program impacts on beneficiary income. The results show that beneficiary income increased by 39.6%. In absolute value, the beneficiary income increased by $1830 a year, or approximately $72. This increase in income is driven by large increases in income from livestock, (48.1%). Income from backyard agricultural production and commerce are not significant but positive (and large) relative to the control group. Program benefit packages included seeds to diversify agricultural production in households backyard, which is traditionally cultivated by women, as well as asset transfers including livestock such as goats or chicken. These results suggest that the changes in employment patterns documented above led to some productivity gains for the beneficiaries. These shifts are mostly observed through changes in occupations within agricultural activities (the lack of significant results for commerce and backyard production can also be due to the limited power of the study).

Table B-2 presents program impact on a range of business practices. There are a number of small but significant improvements in business practices among program beneficiaries. For example, beneficiaries are more likely to record their transactions and to separate business and household accounts. They are also more likely to report that they believe that profits should be reinvested in inputs. In this sense, the program does seem to have provided beneficiaries with additional business skills, consistent with the productivity gains noted earlier.

Table B-9 presents program impacts on overall household income. The results show that income per capita for beneficiary households is higher but not statistically significant. Income from self-agricultural production increased by 40.6% for beneficiary households but there are no other significant differences in any of the other sources of income.

Table B-10 presents impact on total consumption per capita in beneficiary households. Total consumption includes expenditures in food and non-food items, as well as the value of products self-consumed. Non-food consumption includes products ranging from shoes and clothes to razors and detergents. Food consumption is based on a 15-day recall of typical types of foods consumed in the communities. Self-consumption includes items that were not purchased, but were consumed from own production or received as a form of payment in-kind or as a gift. Total consumption levels for beneficiary households is 13.4 percentage points higher than those in the control group, however this result is not statistically significant. Food expenditures, total per capita value of self-consumption, and expenditures in non-food items are also not significantly different for beneficiary households.

Chapter 5: Social Stimuli and Inequities in Aspirations: A Proposal for Experimental Evidence from DC's Underserved Youth

The objective of this study is to understand how role model characteristics affect the aspirations of mentees and consequently the effectiveness of mentoring programs.[33] The overall objective of the study is to understand how aspirations can be included as a practical component of economic policies and role-model programs. Intuitively, I investigate the social signals that those who are underrepresented are exposed to by various role models, including those who are overrepresented. Much of the current research focuses on the characteristics of those who are underrepresented. I am interested in exploring whether there is a role for those who are overrepresented to address the problem of underrepresentation. Underrepresentation tends to be pronounced for disadvantaged youth, particularly males of color. The project is currently being piloted with some suggestive initial findings available.

Variability in social stimuli that role models send can lead to inequality in aspirations and consequently to the underrepresentation of certain populations. The steps for the proposed study are to i) collect data from youth clubs and/or schools in underrepresented communities in the poorest wards of the Greater Washington's Metropolitan Area ii) to test the hypotheses derived from an applied framework which I present in this chapter in order to see whether certain social stimuli can lead to attitudinal reversals and iii) analyze the data in a rigorous manner that is suitable for ordinal data based on my methodology in chapter 3.

An increasing body of research suggests that aspirations are importantly linked to

[33] The proposal for the study won the Philip Dearborn award of the Economic Club in Washington DC.

decision making La Ferrara, and Duryea (2008), Meng (2009), Deci, and Ryan (2000), Tanguy, and Dercon (2008) and economic outcomes Appadurai (2004), Genicot, and Ray (2010), Ray (2006), Altamirano, Lopez-Calva, and Soloaga (2010)). This pilot will invite policy makers to seriously consider the role of social stimuli on aspirations in underrepresented communities. The research targets underrepresented communities in the DC area where high school dropouts are exceptionally high.

I consider an attitudinal poverty trap where intrinsic constraints, which are non-material constraints residing within an individual, are updated by extrinsic social stimuli. I propose to examine how relaxing these non-monetary and non-informational constraints can lead to attitudinal reversals. In particular, I explore the phenomenon by which individuals do not choose available actions due to externally updated intrinsic constraints. An example of such a phenomenon is the underrepresentation of underserved groups in the US, a country that is committed to equal opportunity in employment.

Borrowing insights from case-based decision theory, I model the underrepresentation of groups as a poverty trap which is driven by similarity weighted subjective probabilities. I design an attitudinal experiment for underrepresented communities and hypothesize that salient social stimuli may lead to attitudinal reversals controlling for race, gender and information. Subjective probabilities are measured with self reported data on an ordinal spectrum of likelihood. This study may challenge traditional hypotheses on the importance and use of role-models by also assigning new meaning to those populations that are overrepresented.

5.1 Literature Review of Aspirations

Economists who study the role of aspirations in economic development have been largely influenced by Appadurai (2004)'s anthropology-based discussion of the "capacity to aspire". The underlying tenet is that, in addition to possession of material goods and other forms of empowerment and opportunity, the poor and the non-poor around them also have an attitudinal capacity within themselves to progress towards escape from poverty. Appadurai described the "capacity to aspire" as being fundamentally related to culture. He stated that, "for more than a century, culture has been viewed as a matter of one or other kind of pastness - the keywords here are habit, custom, heritage, tradition. On the other hand, development is always seen in terms of the future - plans, hopes, targets, goals.... Aspirations certainly have something to do with wants, preferences, choices and calculations. And because these factors have been assigned to the discipline of Economics...they have been largely invisible in the study of culture."

Extending Appadurai's ideas, Ray (2006) argued that the condition of poverty is associated with a "failure to aspire." He modeled the acquisition of the capacity to aspire through persons who are in one's cognitive neighborhood, which he termed, an "aspirations window." An individual fails to aspire when successful people are too far away from one's aspirations window. He also introduces the "aspirations gap" which represents the distance from a target aspiration. Ray suggests that polarization may lead to aspirations failure in poverty-ridden environments and that a highly mobile society invites higher aspirational levels. He also noted that aspirations are probably inherently multidimensional.

Genicot, and Ray (2010) provided a measure of the "aspirations gap" in the space of income, based on the poverty gap measure of Foster, Greer, and Thorbecke (1984). In an empirical illustration of their framework, they measure the aspirational level of income as the average income of those above an individual's income level and they successfully calibrated economic growth rates in Brazil based on their model. Dalton, and Mani (2011) examined the conditions under which an appropriate role model could help a poor person escape from poverty. They explore similarity functions to determine the level of similarity that would create a role model that could address aspirations failure. They demonstrated that a poor person would restrict the set from which their role model is selected.

In psychology, for the past 50 years (see Gollwitzer, and Bargh (1996), for an earlier survey) it has been generally accepted that goals constitute a primary source of motivation and largely predict behavior and achievement (for earlier studies, see Elliott, and Dweck (1988), and Ryan, and Connell (1989). The early literature distinguished between goals that were set intrinsically and extrinsically, and studied their effects on wellbeing. Endogenously set actions had an internal locus of causality, and occurred in individuals who were more cognitively flexible (Grolnick, and Ryan (1987)), and had higher self-esteem Deci, and Sheinman (1981). Locke (1995) argued that consciousness and volition must be taken as axiomatic starting points. In response to research done by behaviorists during the 1950's and 60's, Locke, and Latham (2002) used 35 years of empirical research to demonstrate that task difficulty and specificity are of central importance to task performance and satisfaction. Their studies of work performance-related goal setting have been used broadly in the social sciences. A study of the process of goal setting recognized such traits as commitment and conscientiousness in Ajzen, and Flood (2009).

Theoretical models in economics have largely adopted the goal-driven approach to aspirations. Heath, and Wu (1999) modeled goals as "reference points." Their work showed that loss aversion and diminishing sensitivity, characteristics of the value function of prospect theory Kahneman, and Tversky (1979), and Tversky, and

Kahneman (1992), can characterize goals and model phenomena that are noticed in practice, such as the demotivating effect of high goals. Goals also serve as reference points in Koch, and Nafziger (2011) and Suvorov, and Ven (2009) where goals are endogenously determined by present-biased preferences. They derived a condition under which short-term goals are better than long-term goals, and proposed that loss aversion is a means of self-control.

Benabou, and Tirole (2003) followed psychologists and sociologists in viewing the use of punishment as an incentive as counterproductive. They set up a principal agent model, where the agent exerts more effort when offered control of a task, rather than performing a supervised task. The policy implications are that people in authority positions should offer control, rather than assigned tasks. Cho, and Matsui (2005) assumed that a decision maker "satisfies," rather than optimizes in an infinitely repeated game. They assumed that each player sticks to the same action, if the action has a one period payoff that is larger than a predetermined aspirational level where the past history is summarized according to the average of past payoffs. Other game theoretic models compared target aspirations to average payoffs from actions chosen in the past, a process by which the agents update their aspirations (See Bendor, and Ray (2001), Mookherjee, and Napel (2010), and Karandikar, and Vega-Redondo (1998).

In addition to theoretical models, economists have used empirical studies to explore the role of aspirations in various domains, such as subjective wellbeing, behavior, role modeling and program interventions. Tanguy, and Dercon (2008) showed that narrow aspirations windows, measured as the geographical proximity of the respondent's self-reported role model, stimulate very narrow or very wide aspirations gaps, and can lead to aspirations failure. The aspirations gap was measured by a binary variable that depended on whether or not the respondent thinks she can be as successful as her role model. Alois (2004) found that income aspirations, measured by a subjective income evaluation, are negatively correlated with subjective wellbeing, measured as self-reported satisfaction with life.

Exogenous factors affecting aspirations depend to some extent on role models. La Ferrara, and Duryea (2008) examined the effect of media on decision making. They found that individuals in Brazil who were exposed to certain soap operas had significantly smaller families because they used the characters with small families as role models. Hoffner, and Buchanan (2005) concluded that respondents identified with characters perceived as similar to themselves, and these similarities were based on gender and attitudes. Meng (2009) concluded that the bargaining power of children should be taken more seriously in human capital models, which typically assign the role of the decision making processes entirely to the parents. Linear probability models showed that children's aspirations correspond more strongly to measures of ability than parents' aspirations. Children with higher educational aspirations also had a greater likelihood of staying in school.

Psychosocial support can allow individuals to aspire to greater future wellbeing, as demonstrated by the Solidario program in Chile, studied by Carneiro, and Ginja (2010). In their evaluation, aspirations took the form of specific life projects, such as having housing, stable employment, and higher educational achievements for their children, which the families identified together with the social worker. Random assignment of leadership positions to women in Nicaragua not only led to higher aspirations and greater optimism by the assigned leaders, but also had a positive effect on the rest of the participants, who interacted socially with the leaders in Macours, and Vakis (2009).

In Deci, and Ryan (2000) they introduce self-determination theory which models important qualitative aspects of aspirations, a theory that challenges traditional purely goal-driven approaches. According to this theory “an understanding of human motivation requires a consideration of innate psychological needs for competence, autonomy, and relatedness.” Deci, and Ryan (2001) concluded that “social contexts and individual differences that support satisfaction of the basic needs facilitate natural growth processes including intrinsically motivated behavior and integration of extrinsic motivations, whereas those that forestall autonomy, competence or relatedness are associated with poorer motivation, performance and well-being.”

The influence of social surroundings on intrinsic motivation suggests that an individual’s capacity to aspire might depend on his/her social identity. Akerlof, and Kranton (2000) argued that "identity can explain behavior that appears detrimental." Sen (2007) modeled the process of detrimental identity creation where individuals may fail to identify ways in which they are similar to others. The policy implications of this approach are important for role-model programs adopted in schools. The hypotheses demonstrate how the subjective perception of opportunity can lead to valuable choices being dismissed by certain populations even though they are objectively available. Exploring the role of intrinsic stimuli will establish whether they are a useful policy tool.

5.2 The framework Underlying the Hypotheses

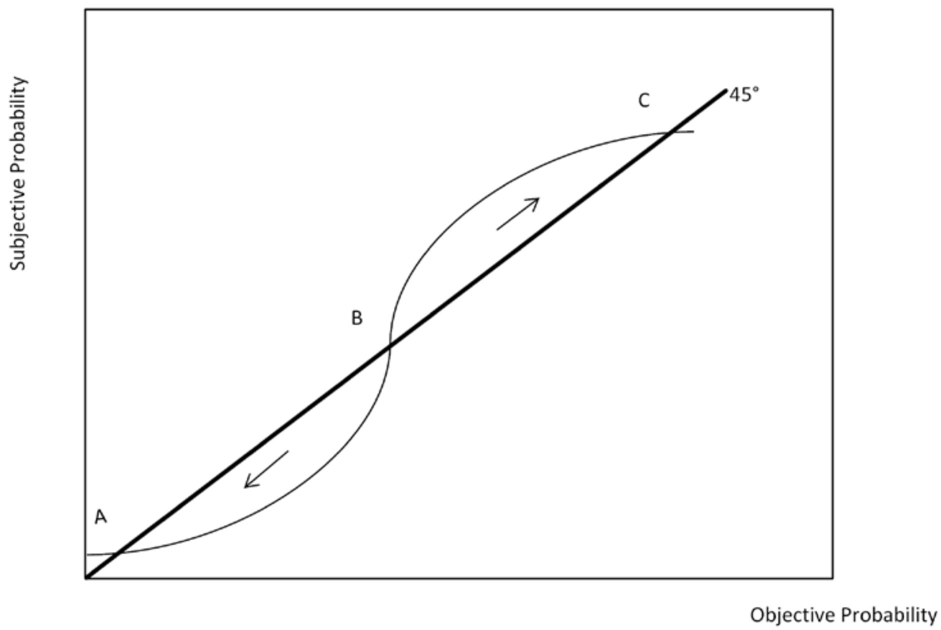
The framework models aspirations, a soft construct recently viewed as having great potential to impact processes that take place early on in life and affect future decision making. It explores the psychosocial foundations of motivation and choice to construct a nuanced model of aspirations that explains persistent low outcomes (poverty traps) and persistent differential outcomes across groups (inequality traps). The hypotheses encompass insights from the capability approach, case-based decision theory (Billot, I., D., and D. (2005)) and identity theory (Sen (2007)). The framework is then used to design a study that aims to pinpoint factors behind aspirational failures among underserved populations and help target policies to address these failures.

The proposal follows a growing trend in empirical and theoretical economics to include constructs that have been previously ignored by economists, due to their perceived "soft" nature. These constructs have been given various labels such as "missing dimensions" (Alkire (2007), Alkire, and Foster (2011a), "subjective wellbeing" (Kahneman, and Deaton (2010)), "soft skills" or "non-cognitive skills" (Heckman (2011)), and "capabilities" (Nussbaum, and Sen (1993)), yet often describe traits and capacities of a similar nature. The central characteristic of these "soft" constructs is that they are not measured by income or other standard economic variables; they reside within the individual and, therefore, are considered to be difficult to measure. Soft constructs are altering the practice and theory of economics in important ways. For example, novel measures of wellbeing are being developed in both multidimensional (Alkire, and Foster (2011a)) and unidimensional (Kahneman, and Deaton (2010)) contexts. In addition, theoretical models that include these constructs and can guide policy in innovative ways are now being considered. There is empirical evidence that soft constructs influence and change preferences (Borghans, Meijers, and Ter Weel), Crano, and Prislin (2008)), and have strong predictive power regarding choice outcomes (Heckman, J., and S. (2006)). Soft constructs are increasingly being recognized as important tools for policy makers.

The capacity to aspire is linked to the process of choice making, a process that ultimately determines choice outcomes. Economists and anthropologists have long recognized the central role of choice in development Appadurai (2004), Sen (1988, (1991)). If development is viewed as freedom (Foster, and Sen (1997), Sen (1999), the role of choice can be better understood when the focus is on the process of choice making, rather than on choice outcomes alone (Sen (2004), Foster (2011), Alkire, and Foster (2011a). In the present context, the process includes a step wherein choice alternatives are evaluated in terms of their perceived availability. This internal screening process removes options or reduces the probability of their being available to persons like the choice maker. A greater capacity to aspire is associated with a broader range of choices being available; I define “failure to aspire” as the phenomenon by which a decision maker who has reason to value certain feasible choices does not perceive those choices as being available. The extent of the failure to aspire can then be measured as the degree to which a set of alternatives is effectively diminished.

In economic models, human motivation is often assumed to be driven by rewards and incentives, a process otherwise known as extrinsic motivation. In an alternative view of human motivation, intrinsic motivation is driven by innate psychological needs that individuals seek to meet Deci, and Ryan (2000) and empirical findings suggest that rewards and incentives have hidden costs and can effectively be demotivating (see Deci, Koestner, and Ryan (1999), Deci (1975)). Thus, there has been a longstanding call for economics models that incorporate intrinsic stimuli in the process of choice making Kreps (1997). Consider for example an environment where equality of opportunity hypothetically exists, yet inequality of participation is observed for certain choice outcomes across the population; consequently, a particular population would be underrepresented. An extrinsic motivational approach would lead to the paradoxical assertion that this population, in the aggregate, does not respond to incentives or is unwilling to exert effort given the cost-to-effort. In other words, the underrepresented population is assumed to be either irrational or lazy. The model clarifies how studying the process of choice making, as it influences the capacity to aspire, is crucial to understanding why some populations do not consider available choices, even though they might have reason to value them Sen (2007).

Figure 5-1: Underrepresentation Trap



In Figure 5-1 the “aspirations failure” trap is represented by point A. Let $y \in 0,1$ be the success or failure of the role models, the people that followed a certain pathway the decision-making agent is considering. The vector of characteristics of the role models are $\mathbf{x} = (x^1, ..., x^m)$ where some or all of them may be discrete. The similarity function s is such that for two vectors of characteristics (or identities) $\mathbf{x}_i = (x_i^1, ..., x_i^m)$ and $\mathbf{x}_j = (x_j^1, ..., x_j^m)$ it measures the similarity between agent i and agent j with identity vectors $\mathbf{x}_i$ and $\mathbf{x}_j$. Define the subjective probability $\hat{y}_{n+1} = \frac{\sum_{i \le n} s(\mathbf{x}_i, \mathbf{x}_{n+1}) y_i}{\sum_{i \le n} s(\mathbf{x}_i, \mathbf{x}_{n+1})}$ where the probability that the agent assigns to a pathway she considering is the s-weight of all past successful cases of an individual following that particular pathway divided by the total s-weight of all past cases, successes and failures. The updating and evolution of probabilities occurs until the subjective probability is equal to the objective probability at point C. One channel of similarity updating that this study explores is through social stimuli that witnesses (or role models in our study) send. When at point A, no agent makes the first step of assigning a high likelihood to an alternative pathway that could help them escape poverty. This could mean that even though an alternative is available to the agent, it is not considered.

I support that the choice of the functional form that determines the threshold of a target must be behaviorally justified, even if the functional form depends on an individual's social neighborhood. An average of payoffs or incomes, for example, models the process of goal setting in a mechanistic, rather than a behavioral, manner. In addition to the creation of role models - those who share the same identity characteristics and can be added to one's aspirations window - people who do not share many of the same identity characteristics and are outside of an individual's aspirations window can also represent a crucial tool for policy. Though this may seem unconventional, the proposed research may clarify how such people can be useful for addressing the failure to aspire.

The applied framework draws upon case-based decision theory where similarity-weighted frequencies are introduced as subjective probabilities Billot, I., D., and D. (2005). In psychology, a similarity weighted probability also models exemplar learning, and is assumed to represent a belief Gilboa (2009). Tools from the attitude and attitude change literature are also used to model this perception. Crano, and Prislin (2008) defined an attitude as "an evaluative integration of cognitions and affects experienced in relation to an object." In the proposed model, social stimuli constitute the attitude object.

Failure to aspire therefore represents the extent to which opportunity sets are smaller under subjective beliefs than under objective frequencies. The counterfactual world is represented by what preferences could have been under intrinsic signals. Hence, changes in choice outcomes can occur if i) the role models were different or ii) stimuli sent by existing role-models were different. Research in the attitude literature (an extensive survey can be found in Crano, and Prislin (2008) determined that certain attitudes are driven by the central processing or the peripheral processing of messages related to the attitude object. In the present context, a decision maker considers different pathways based on evidence of role model successes or failures. Each role model is weighted by the similarity between the characteristic vectors and the framework uses as a variable the central and peripheral processing of witness characteristics that determine an individuals belief that an option is available.

This framework explicitly models attitudes, which are constructs subject to manipulation and hence powerful policy tools, and which up to now, have been largely viewed as passive in economics. It identifies policy tools that can be used to change attitudes, perceived opportunities, effective preferences and finally choices. Instead, wellbeing and decisions are seen as depending on perceived opportunities, which can fall short of objectively viewed opportunities in the case of a failure to aspire.

The pilot questionnaire, which is presented in Appendix C, is guided by this framework. The survey is designed to examine how relaxing non-monetary and non-informational constraints can lead to attitudinal reversals. The findings of this study may shed light on existing hypotheses for the importance of aspirations that leads individuals to invest in their future – such as showing up for school.

5.3 Experimental Design

Though the original proposal for implementing this study was designed to capture entrepreneurial aspirations, in meetings with DCPS staff and the Deputy Mayor for Education, I was informed that one of the most pressing problems they would like to address is student absenteeism. They invited me to consider changing the survey design so that it offers insight to this particular problem. Student absenteeism is a good predictor of dropping out of high school, which in turn predicts low outcomes in many dimensions later on in life. Empirical findings in development economics point towards a "man-made trap" as discussed in Duflo (2012). In this man-made trap expectations of the benefits of a high school diploma may be far-fetched, compared to realistic expectations of the basic skills a high school diploma offers. I propose to investigate the external validity of these findings in the United States. The surveys have been designed so that these insights are incorporated.

The collection of data involves presenting a role model describing their success story on a screen where each sentence of the story appears sequentially. The students in the schools and youth clubs targeted for the study are 99 per cent Black and academically under-performing. The study has received IRB approval and the research involves the proper protocol of finding the permission of guardians. The IRB has classified the study to be of minimal risk to the human subjects under study.

Currently, information from DCPS students is being collected in social media settings where they are invited to state why they do not show up for school. Based on the insights that this exercise provided, the narrative of the role-model stories were designed and presented in Appendix C. The design allows us to simulate three levels of similarity s, $s \in [0,1,2]$ and two types of social stimuli $S \in [i,r]$ where i denotes intrinsic and r denotes extrinsic. Extrinsic stimuli to stay in school focus on the end result of attaining a high school diploma. Intrinsic stimuli pertain to a more realistic account of the process through which a high school diploma is attained. Additionally, aspirational targets can be either white or blue collar success stories. The role models are also differentiated by gender G. The set of hypotheses is

$$\mathbf{H}_0 : \mathbf{E}[y \mid s, G, S] - \mathbf{E}[y \mid s^*, G^*, S^*] = 0 \tag{5-1}$$

Block randomization will result in $\mathbf{2^4}$ categories of variation defined by gender, race, type of stimulus and aspirational target. Intuitively, the set of hypotheses tests whether attitudes towards the role models will reverse after controlling for extrinsic and intrinsic stimuli and whether these reversals are stronger as similarity with the role model increases. Rejection of each of the hypotheses will support that incentives, such as daily raffles currently used by DCPS or unrealistic perceptions hidden in acquiring a high school diploma, need to be revisited. Identifying attitudinal reversals involves estimating the likelihoods with which subjects report that they could do what the role model has done based on their gender similarity, race similarity and whether they were exposed to intrinsic or extrinsic stimuli and to a white-collar or blue-collar success story.

As with all self-reported ordinal data, self-reporting bias and interpersonal comparability concerns remain. However, such considerations are minimized through randomization. The design entails the calibration of subjective and objective probabilities based on the degree to which perceived similarities are distorted by the peripheral processing of role-models who may not possess the same characteristics that constitute the identity of the underrepresented population. In the same spirit, Bowles, and Hoff (2006) note that “similar individuals in dissimilar socioeconomic environments develop different preferences and beliefs that can transmit poverty or affluence from generation to generation.”

5.4 Early Findings and Concluding Remarks

Early findings from a pilot exercise in the Girls and Boys youth club in Anacostia have suggested that white male role models are discouraging to young males of color yet equivalent findings do not exist for females. When the same aspirational story is delivered by a white male, black youth seem to be more discouraged than when a non-white person of any gender shares the same story.

The findings of this study will shed light on how social stimuli, often in a highly-frequent and compounded manner, lead to the construction of internal binding constraints. Understanding how the process of aspirations creation can lead to inequality in economic outcomes will add to our understanding of early processes that take place in one's environment. Changing existing attitudes may lead to better economic choices for individuals, such as showing up for school. The findings of this pilot may inspire larger scale policy-driven research for the problem of student absenteeism in the DC area.

Depending on the findings, this research may invite change to the national standards for designing role model programs. Material based on this research was incorporated into two courses of the undergraduate economics 2151 class at the George Washington University. A team of selected undergraduate students has contributed to the questionnaire and the narratives design. The model and measures will be made publicly available for practitioners, researchers and policy makers to use. The research will help shape national and policies concerning vulnerable populations in the US.

The proposed research could also contribute to the field of education where soft constructs have taken on a greater salience recently. The capacity to aspire is an important non-cognitive skill in K-12 schooling (see Heckman and Rubinstein 2001 for a discussion of non-cognitive skills) and, as previously discussed, such skills can be powerful policy tools. Although there is an abundance of psychometric measures for non-cognitive skills, there remains a need for intuitive aggregate measures that are transparent and readily used by policy makers involved in evaluating and improving the quality of education.

References

ABUL NAGA, R. H., AND T. YALCIN (2004): "Inequality Measurement for Ordered Response Health Data," *Journal of Health Economics*, 23, 1614-1625.

AJZEN, I. C. C., AND M. FLOOD (2009): ""From Intentions to Behavior: Implementation Intention, Commitment, and Conscientiousness," *Journal of Applied Social Psychology*, 39, 1356-1372.

AKERLOF, G. A., AND R. E. KRANTON (2000): "Economics and Identity," *The Quarterly Journal of Economics*, 115, 715-753.

ALKIRE, S. (2007): "The Missing Dimensions of Poverty Data: Introduction to the Special Issue," *Oxford Development Studies*, 35 347-359.

ALKIRE, S., AND J. FOSTER (2011): "Counting and Multidimensional Poverty Measurement," *Journal of Public Economics*, 95, 476-487.

— (2011): "Understandings and Misunderstandings of Multidimensional Poverty Measurement," *Journal of Economic Inequality*, 9, 289-314.

ALLISON, R. A., AND J. E. FOSTER (2004): "Measuring Health Inequality Using Qualitative Data," *Journal of Health Economics*, 27, 505-524.

ALOIS, S. (2004): "The Role of Income Aspirations in Individual Happiness," *Journal of Economic Behavior and Organization*, 54, 89-109.

ALTAMIRANO, A., L. F. LOPEZ-CALVA, AND I. SOLOAGA (2010): "Inequality and Teenagers' Educational Aspirations in Urban Mexico," El Colegio de México: Centro de Estudios Económicos.

APOSTOL, T. M. (1974): *Mathematical Analysis*. Addison-Wesley.

APPADURAI, A. (2004): "The Capacity to Aspire: Culture and the Terms of Recognition," Stanford University press, 59-84.

BANERJEE, A., E. DUFLO, R. GLENNESTER, AND D. KOTHARI (2010): "Improving Immunization Coverage in Rural India: Clustered Randomised Controlled Evaluation of Immunization Campaigns with and without Incentives," *British Medical Journal*, 340, c2220-c2220.

BARDHAN, P., AND C. UDRY (1990): *Development Microeconomics*. Oxford University Press.

BASU, K. (1999): "Child Labor: Cause, Consequence, and Cure, with Remarks on International Labor Standards," *Journal of Economic Literature*, 37.

— (2006): "Gender and Say: A Model of Household Behaviour with Endogenously Determined Balance of Power," *The Economic Journal*, 116, 558-580.

BENABOU, R., AND J. TIROLE (2003): "Intrinsic and Extrinsic Motivation," *Review of*

Economic Studies, 70, 489-520.

BENDOR, J. M. D., AND D. RAY (2001): "Aspiration-Based Reinforcement Learning in Repeated Interaction Games: An Overview," *International Game Theory Review*, 3, 159-174.

BENNETT, C. J., AND C. HATZIMASOURA (2011): "Poverty Measurement with Ordinal Data," *The Institute for International Economic Policy*, Working Paper 2011-14.

BERTRAND, M., E. DUFLO, AND S. MULLAINATHAN (2004): "How Much Should We Trust Differences-in-Differences Estimates?," *Quarterly Journal of Economics*, 119(1), 249-275.

BILLOT, A., G. I., S. D., AND S. D. (2005): "Probabilities as Similarity-Weighted Frequencies " *Econometrica*, 1125-1136.

BORGHANS, L., H. MEIJERS, AND B. A. S. TER WEEL ("The Role of Noncognitive Skills in Explaining Cognitive Test Scores," *Economic Inquiry*, 46, 2-12.

BOWLES, S. S. N. D., AND K. HOFF (2006): *Poverty Traps*. Princeton University Press.

BRUHN, M., AND D. MCKENZIE (2008): "In Pursuit of Balance: Randomization in Practice in Development Field Experiments," *World Bank Policy Research Working Paper*, WPS4752.

CARNEIRO, P. G. E., AND R. GINJA (2010): "The Impact of Providing Psycho-Social Support to Indigent Families and Increasing Their Access to Social Services: Evaluating Chile Solidario," *Evaluation report submitted to Chilean Government*.

CHIAPPORI, P. A. (1997): "Introducing Household Production in Collective Models of Labor Supply," *Journal of Political Economy*, 105, pp. 191-209-209.

CHO, I. K., AND A. MATSUI (2005): "Learning Aspiration in Repeated Games," *Journal of Economic Theory*, 124, 171-201.

CRANO, W. D., AND R. PRISLIN (2008): *Attitudes and Attitude Change*. New York: Psychology Press.

DALTON, P. S. S. G., AND A. MANI (2011): "Poverty and Aspirations Failure," *Working Paper, Competitive Advantage in the Global Economy (CAGE)*.

DECI, E. L. (1975): *Intrinsic Motivation*. New York: Plenum Press.

DECI, E. L., R. KOESTNER, AND R. M. RYAN (1999): "A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation," *Psychological Bulletin*, 6, 627-668.

DECI, E. L., AND R. M. RYAN (2000): "The ``What" and ``Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior," *Psychological Inquiry*, 11, 227-268.

DECI, E. L. J. N., AND L. SHEINMAN (1981): "Characteristics of the Rewarder and Intrinsic Motivation of the Rewardee," *Journal of Personality and Social Psychology*, 40, 1-10.

DECI, E. L. K. R., AND R. M. RYAN (2001): "Extrinsic Rewards and Intrinsic Motivation in Education : Reconsidered Once Again," *Washington, DC: American Educational Research Association*.

DONALDSON, D., AND J. A. WEYMARK (1986): "Properties of Fixed-Population Poverty Indices," *International Economic Review*, 27, 667-688.

DUFLO, E. (2012): "Women Empowerment and Economic Development," *Journal of Economic Literature*, 50, 1051-1079.

DUFLO, E., AND C. UDRY (2004): "Intrahousehold Resource Allocation in Cote Divoire: Social Norms, Separate Accounts and Consumption Choices," *NBER Working Paper*.

ELLIOTT, E. S., AND C. S. DWECK (1988): "Goals: An Approach to Motivation and Achievement," *Journal of Personality and Social Psychology*, 54, 5-12.

FOSTER, J., J. GREER, AND E. THORBECKE (1984): "A Class of Decomposable Poverty Measures," *Econometrica*, 52, 761-766.

— (2010): "The Foster-Greer-Thorbecke (Fgt) Poverty Measures: 25 Years Later," *Journal of Economic Inequality*, 8, 491-524.

FOSTER, J. E. (2011): "Freedom, Opportunity, and Well-Being," in *Handbook of Social Choice and Welfare*, ed. by A. K. J., S. A., and S. K. Oxford: North Holland, 687-728.

FOSTER, J. E., M. MCGILLIVRAY, AND S. SETH (2012): "Rank Robustness of Composite Indices: Dominance and Ambiguity," *OPHI Woking Paper No. 26*.

FOSTER, J. E., AND A. SEN (1997): "Annex to the Enlarged Edition of on Economic Inequality, ," in *On Economic Inequality: After a Quarter Century*, ed. by A. Sen. Oxford: Clarendon Press.

GENICOT, G., AND D. RAY (2010): "Aspirations and Inequality," *Georgetown Working Paper*.

GILBOA, I. (2009): *Theory of Decision under Uncertainty* Cambridge: Cambridge University Press.

GOLLWITZER, P. M., AND J. A. BARGH (1996): *The Psychology of Action: Linking Cognition and Motivation to Behavior*. Guilford Pressl.

GROLNICK, W. S., AND R. M. RYAN (1987): "Autonomy in Children's Learning: An Experimental and Individual Difference Investigation," *Journal of Personality and Social Psychology*, 52, 890-898.

HANSEN, M., M. LEMKE, AND N. SORENSEN (2014): "Combining Multiple Performance Measures: Do Common Approaches Undermine Districts' Personnel Evaluation Systems?," *National Center for Analysis of Longitudinal Data in Education Research, Working Paper 118*.

HARSANYI, J. C. (1953): "Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking," *Journal of Political Economy*, 61, 434-434.

HEATH, C. R. P. L., AND G. WU (1999): "Goals as Reference Points," *Cognitive Psychology*, 38, 79-109.

HECKMAN, J. J. (2011): *Hard Evidence on Soft Skills: The Ged and the Problem of Soft Skills in America*. Chicago: University of Chicago Press.

HECKMAN, J. J., S. J., AND U. S. S. (2006): "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior " *Journal of Labor Economics*, 411-482.

HOFFNER, C., AND M. BUCHANAN (2005): "Young Adults' Wishful Identification with Television Characters: The Role of Perceived Similarity and Character Attributes," *Media Psychology*, 7, 325-351.

KAHNEMAN, D., AND A. DEATON (2010): "High Income Improves Evaluation of Life but Not Emotional Well-Being," 16489-93.

KAHNEMAN, D., AND A. KRUEGER (2006): "Development in the Measurement of Subjective Wellbeing," *Journal of Economic Perspectives*, 20, 3-24.

KAHNEMAN, D., AND A. TVERSKY (1979): "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47, 263-291.

KARANDIKAR, R. M. D., AND F. VEGA-REDONDO (1998): "Evolving Aspirations and Cooperation," *Journal of Economic Theory*, 80, 292-331.

KOCH, A. K., AND J. NAFZIGER (2011): "Self-Regulation through Goal Setting," *Scandinavian Journal of Economics*, 113, 212-227.

KREPS, D. M. (1997): "Intrinsic Motivation and Extrinsic Incentives," *American Economic Review*, 359-64.

LA FERRARA, E. C. A., AND S. DURYEA (2008): "Soap Operas and Fertility: Evidence from Brazil," *IDB Working Paper No. 633*.

LOCKE, E. A. (1995): "Beyond Determinism and Materialism, or Isn't It Time We Took Consciousness Seriously?," *Journal of Behavior Therapy and Experimental Psychiatry*, 26, 265-273.

LOCKE, E. A., AND G. P. LATHAM (2002): "Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-Year Odyssey," *American Psychologist*, 57, 705-717.

MACOURS, K., AND R. VAKIS (2009): "Changing Households' investments and Aspirations through Social Interactions: Evidence from a Randomized Transfer Program," *Policy Research Working Paper 5137, The World Bank*.

MADDEN, D. (2010): "Ordinal and Cardinal Measures of Health Inequality: An Empirical Comparison," *Health Economics*, 19, 243-250.

MENG, J. (2009): ""Aspirations and Schooling: Analysis of the Formation and Intra-Household Impact of Educational Aspirations in Rural China," *Master of Science in Economics, University of Oxford*.

MOOKHERJEE, D. D. R., AND S. NAPEL (2010): "Aspirations, Segregation, and Occupational Choice," *Journal of the European Economic Association*, 8, 139-168.

NEWEY, W. K. (2007): "Treatment Effects," *MIT Mimeo*.

NUSSBAUM, M., AND A. SEN (1993): *The Quality of Life* Oxford: Oxford University Press.

RANGEL, M. A., AND D. THOMAS (2012): "Gender, Production and Consumption: Allocative Efficiency Wthin Farm Households," *Princeton Working Paper*.

RAWLS, J. (1999): *A Theory of Justice*. Oxford University Press.

RAY, D. (2006): "Aspirations, Poverty and Economic Change," Oxford University Press, 59-84.

RYAN, R. M., AND J. P. CONNELL (1989): "Perceived Locus of Causality and Internalization: Examining Reasons for Acting in Two Domains," *Journal of Personality and Social Psychology*, 57, 749-761.

SEN, A. (1976): "Poverty: An Ordinal Approach to Measurement," *Econometrica*, 44, 219-231.

SEN, A. (1988): "Freedom of Choice: Concept and Content," *European Economic Review*, 269-294.

— (1991): "Welfare, Preference and Freedom " *Journal of Econometrics*, 15-29.

— (2004): *Rationality and Freedom* Cambridge: Harvard University Press.

SEN, A. K. (1999): *On Economic Inequality with a Substantial Annexe `after a Quarter Century' by J. Foster and A. Sen*. Clarendon Press.

— (2007): *Identity and Violence: The Illusion of Destiny*. W. W. Norton & Company.

STRAUSS, J., AND D. THOMAS (1996): "Measurement and Mismeasurement of Social Indicators," *The American Economic Review*, 86, pp. 30-34-34.

SUVOROV, A., AND J. V. D. VEN (2009): "Goal Setting as a Self-Regulation Mechanism," *SSRN Working Paper*.

TANGUY, B. T. A. S., AND S. DERCON (2008): "Aspirations Failure and Well-Being Outcomes in Ethiopia: Towards an Empirical Exploration," *Working Paper 28, Oxford University*.

TODARO, M. P., AND S. SMITH (2014): *Economic Development*. Prentice Hall.

TVERSKY, A., AND D. KAHNEMAN (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5, 297-323.

WOOLDRIDGE, J. M. (2002): *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

ZHENG, B. (1997): "Aggregate Poverty Measures," *Journal of Economic Surveys*, 11, 123-162.

— (2008): "Measuring Inequality with Ordinal Data: A Note," in, ed. by J. Bishop, and B. Zheng: Emerald Group Publishing Limited, 177-188.

Appendix A: Evaluation Reversals Examples

Scale:

10.0000 10.5000 16.7500 17.0000 17.2500

Percentage of reversals: 0.1338

Scale:

1 4 9 16 25

Percentage of reversals: 0.0481

Scale:

0 0.3010 0.4771 0.6021 0.6990

Percentage of reversals: 0.0531

Scale:

1.0000 2.5000 3.0000 3.5000 5.0000

Percentage of reversals: 0.0600

Appendix B: RCT Evaluation Results

Table B-1: Program impacts on Beneficiary's Income (Logged)

Variable Description	N	Control	Treatment	Difference	SE
Total Beneficiary income	877	5.037	5.389	0.396**	(0.177)
Total annual income from agricultural wage employment	877	1.241	0.788	-0.105	(0.137)
Total annual income for non-agricultural self-employment	877	2.690	2.866	0.179	(0.267)
Total annual income for non-agricultural wage employment	877	0.537	0.518	-0.017	(0.145)
Income from sale of titular's own animals	877	2.010	2.851	0.481**	(0.195)
Income from backyard agricultural production	877	3.102	3.879	0.315	(0.261)
Income from temporary migration	877	1.343	1.376	-0.078	(0.195)

Table B-2: Program Impacts on Business practices

Variable Description	N	Control	Treatment	Difference	SE
Business Practices					
How many customers do you have	877	7.600	9.415	2.840	(1.759)
Proportion of customers that are allowed to borrow	872	0.185	0.185	-0.009	(0.022)
Are activities recorded	877	0.167	0.222	0.060***	(0.022)
Titular records accounts	877	0.129	0.160	0.036**	(0.018)
Do you calculate profits of main activity	877	0.320	0.415	0.093***	(0.034)
Separate accounts within the household	877	0.210	0.274	0.050*	(0.026)
Household and business money saved separately	877	0.235	0.289	0.043*	(0.026)
Are household accounts written down	877	0.091	0.091	0.005	(0.018)
Business Practice Perceptions					
It's important to diversify	833	0.094	0.057	-0.032	(0.022)
It's important to practice accounting	833	0.135	0.119	-0.039**	(0.019)
Businesses fail because owners don't work hard enough	832	0.095	0.075	-0.035*	(0.019)
Customer service can always improve	832	0.054	0.041	-0.022	(0.015)
Businesses should discourage credit	877	0.227	0.247	0.021	(0.029)
Business profits should be reinvested in merchandise	877	0.121	0.170	0.049*	(0.026)

Table B-3: Program impacts on Beneficiary's Employment

Variable Description	N	Control	Treatment	Difference	SE
Last 12 Months					
Was self-employed in agricultural activity	835	0.665	0.651	-0.041	(0.032)
Was self-employed in agricultural activity (including livestock)	835	0.946	0.967	0.016	(0.011)
Was self-employed in non-agricultural activity	835	0.371	0.372	0.003	(0.031)
Worked in agricultural wage job	835	0.182	0.138	0.007	(0.021)
Worked in non-agricultural wage job)	835	0.076	0.072	-0.003	(0.018)
Was self-employed in livestock	835	0.892	0.933	0.022	(0.018)
Worked as wage employed in agriculture	835	0.182	0.138	0.007	(0.021)
Worked as self-employment in food production	835	0.231	0.228	-0.009	(0.027)
Worked as self-employment in manufacturing	835	0.016	0.008	-0.010	(0.006)
Worked as self-employment in trading	835	0.137	0.182	0.059***	(0.020)
Worked as self-employment in services	835	0.052	0.056	-0.003	(0.011)
Worked as wage employed in private jobs	835	0.038	0.028	-0.000	(0.012)
Worked as wage employed in public jobs	835	0.038	0.044	-0.003	(0.011)
Total no of days worked (excluding agricultural activity)	835	110.865	107.685	-3.939	(9.998)
Total # days worked in non-agricultural self-employment	835	85.661	84.523	-7.027	(10.742)
Total # days worked in agricultural wage job	835	10.580	6.228	0.454	(1.271)
Total # days worked in non-agricultural wage job	835	14.625	16.933	2.634	(4.727)
Total # days worked as self-employment in food production	835	40.957	34.295	-12.987*	(6.920)
Total # days worked as self-employment in manufacturing	835	3.640	1.205	-2.215*	(1.315)

Variable Description	N	Control	Treatment	Difference	SE
Total # days worked as self-employment in commerce	835	33.769	43.123	12.022**	(6.113)
Total # days worked as self-employment in services	835	7.294	5.900	-3.847**	(1.869)
Total # days worked as wage employed in private jobs	835	3.081	5.462	4.573*	(2.598)
Total # days worked as wage employed in public jobs	835	11.544	11.472	-1.940	(3.687)
Last Week					
Worked last week -including unpaid work	835	0.391	0.423	0.040	(0.038)
Hours worked	835	8.609	7.705	0.299	(1.047)
Hours worked as self-employed in agriculture	835	2.348	3.255	1.131**	(0.497)
Hours worked as a wage employed in agriculture	835	1.110	0.258	-0.656***	(0.225)
Hours worked in non-agricultural self-employment	835	2.722	2.043	-0.347	(0.693)
Hours worked as in non-agricultural wage jobs	835	1.370	1.500	0.314	(0.535)
Small Business Outcomes (last 12 months)					
Household has a small food production business	834	0.220	0.226	-0.009	(0.026)
Household has a small trading business	835	0.166	0.206	0.054**	(0.023)
Household has a small service business	835	0.094	0.098	-0.003	(0.013)
Household has any types of business	877	0.623	0.689	0.056***	(0.022)
Number of businesses owned by the household	834	0.913	1.121	0.101	(0.067)
Household still has this business	834	0.547	0.616	0.035	(0.025)
Value of equipment/tools (log)	837	4.062	4.537	0.207	(0.217)
Expects the business will run in December (peak season)	834	0.507	0.603	0.085***	(0.032)

Table B-4: Program Impacts on Savings

Variable Description	N	Control	Treatment	Difference	SE
Did you save in the past 12 months?	835	0.105	0.237	0.115***	(0.021)
How much did you save last month? (log)	838	0.247	0.586	0.338***	(0.099)
Are you currently saving?	830	0.081	0.191	0.090***	(0.015)
Does titular have separate household savings from her savings?	838	0.040	0.115	0.057***	(0.019)

Table B-5: Program Impact on Credit

Variable Description	N	Control	Treatment	Difference	SE
Did you use a seed bank in the last 12 months?	831	0.013	0.324	0.275***	(0.036)
In past 12 months has anyone taken out a loan	831	0.419	0.661	0.205***	(0.035)
Did you get loan from FUMDEC	832	0.000	0.323	0.321***	(0.040)
Did you get loan from elsewhere (formal)	831	0.132	0.113	-0.019	(0.029)
Did you take a loan from an informal source (including communal banks)	831	0.009	0.069	0.059***	(0.012)
Purpose of loan was agricultural activities	486	0.230	0.462	0.217***	(0.046)
Purpose of loan was non-agricultural activities	486	0.033	0.158	0.157***	(0.024)
Purpose of loan was for the household	486	0.038	0.048	0.008	(0.015)
Do you desire a loan in the coming months?	831	0.103	0.383	0.280***	(0.034)

Table B-6: Employment of household adults (excluding beneficiary and children)

Variable Description	N	Control	Treatment	Difference	SE
Last 12 Months					
Was self-employed in agricultural activity	1,565	0.751	0.780	0.025	(0.019)
Was self-employed in agricultural activity (including livestock)	1,565	0.841	0.886	0.034**	(0.016)
Was self-employed in non-agricultural activity	1,565	0.104	0.121	0.027*	(0.015)
Worked in agricultural wage job	1,565	0.434	0.400	0.002	(0.032)
Worked in non-agricultural wage job)	1,565	0.067	0.078	0.005	(0.013)
Was self-employed in livestock	1,565	0.360	0.414	-0.004	(0.021)
Worked as wage employed in agriculture	1,565	0.434	0.400	0.002	(0.032)
Worked as self-employment in food production	1,565	0.023	0.027	0.007	(0.007)
Worked as self-employment in manufacturing	1,565	0.004	0.003	-0.001	(0.002)
Worked as self-employment in trading	1,565	0.051	0.052	0.010	(0.012)
Worked as self-employment in services	1,565	0.033	0.044	0.011	(0.008)
Worked as wage employed in private jobs	1,565	0.052	0.049	-0.005	(0.011)
Worked as wage employed in public jobs	1,565	0.016	0.029	0.010	(0.006)
Total no of days worked (excluding agricultural activity)	1,565	86.076	74.158	1.436	(7.868)
Total # days worked in non-agricultural self-employment	1,565	21.717	20.486	2.995	(4.287)
Total # days worked in agricultural wage job	1,565	56.171	44.370	-2.429	(6.404)
Total # days worked in non-agricultural wage job	1,565	8.188	9.302	0.870	(1.804)
Total # days worked as self-employment in food production	1,565	4.402	3.826	-0.200	(1.822)
Total # days worked as self-employment in manufacturing	1,565	0.357	0.041	-0.044	(0.194)
Total # days worked as self-employment in commerce	1,565	12.287	10.740	1.250	(2.637)
Total # days worked as self-employment in services	1,565	4.671	5.879	1.988	(1.217)
Total # days worked as wage employed in private jobs	1,565	4.718	3.639	-0.759	(1.640)

Variable Description	N	Control	Treatment	Difference	SE
Total # days worked as wage employed in public jobs	1,565	3.470	5.663	1.629	(1.307)
Last Week					
Worked last week -including unpaid work	1,565	0.773	0.811	0.040**	(0.018)
Hours worked	1,565	31.888	33.547	1.525*	(0.859)
Hours worked as self-employed in agriculture	2,182	10.734	13.098	1.567**	(0.799)
Hours worked as a wage employed in agriculture	2,182	5.521	4.097	-1.127	(0.769)
Hours worked in non-agricultural self-employment	2,182	0.722	0.781	0.207	(0.213)
Hours worked as in non-agricultural wage jobs	2,182	1.487	1.924	0.123	(0.380)

Table B-7: Household Income for Adults excluding the beneficiary (Logged)

Variable Description	N	Control	Treatment	Difference	SE
Total income from non-agricultural activity (4,5,6,7)	2,182	0.411	0.548	0.134	(0.088)
Total income from agricultural wage job (3, peon)	1,564	3.394	3.003	-0.083	(0.257)
Total income from non-agricultural wage job (8,9)	2,182	0.297	0.357	0.046	(0.067)
Total income from self-employment in food production	1,564	0.093	0.106	-0.002	(0.035)
Total income from self-employment in manufacturing	1,564	0.023	0.023	0.003	(0.015)
Total income from self-employment in trading	1,564	0.269	0.305	0.076	(0.103)
Total income from self-employment in services	1,564	0.238	0.335	0.087	(0.064)
Total income from wage employment in private jobs	1,564	0.293	0.259	-0.011	(0.076)
Total income from wage employment in public jobs	1,564	0.131	0.227	0.082	(0.059)

Table B-8: Child Labor and Schooling (children aged 16 and below)

Variable Description	N	Control	Treatment	Difference	SE
School attendance (current school year)	1,222	0.717	0.785	0.061**	(0.026)
Any work	1,217	0.760	0.767	-0.018	(0.026)
Usual hours of work per week	1,183	8.709	11.104	1.449	(1.073)
School only	1,217	0.195	0.204	0.023	(0.024)
Work only	1,217	0.237	0.185	-0.053**	(0.025)
Both in work and in school	1,217	0.523	0.581	0.036	(0.029)
Neither in work nor in school	1,217	0.045	0.030	-0.005	(0.012)

Table B-9: Household and Per Capita Income (Logged)

Variable Description	N	Control	Treatment	Difference	SE
Total Annual income Per Capita (logged)	877	7.617	7.884	0.156	(0.128)
Total Per Capita annual income from agricultural wage job	877	4.753	4.124	-0.188	(0.278)
Total Per Capita annual income for non-ag self-employment	877	3.392	3.437	0.143	(0.310)
Total Per Capita annual income for non-ag wage job	877	1.558	1.526	-0.090	(0.198)
Total Per Capita annual income from agricultural self-employment	877	3.336	4.201	0.406**	(0.158)
Total Per Capita income from temporary migration	877	1.853	1.698	-0.078	(0.195)
Total Per Capita annual income from livestock	877	3.394	4.009	0.305	(0.296)
Total Per Capita Backyard Income	877	2.520	3.144	0.247	(0.222)

Table B-10: Consumption Per Capita (Logged)

Variable Description	N	Control	Treatment	Difference	SE
Total per capita food and non-food consumption (logged)	877	8.884	9.078	0.134	(0.088)
Total per capita food consumption overall (logged)	877	8.76	8.934	-0.083	(0.257)
Total per capita food self-consumption (logged)	877	8.567	8.749	0.046	(0.067)
Total per capita food purchased (logged)	877	6.901	6.931	-0.002	(0.035)
Total per capita non-food consumption (logged)	877	6.66	6.928	0.003	(0.015)

Table B-11: Intra-household allocation of private consumption goods

Variable Description	N	Control	Treatment	Difference	SE
Titular's share of all non-food items	759	0.344	0.372	0.051***	(0.018)
Husband's share of all non-food items	759	0.318	0.275	-0.053***	(0.019)
Others' share of all non-food items	759	0.338	0.353	0.002	(0.020)

Table B-12: Aggregate Indices for Agency Proc

Variable Description	N	Control	Treatment	Difference	SE
Intrinsic Motivation (Aspirations) Index	830	0.808	0.826	0.010	(0.006)
Locus of Control Index	831	0.778	0.793	0.014*	(0.009)
Self-Esteem Index	830	0.844	0.864	0.005	(0.007)
Aggregate days of feeling positive	831	10.877	11.333	-0.107	(0.327)
Aggregate days of feeling negative	831	27.127	26.249	-0.529	(1.352)

Table B-13: Assets

Variable Description	N	Control	Treatment	Difference	SE
Index of durables	832	0.442	0.446	-0.026***	(0.005)
Owns a radio	833	0.134	0.131	-0.019*	(0.011)
Owns a stove	833	0.049	0.031	-0.026***	(0.010)
Owns a vehicle	833	0.123	0.126	-0.023	(0.020)
Owns a refrigerator	833	0.056	0.021	-0.035***	(0.012)
Owns a fan	832	0.036	0.015	-0.024***	(0.008)
Owns a grinder	833	0.940	0.956	0.001	(0.014)
Owns an iron	833	0.152	0.072	-0.097***	(0.023)
Owns a tv	833	0.190	0.218	0.012	(0.026)
Owns a bicycle	833	0.266	0.290	-0.015	(0.031)
Owns dishes	833	0.996	0.992	-0.004	(0.007)
Owns cookware	833	0.991	0.982	-0.010	(0.007)
Owns a table	833	0.729	0.731	-0.083***	(0.023)
Owns a chair	833	0.736	0.767	-0.059**	(0.026)

Variable Description	N	Control	Treatment	Difference	SE
Owns a truck	833	0.083	0.077	-0.032*	(0.019)
Owns a work animal	833	0.499	0.521	-0.054**	(0.025)
Owns a water bomb	833	0.886	0.933	0.009	(0.013)
Owns tools	833	0.893	0.959	0.034**	(0.015)
Owns a sewing machine	833	0.166	0.159	-0.060**	(0.025)
Owns an oven	833	0.298	0.356	0.010	(0.032)
Livestock					
Raised cattle, bulls, calves	833	0.262	0.351	-0.002	(0.030)
Raised pigs	833	0.582	0.741	0.081***	(0.027)
Raised chicken	833	0.922	0.969	0.031**	(0.014)
Raised horses, donkeys, mules	831	0.289	0.366	-0.004	(0.025)
Raised other animals	831	0.031	0.082	0.039*	(0.022)

Table B-14: Intra-household Decision making

Variable Description	N	Control	Treatment	Difference	SE
Aggregate sole decision making by titular	877	0.412	0.371	-0.003	(0.015)
Aggregate joint decision making	877	0.193	0.314	0.083***	(0.024)
Aggregate sole decision making by husband	877	0.330	0.277	-0.049***	(0.017)
Relationship Index	730	0.654	0.717	0.044**	(0.019)

Table B-15: Social Participation

Variable Description	N	Control	Treatment	Difference	SE
Participation in social activities					
Overall likelihood of attending social activities	873	0.317	0.463	0.149***	(0.012)
Number of social activities attended	873	9.235	16.207	6.809***	(0.823)
Attended training in the community	827	0.202	0.621	0.412***	(0.028)
Attended meetings in community	827	0.511	0.655	0.151***	(0.020)
Attended community association meeting	827	0.114	0.447	0.340***	(0.035)
Attended parents' association meeting	827	0.256	0.281	0.037	(0.026)
Attended religious association meeting	827	0.287	0.291	-0.026	(0.037)
Attended school meal preparation	827	0.641	0.631	0.014	(0.025)
Number of trainings attended	873	0.822	4.427	3.470***	(0.353)
Number of community meetings attended	873	2.091	3.593	1.394***	(0.193)
Number of community association meetings attended	873	0.324	1.563	1.188***	(0.181)
Number of parents' association meetings attended	873	0.917	1.017	0.103	(0.096)
Number of religious meetings attended	873	2.011	2.504	0.623	(0.493)
Number of school meal preparation attended	873	3.070	3.104	0.030	(0.287)
Intensity of social interactions					
Talked to community leader in last week	810	0.155	0.134	-0.000	(0.021)
Talked to alcaldito in last week	765	0.053	0.055	0.013	(0.012)
Talked to health worker in last week	807	0.234	0.259	0.035	(0.028)
Talked to teachers in last week	824	0.369	0.398	0.011	(0.038)
Talked to religious leaders in last week	819	0.400	0.397	-0.031	(0.031)
Talked to relatives in last week	827	0.789	0.800	0.008	(0.023)
Talked to neighbors in last week	826	0.879	0.919	0.054***	(0.016)
Talked to others about food prices	827	0.516	0.590	0.076***	(0.026)
Talked to others about input prices	827	0.166	0.262	0.087**	(0.039)

Variable Description	N	Control	Treatment	Difference	SE
Talked to others about small businesses	827	0.217	0.291	0.080***	(0.029)
Talked to others about new agricultural practices	827	0.076	0.130	0.038**	(0.019)
Talked to others about relationships with men	827	0.150	0.195	0.033	(0.028)
Talked to others about other women issues	827	0.307	0.371	0.049	(0.039)
Trust and Social cohesion					
Degree of community cohesion (1-5)	826	3.789	4.049	0.268*	(0.144)
Degree of collaboration in community (1-5)	826	4.146	4.270	0.140*	(0.068)
Degree of trust in relatives (1-5)	826	4.249	4.288	0.069	(0.061)
Degree of trust in neighbors (1-5)	826	3.344	3.317	-0.020	(0.061)
Degree of trust in women leaders (1-5)	826	3.384	3.701	0.280***	(0.076)
Degree of trust in men leaders (1-5)	826	2.894	3.132	0.227*	(0.116)
Degree of trust in other women in communities (1-5)	826	3.130	3.208	0.065	(0.073)
Degree of trust in other men in communities (1-5)	826	2.490	2.592	0.151	(0.092)
Degree of trust in people from outside the community (1-5)	825	2.583	2.587	0.008	(0.082)

Appendix C: Questionnaire for Pilot Study

Appendix C-1: Narratives with commentary

The sentences for the narrative were chosen very strategically. We wanted to make the narratives as standardized as possible, with very subtle variation, in order to control for information as much as is possible. However, we made sure that the discrepancies in the intrinsic and extrinsic stimuli were subtle yet present. In the White Collar Intrinsic narrative, the sentence "Teachers were boring, and I struggled to stay interested in class, but I kept at it," by putting emphasis on the fact that they 'kept at it', suggests that the success of the student came from their own perseverance, as opposed to encouragement from their teachers or the school. With this method of analysis, we constructed the narrative. Because both intrinsic and extrinsic are the same aside from the job differences, I only present notes on one of each.

Appendix C-2: White Collar Intrinsic

"I was not the best at school growing up." This familiarizes the role model.

"Teachers were boring, and I struggled to stay interested in class, but I kept at it." This measures how the motivation to stay in school came from the individual.

"I surprised a lot of people when I got into college and decided to go. I wanted to become a doctor. I applied to as many medical schools as I could." This represents a possibility of poor circumstance because people were surprised that the role model got

into college. They apply to as many as they can, hinting that they are not overconfident.

“I got into one, and I’ll be starting next year.” Shows that while they applied to many, they got into one, making them more relatable. “My friends are pretty surprised that I’ll be a doctor.” Again, indicating not a very supportive environment growing up.

“I think that anyone can achieve what I have done as long as they work hard.” An ending which invites the mentee to feel that they can be confident about their future and achieve what the role model did with hard work.

Appendix C-3: White Collar Extrinsic

“I was not the best at school growing up.” This familiarizes the role model (staying same as the first one).

“Teachers were boring, and I struggled to stay interested in class, but I knew I was smart enough to succeed.” This measures the individual’s confidence level in themselves.

“I surprised a lot of people when I got into college and decided to go. I wanted to become a doctor. I applied to medical school.” This represents a possibility of poor circumstance because people were surprised that the role model got into college. However, in this sentence the role model does not specify that he/she applied to many schools, subtly indicating that they are confident in themselves to not have to apply to many.

"I got in and I am starting next year." Short and matter of fact, represents exclusivity. *"My friends are pretty surprised that I'll be a doctor."* Again, indicating not a very supportive environment growing up.

"I achieved what I did because I worked hard." An exclusive ending, not inclusive of others because the role model is only focused on themselves.

Appendix C-4: Blue Collar Intrinsic

"I was not the best at school growing up." Same as previous. All first lines are the same to embody overall background of the models.

"Teachers were boring, and I struggled to stay interested in class, but I kept at it." Measures individual's tenacity at school.

"I wanted to be a security guard. I looked at security companies and training programs, and they all required a high school diploma." This presents the attainable post-high school possibilities for the individual. A different route from college and one that should be familiar to the reader. A blue collar job that does not seem so far-fetched or unattainable.

"I got accepted to a training program. I'll be starting next year! My friends are happy for me." Role model helps the reader envision the same outcome for themselves. A program that accepts regardless of school merits (just diploma needed) and the swiftness of the role model's enrollment. The reader is able to envision their friends

being happy for them. The role model's outlook is not only encouraging but can be obtained.

"I think that anyone can achieve what I have done as long as they work hard." Role model is externally focused. Encourages reader that it is not impossible to reach for the same goals.

Appendix C-5: Blue Collar Extrinsic

"I was not the best at school growing up." Same as previous. All first lines are the same to embody overall background of the models.

"Teachers were boring, and I struggled to stay interested in class, but I knew I was smart enough to succeed." Measures role model's level of self-confidence as well as tenacity.

"I wanted to be a security guard. I looked at security companies and training programs, and all they required was a high school diploma." Role model is more self-assured in decision to become security guard; it is not just about circumstance and what is attainable.

"I got accepted to a training program." Abrupt compared to blue collar encouraging. No information about the program to give reader inspiration. Role model is unrelatable. Comes off as more self-focused.

"I achieved what I did because I worked hard." The role model's story is internally focused. The reader will not be encouraged to take the same approach as they did. The

role model is lecturing the reader just as teachers do. The mentee will compare their work ethic to role model's and theirs (the readers ethic) could be measured as less.

Appendix C-6: Pilot Questionnaire

1. How likely is it that you could do what NAME OF ROLE MODEL did?

   1 2 3 4 5 scale on a spectrum of "Very Unlikely" to "Very Likely"

2. How similar are you to NAME OF ROLE MODEL?

   1 2 3 4 5 scale on a spectrum of "Very Unlikely" to "Very Likely"

3. Would you like to do what NAME OF ROLE MODEL did?

   YES NO

4. Do you control your school performance?

   YES NO

5. Was ROLE MODEL'S NAME success due to luck?

   YES NO

6. How likely is it that SCHOOL NAME/YOUTH CLUB can help me achieve what ROLE MODEL NAME has achieved?

   1 2 3 4 5 scale on a spectrum of "Very Unlikely" to "Very Likely"

7. Is it possible to impact a situation as it happens?

   YES NO

What do you like about school/youth club? **

What do you not like about school/youth club? **

** Open response questions for focus group questions.

Appendix C-7: Final Narratives (without commentary)

White Collar Intrinsic:

I was not the best at school growing up.

Teachers were boring, and I struggled to stay interested in class, but I kept at it.

I surprised a lot of people when I got into college and decided to go. I wanted to become a doctor. I applied to as many medical schools as I could.

I got into one, and I'll be starting next year. My friends are pretty surprised that I'll be a doctor.

I think that anyone can achieve what I have done as long as they work hard.

White Collar Extrinsic:

I was not the best at school growing up.

Teachers were boring, and I struggled to stay interested in class, but I knew I was smart enough to succeed.

I surprised a lot of people when I got into college and decided to go. I wanted to become a doctor. I applied to medical school.

I got in and I am starting next year. My friends are pretty surprised that I'll be a doctor.

I achieved what I did because I worked hard.

Blue Collar Intrinsic:

I was not the best at school growing up.

Teachers were boring, and I struggled to stay interested in class, but I kept at it.

I wanted to be a security guard. I looked at security companies and training programs, and they all required a high school diploma.

I got accepted to a training program. I'll be starting next year! My friends are happy for me.

I think that anyone can achieve what I have done as long as they work hard.

Blue Collar Extrinsic:

I was not the best at school growing up.

Teachers were boring, and I struggled to stay interested in class, but I knew I was smart enough to succeed.

I wanted to be a security guard. I looked at security companies and training programs, and all they required was a high school diploma.

I got accepted to a training program. My friends are happy for me.

I achieved what I did because I worked hard.

Figure C-5-1: Sample Front Side of Card

Figure C-5-2: Sample Flipped Side of Card

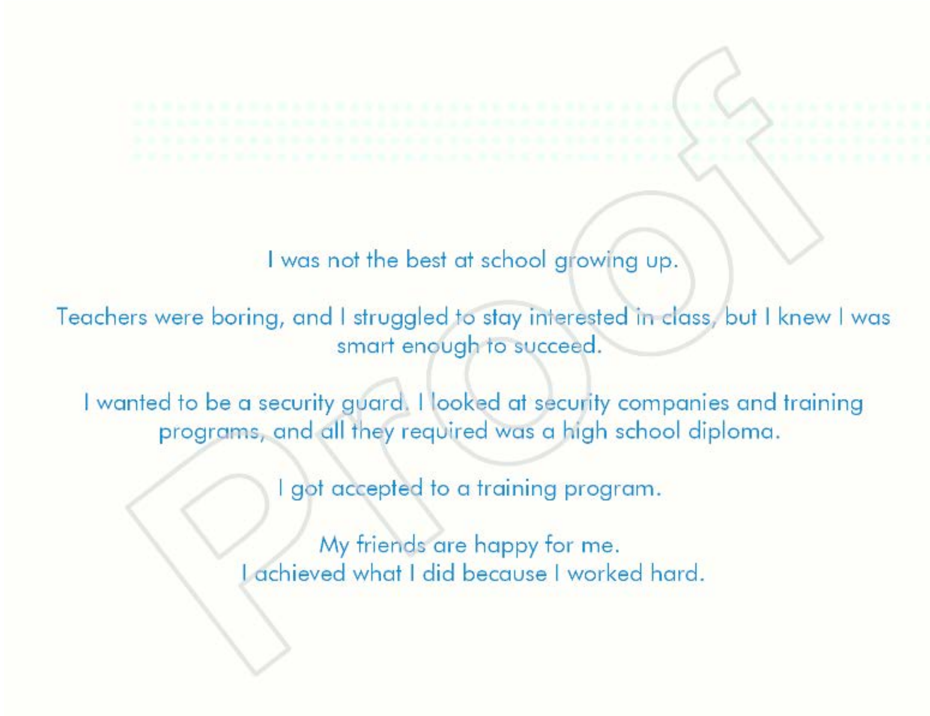


Figure C-5-3: Sample Front Side of Card

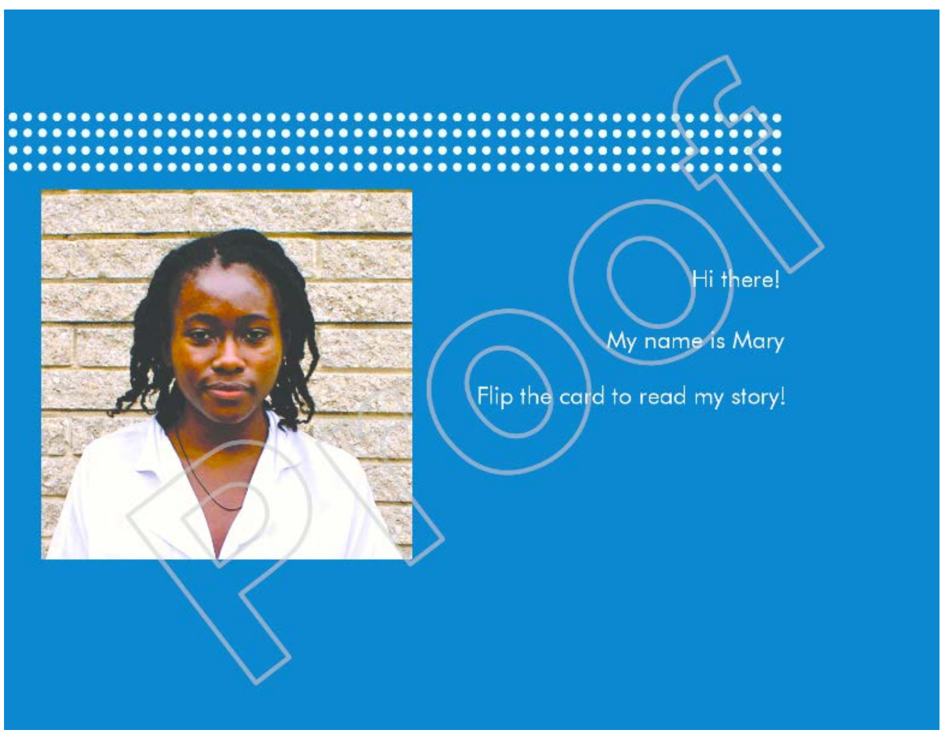
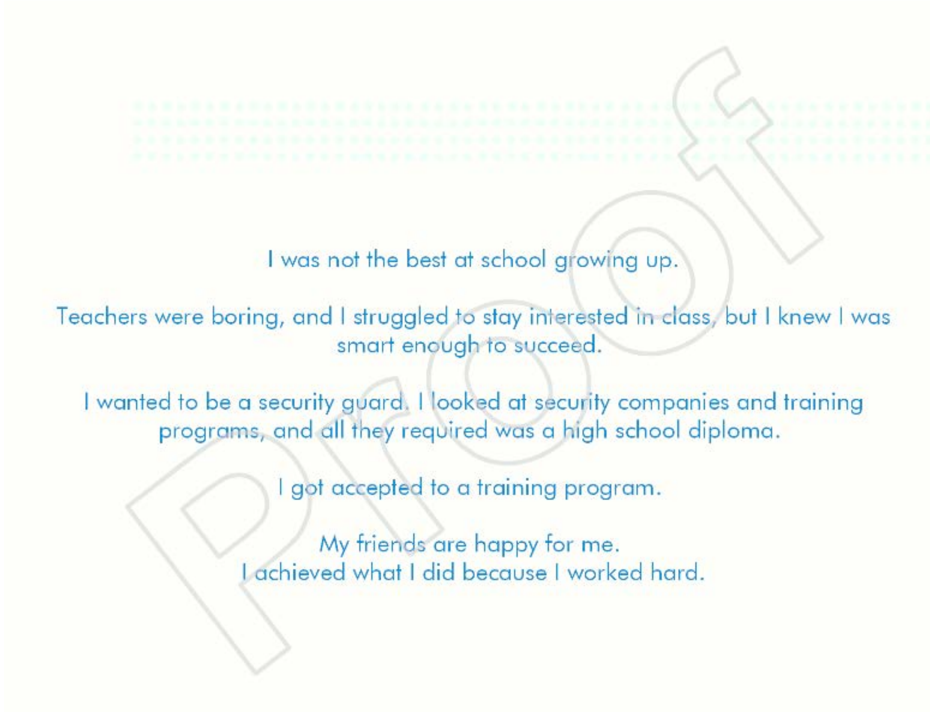


Figure C-5-4: Sample Flipped Side of Card


I was not the best at school growing up.

Teachers were boring, and I struggled to stay interested in class, but I knew I was smart enough to succeed.

I wanted to be a security guard. I looked at security companies and training programs, and all they required was a high school diploma.

I got accepted to a training program.

My friends are happy for me.
I achieved what I did because I worked hard.