The Analysis of Well-Being Using the Income-Adjusted Multidimensional Synthesis of Indicators: The Case of China

Published Version:
Mario Biggeri, Luca Bortolotti, Vincenzo Mauro (2021). The Analysis of Well-Being Using the Income-Adjusted Multidimensional Synthesis of Indicators: The Case of China. REVIEW OF INCOME AND WEALTH, 67(3), 684-704 [10.1111/roiw.12488].

Availability:
This version is available at: https://hdl.handle.net/11585/921936 since: 2023-04-04

Published:
DOI: http://doi.org/10.1111/roiw.12488

(Article begins on next page)

The Analysis of Well-being using the Income-adjusted Multidimensional Synthesis of Indicators: the Case of China

Mario Biggeri*, Luca Bortolotti* and Vincenzo Mauro°

** University of Florence, ° University of Pisa*

Abstract

In multidimensional indexes employed to measure well-being and deprivation, income is sometimes included and sometimes excluded. The aim of this paper is to reconsider the role of income in the measurement of multidimensional well-being by recognizing that it can indirectly contribute to individual well-being, even if it is not regarded as a goal in itself. This involves introducing a new composite index: the Income-adjusted Multidimensional Synthesis of Indicators (*I-MSI*). To illustrate this index, individual-level data from the 2015 China Household and Nutrition Survey (CHNS) are analyzed. Results confirm the soundness of *I-MSI* approach as a multidimensional aggregation method and show that it can capture disparities across Chinese macro-regions and variations among different segments of society.

JEL: I31, O53

Keywords: Multidimensional well-being, Income-adjusted Multidimensional Synthesis of Indicators, China.

1. Prof. Mario Biggeri, Ph.D, Department of Economics and Management, University of Florence.
Orcid 0000-0001-7986-380X

2. Luca Bortolotti, PhD; Department of Economics and Management, University of Florence.
Orcid 0000-0003-2632-6677

3. Vincenzo Mauro, Ph.D; Department of Economics, University of Pisa.
Orcid 0000-0001-9059-855X

1 Introduction

In recent decades, academic research has increasingly looked beyond traditional income-based measures of well-being and deprivation to embrace a multidimensional perspective (Maasoumi, 1986; Sen, 1999; Chakravarty and D'Ambrosio, 2006; Alkire, and Foster, 2011; Bossert *et al*., 2013; Mazziotta and Pareto, 2013; see Burchi et al., 2018 for a review).[1] Three main issues emerge in the debate on how to build appropriate composite indexes: firstly the selection of dimensions, secondly the choice of weights and, thirdly, the choice of the aggregation method, which is the main focus of this paper. The selection of dimensions is dealt with in Section 3; while the choice of weights, traditionally based on normative assumptions or empirical approximations, is simplified by adopting equal weights. However, it is important to acknowledge that different weighting systems can be applied. For instance, following Maasoumi (1986) it is possible to use data to weigh factors of well-being (empirical weights). For a review of this approach, see also Decancq and Lugo (2013). For Maasoumi and Racine (2016) heterogeneous weights are found *ex-post*, based on the distribution of outcomes across different populations in their sample. In contrast, Alkire and Foster (2011) defend the use of equal weighting, thus adopting a normative approach.[2]

Another strand of literature focuses on the role of income in measuring multidimensional well-being (Maasoumi, 1986; Anand and Sen, 1997; Kovacevic, 2010). Income is sometimes included and in other cases excluded from the selection of relevant multidimensional indicators depending on whether it is viewed as a means or as a goal in its own right (Sen, 1999). For instance, income has always been a component in the Human Development Index (HDI) as a proxy for living standards. Conversely, in the Multidimensional Poverty Index (MPI) and Multiple Overlapping Deprivation Analysis (MODA), assets, rather than income, are taken into account. Nevertheless, the role of

---

[1] A similar debate took place in several international institutions and forums from the end of the 1980s including the UNDP, FAO and WFP. More recently, the Commission on the Measurement of Economic Performance and Social Progress (Stiglitz *et al*., 2009), the 2030 Agenda for Sustainable Development (SDG 1.3), the OECD Better Life Index. Moreover, some National Statistical Offices such as CONEVAL in Mexico, ISTAT in Italy and Eurostat in the EU, have developed multidimensional measures of well-being that include the income dimension.

[2] It is beyond the scope of this paper to consider empirical weights within our approach. Future investigations may consider combining our technique with empirical weights so that preferences for different dimensions may vary across different groups of individuals.

income with respect to well-being is still unresolved and open to debate (Laderchi, 2003; Pogge, 2010).

In terms of the aggregation method, two distinct approaches are usually considered. One approach provides a continuous measure of multidimensional well-being, eventually resorting to arithmetic means across dimensions or more sophisticated operations that penalize the heterogeneity of outcomes across dimensions. A well-known example is the post-2010 HDI, which now adopts the geometric mean (Klugman *et al*., 2011). The rationale is to make an allowance for greater homogeneity in outcomes across dimensions as a valuable achievement in itself. To improve these measures, Mauro *et al*. (2018) suggest a new aggregation method to address problems associated with the geometric mean (Klugman *et al*., 2011). The second approach utilises a binary distinction between multidimensionally poor and non-poor individuals. This is obtained by considering a set of deprivation indicators that can be aggregated through the union approach, the intersection approach, or more often, using a dual cut-off procedure. This procedure has been used for measuring social exclusion (Chakravarty and D'Ambrosio, 2006), estimating the overlap in deprivations among children (Gordon *et al*., 2003 and UNICEF MODA, de Neubourg et al., 2013) and other multidimensional poverty indexes (e.g. MPI, Alkire and Foster, 2011).

The objective of this paper is to try to reconcile the role of income in the measurement of multidimensional well-being by incorporating it as a means to other ends and not as an end in itself, and by introducing a new composite index: The Income-adjusted Multidimensional Synthesis of Indicators (***I-MSI***).

The ***I-MSI*** is linked to the Multidimensional Synthesis of Indicators (***MSI***) approach introduced by Mauro *et al*. (2018) and elaborates on Bourguignon and Chakravarty's (2003) seminal contribution. The ***I-MSI*** introduces a new formula to elaborate on the role of income while maintaining the properties of the ***MSI*.** The ***MSI*** respects key properties such as full sensitiveness (monotonicity), continuity, and flexibility in the structure of substitutability between outcomes. These properties recognize that: the aggregation procedure must take into account the heterogeneity among indicators; each improvement (or deterioration) in any individual unidimensional indicator of well-being results in an increase (or decrease) in the overall index; and a "small" change in any indicator is associated with similar "small" changes in the overall index (reducing discontinuities).[3] With respect to the latter, the degree of substitutability between dimensions is defined through a function of individual characteristics (see also Bourguignon and Chakravarti, 2003). In other words, the ***MSI*** index takes into account the heterogeneity of outcomes by allowing compensation across dimensions to vary in accordance with a person's individual characteristics. In this respect, our approach helps reduce the so-called degree of "inexplicable arbitrariness" entrenched in the substitutability between indicators (Anand and Sen, 1997, p. 17; see also Mauro *et al*., 2018). In addition, our approach overcomes the tendency of the geometric mean to collapse to zero in cases where performance in one or more dimensions is close to zero (Mauro *et al*., 2018). The ***MSI*** also remains robust at each level of heterogeneity and for every overall level of achievement compared to other aggregation functions (see Appendix A). The ***MSI*** can select either household/individuals or countries/regions as the unit of analysis using micro- or macro-data respectively. So far, it has been only applied at the cross-country level (Biggeri and Mauro, 2018; Biggeri *et al*., 2019; Biggeri and Bortolotti, 2020); it follows, therefore, that the application of our ***MSI***-adjusted technique to micro-data is a contribution in itself.

In order to illustrate the ***I-MSI*** approach to measure well-being, China is taken as a case study for the following four reasons. Firstly, China has played a central role in worldwide poverty reduction. The second, related reason, is the surge in inequalities, which includes growing income disparities as well as other inequalities emerging at the individual and household levels (Ward, 2014; Nicholas *et al*., 2017; Biggeri and Bortolotti 2020). Thirdly, Chinese policymakers are increasingly interested in multidimensionality (Feng et al., 2016; Xue *et al*., 2018). The fourth reason is the availability of micro-data — specifically the 2015 China Health and Nutrition Survey (CHNS).

This paper is divided into five sections. Following the introduction, Section 2 introduces the role of income in multidimensional well-being and presents the ***I-MSI*** index (see also Appendix A). Section 3 describes the illustrative case of China and the dataset used in the analysis (further details can be found in Appendix B). Section 4 reports the main results produced by the ***I-MSI***, comparing different sections of the population, other methods of aggregation, and investigating correlations with subjective well-being. Finally, Section 5 draws the main conclusions.

---

[3] This allows to compensate for well-being changes across dimensions in a flexible manner. In this respect, the *MSI* represents an alternative to the assumption of full substitutability between dimensions (see also Mazziotta and Pareto, 2013)—an implicit idea in indexes based on the arithmetic mean.

2 The Role of Income in Multidimensional Well-Being and the I-MSI

Attempts to create a multidimensional index of well-being stem from the idea that a single indicator (even if as relevant as income), cannot serve as an adequate proxy for all aspects of human development. This contention has spurred an enormous literature in economics and development studies (Streeten, 1981; Anand and Sen, 1997; Bourguignon and Chakravarty, 2003; Alkire and Foster, 2011; Bossert *et al*., 2013; Mazziotta and Pareto, 2013; Maasoumi and Racine, 2016; Mauro *et al*., 2018). If, for example, income is included as a component of a multidimensional index (which is the case with the HDI), it may well be regarded as a proxy for other "unmeasurable" aspects of well-being rather than a goal in itself (e.g. Sen, 2007; for further debate see Kovacevic, 2010). It follows that any comprehensive attempt to tackle deprivation[4] cannot focus exclusively on income. Meeting basic needs in health, education, social security and so on and so forth is widely recognized as an essential component of any poverty mitigation strategy.[5]

The idea of "looking beyond" income towards multidimensionality and considering income as a means, have become a milestone in the literature on the Capability Approach to multidimensional well-being and poverty analysis. According to Sen (1999), income has a fundamental instrumental role, because it can expand the capabilities (opportunities/abilities) of individuals, but it is not an end in itself. In other words, in the Capability Approach, income is relevant for well-being but is not an intrinsically valuable end. For this reason, it should not count as a dimension in the multidimensional analysis of well-being (Deneulin and Shahani, 2009; Alkire and Foster, 2011), one exception being in the analysis of children well-being and poverty (Sen, 2005; Trani *et al*., 2013).

---

[4] The analysis of multidimensional poverty has evolved from distinguishing the multidimensional poor and the multidimensional non-poor, just as monetary poverty lines sharply distinguish the income poor from the income non-poor (Bourguignon and Chakravarty, 2003; Alkire and Foster, 2011). This strategy provides an immediate indication of who is poor but neglects the importance of any progress above the poverty line (Sen, 1976; Clark and Hulme, 2010, p. 349). While in a continuous poverty scale every improvement in well-being contributes to alleviate the multidimensional deprivations (Burchi *et al*., 2018).

[5] Although there are many concepts and measure of poverty, we follow the 2030 Sustainable Development Agenda, which aims to "end poverty in all its forms everywhere" as one of the first SDG targets: "Poverty entails more than the lack of income and productive resources to ensure sustainable livelihoods. Its manifestations include hunger and malnutrition, limited access to education and other basic services, social discrimination and exclusion as well as the lack of participation in decision-making. Various social groups bear disproportionate burden of poverty." Source: https://www.un.org/development/desa/dspd/poverty-social-policy-and-development-division.html.

Generally, existing multidimensional indexes of poverty or well-being in the economic literature either (1) assign income the same theoretical and computational role as other components or dimensions, or (2) exclude income from the metric altogether.[6] This paper tries to address the inclusion of income in the measurement of multidimensional well-being by emphasizing its instrumental role. In terms of the ***I-MSI***, if two persons achieve the same level in each dimension, but one person has a higher income (and thus greater opportunity to choose between capabilities), this person will reach a higher well-being score. Although income may not directly contribute to well-being (as many scholars argue), it continues to play an important role within the ***I-MSI*** by facilitating overall well-being. In this respect, the ***I-MSI*** helps to reconcile the aforementioned ongoing debates on the role of income.

2.1 The Income-Adjusted Multidimensional Synthesis of Indicators (I-MSI)

The ***I-MSI*** is an original composite index which assigns income a relevant role within the multidimensional synthesis of well-being without including it among the well-being components. It exploits the high degree of generalizability of the ***MSI*** formula in which (the transformation of) income plays an instrumental role.

Following Mauro *et al.* (2018), the general formulation of the ***MSI*** uses a standard data matrix, with n observations and k dimensions. Each outcome represents one of the surveyed dimensions. For a generic individual i, the ***MSI*** score is:

$$MSI_i = 1 - \left[\frac{1}{k}\sum_j (1 - x_{ij})^{g(x_i)}\right]^{\frac{1}{g(x_i)}} \qquad (1)$$

where the generic entry x_{ij}, is a continuous number bounded between $0 \leq x_{ij} \leq 1$, which measures the outcome of individual i in dimension j. As shown in Appendix B, $x_{ij} = 0$ corresponds to the lowest level of unidimensional well-being and $x_{ij} = 1$ to the highest level. The vector of outcomes x_i is aggregated into the value $\boldsymbol{MSI_i}$ adopting a function $g(x_i) \geq 1$, indicating to what extent individual i can substitute different dimensions. In fact, $g_i \rightarrow \infty$ corresponds to the complete inability to compensate for well-being in one dimension by drawing on others; and $g(x_i) = 1$ corresponds to a prefect "substitutability" of well-being among dimensions for individual i (equivalent to the

[6] More generally, in the composite index literature (Maasoumi and d Nickelsburg, 1988; Decancq and Lugo, 2013), the various dimensions considered are typically treated equally. No theoretical differences between income and other dimensions are usually introduced *ex-ante*.

arithmetic mean between x_i). The introduction of "thresholds," named *a* and *b* (Mauro *et al*., 2018), can prevent function $g(x_i)$ from assuming extreme values by setting an appropriate range in which the substitutability of dimensions varies (Biggeri *et al*., 2019). The simplest form of the index, as suggested by Mauro *et al*. (2018), defines the function g_i as the inverse of the average level of the constituent components, i.e. $g_i = \left(\frac{1}{k}\sum_{j=1}^{k} x_{ij}\right)^{-1}$.

The ***I-MSI*** is obtained by substituting the parameter $g(x_i)$ in equation (1) according to the transformation of income of individual i, y_i (which is not included in the *k* dimensions). In this respect, income assumes an instrumental role within the *MSI* formula. Equation (2) describes the ***I-MSI***.

$$I - MSI_i = 1 - \left[\frac{1}{k}\sum_{j=1}^{k}\left(1 - x_{ij}\right)^{\frac{1}{f_i}}\right]^{f_i} \qquad (2)$$

$$where\ f_i = \begin{cases} \frac{\ln(a)}{\ln(b)} & if y_i < a \\ \frac{\ln(y_i a)}{\ln(b)} & if a \leq y_i < b \\ 1 & if\ y_i \geq b \end{cases} \qquad (3)$$

Equation (3) introduces income y_i and transforms it in a value bounded between 0 and 1 that determines substitutability across dimensions: the higher the income, the higher the substitutability across dimensions. In this function, two thresholds, *a* and *b*, correspond to the levels of income for which substitutability reaches its minimum and its maximum, respectively. The rationale behind the choice of these two thresholds is the attempt to avoid extreme situations (i.e. zero or infinite substitutability) that could bias the results. To account for the subjectivity that is inevitably induced by selecting the two parameters *a* and *b*, a sensitivity analysis was conducted to test the robustness of the results, and the latter confirmed the soundness of the index.

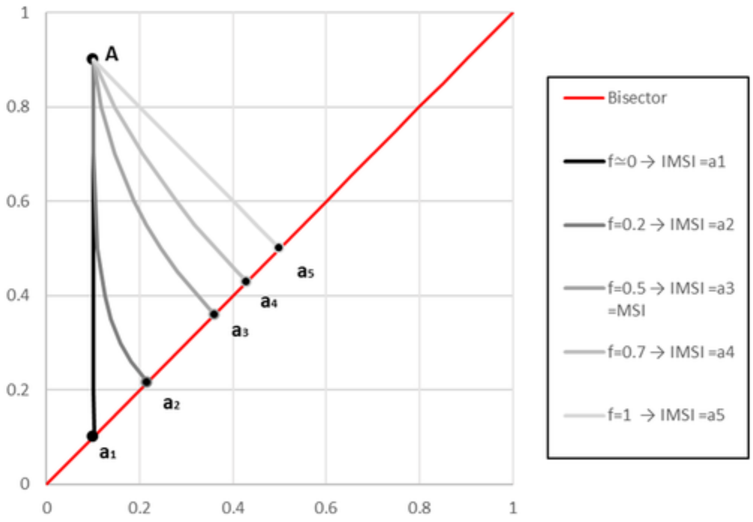
Equations (1) and (2) refer to each individual. The aggregation of all the individual ***I-MSI*** scores into a single value (vertical aggregation, i.e. between individuals) is beyond the scope of our analysis. In this paper, however, the differences between social groups are calculated comparing their average ***I-MSI*** scores. Indeed, for each group we computed the average outcomes $\overline{I - MSI} = \frac{1}{m}\sum_{i=1}^{m} I - MSI_i$, where individuals from 1 to m are the only individuals belonging to the given group.[7]

[7] Alternatively, vertical aggregation can operate through more elaborated aggregation functions including the *MSI* itself (Biggeri and Mauro, 2018) or the *I-MSI*.

The ***I-MSI*** theoretically differs from the ***MSI*** indexes presented so far and represents a contribution to the debate on how the income dimension could be incorporated within a multidimensional index.

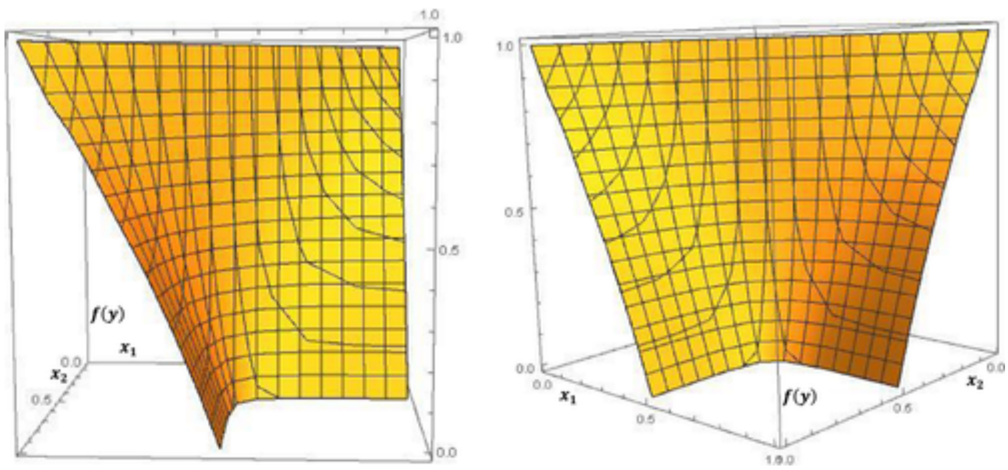
To illustrate the behavior of the ***I-MSI*** index, a two-dimension case (e.g. health and education) having income as a mean is considered. Figure 1 shows the projection in the bisector of the ***I-MSI*** for different levels of income for individual A considering a high heterogeneity between two dimensions, for instance A (0.1; 0.9). In the ***I-MSI***, in contrast to the ***MSI***, the coordinates of point A provide a *range* of infinite possible outcomes (from point a_1 to point a_5), depending on the individual average score of the dimensions and their heterogeneity, while the income level determines the exact point in the bisector in the range (a_1:a_5) which corresponds to ***I-MSI***. In other words, the ***I-MSI*** value varies according to the function of income f_i: poorer individuals have higher penalization rates than richer individuals. At point a_1, with income approaching 0, point A identifies with ***I-MSI = 0.1*** (i.e. the minimum across dimensions). Conversely, if income is close to 1 (point a_5) the ***I-MSI*** converges on the arithmetic mean. In case f_i =0.5 is equal to the arithmetic mean, the ***MSI*** and ***I-MSI*** coincide. For individuals such as A, with highly heterogeneous outcomes (which are distant from the bisector), the capacity to compensate for well-being across dimensions through income becomes crucial. In contrast, individuals closer to the bisector (with more homogeneous outcomes) are less affected by the level of income.

Figure 1 - *I-MSI* Scores and Income



A three-dimensional illustration of the ***I-MSI*** is presented in Figure 2. This figure shows an iso-quant surface for a fixed level of the index (***I-MSI* = *0.5***) in the synthesis between two dimensions (on axis x and axis y respectively) according to function f_i (on axis z).

Figure 2 - Iso-Quant *I-MSI* = 0.5 Surface



Note: A transformation of income $f(y)$ is measured on the vertical axis and contributes to determine the synthesis between dimension x_1 and x_2.

To better clarify the mechanism behind the ***I-MSI***, we consider two individuals with the same score in each of the k dimensions and different incomes, namely an income poor and a non-income poor. Supposing, for the sake of simplicity, that there is a negative shock in the asset dimension (e.g. their cars break down), the rich individual will have more instruments to absorb the shock (i.e. greater resilience) and will be better placed to maintain outcomes in other dimensions (such as leisure, work and health, etc.). In short, the rich individual, can count on more coping strategies. The same reasoning suggests that the deterioration of well-being in any given dimension will result in a more serious drop in overall well-being of the income poor individual (i.e. less resilience to maintain overall functioning).$Digitare\ l'equazione\ qui.$

The properties of the ***I-MSI*** can be summarized as follows:

1. *Heterogeneity Penalization*. For any fixed level of income, and any average level of well-being, the ***I-MSI*** is a continuous and strictly monotonic function of variability across dimensions. Any increase in the level of heterogeneity produces greater penalization of the overall score.[8]
2. *Income Poverty Penalization*. The index is a continuous and strictly monotonic function of the income for units with a positive variance of the vector of outcomes x_i. In other words, if the heterogeneity across dimensions is not zero, a reduction in income will increase the penalization of the overall score. An income increase helps a person to get closer to the "full potential" of well-being obtained from specific dimensions as point a_5 illustrates (Figure 1).
3. *Full Well-being Satisfaction*. The highest possible level of the ***I-MSI*** (***I-MSI = 1***) can only be achieved when all non-income dimensions are fully satisfied ($x_{ij} = 1 \forall j$), irrespectively of the income level.
4. *Full Well-being Deprivation*. The lowest possible level of ***I-MSI*** (***I-MSI = 0***) can be achieved when at least one dimension is 0 *and* income is 0 (as well as in the trivial case of $x_{ij} = 0 \forall j$, irrespective of the income level).[9]
5. *Strict **I-MSI** monotonicity*. For any given level of income where $f_i > 0$, any increase in the level of any component results in an increase in the overall ***I-MSI*** level, provided $a > 0$.[10]

3 The Chinese Multidimensional Well-Being: Data, Dimensions and Indicators

In order to operationalize the ***I-MSI*** index, as anticipated in the introduction, China has been selected as a case study for four reasons. First, thanks to well-known reforms implemented since 1978, the

[8] Penalization is lower for income rich individuals. At the extreme, a rich individual with $f_i = 1$ is not affected by heterogeneity of non-income outcomes; instead the aggregation index only reflects the average score across dimensions.

[9] This means that $f_i \neq 0$; otherwise a value ***I-MSI*** $= 0$ occurs whenever an individual is fully deprived in at least one dimension. This is linked the properties of strict and weak monotonicity.

[10] As in the previous footnote, this means that $f_i \neq 0$; otherwise monotonicity is weak rather than strict.

Chinese development has been a major driver of global income poverty reduction. Second, the presence of strong heterogeneity across regions, social groups and different dimensions of well-being helps bring strong evidence to the strengths of the ***I-MSI***. In the Chinese case, this heterogeneity is due to prevailing resource endowments as well as pragmatic "trial and error" policies and the unbalanced approach of reforms (Goodman and Segal, 2002; Ravallion and Chen, 2007; Shue and Wong, 2007). Third, Chinese policy makers are increasingly interested in multidimensionality. Under the presidency of Hu Jintao (2003–2013), China adopted a strategy leading to a "harmonious society," where income growth should be balanced by social and environmental development (Li *et al*., 2016). Finally, the China Household and Nutrition Survey (CHNS) recently released reliable micro-data for 2015 allowing for a detailed analysis.

The economic literature on Chinese multidimensional development (Labar and Bresson, 2011; Ray and Mishra, 2012; Alkire and Shen, 2017; Bin, 2016; Feng *et al*., 2016; Nicholas *et al*., 2017; You *et al*., 2018) has often adopted a dichotomous (and therefore discontinuous) notion of poverty where individuals are either regarded as "poor" or "non-poor," rather than opting for a measure where the notion of poverty is defined as a continuous (latent) variable. According to this literature, inequalities have expanded, and multidimensional poverty persists despite economic growth (Sicular *et al*., 2007; Ray and Mishra, 2012; Li *et al*., 2013; Ward, 2014; Alkire and Shen, 2017). A recent study by Nicholas *et al*. (2017), for example, seeks to decompose poverty according to the proportion of deprivation attributable to specific time-periods and particular dimensions of poverty. This study identifies large urban–rural disparities, coastal/inner gaps, high blood pressure levels, lack of education and drinking water issues as key drivers of multidimensional and recurring poverty. An issue not adequately investigated from a multidimensional perspective to date relates to the internal migration of rural citizens without urban household registration ("hukou"), having limited access to goods and services and generally enjoying fewer rights and lower incomes and wages (Knight and Gunatilaka, 2010; Pakrashi and Frijters, 2017; Combes *et al*., 2019).

To investigate multidimensional well-being in China from a "harmonious society" perspective, the ***I-MSI*** embraces the following four principles: (1) the homogeneity of outcomes is adopted over unbalanced outcomes (reflected in a flexible structure of substitutability), (2) the granular sensitivity to the diverse conditions affecting different social groups is adopted instead of utilizing a dichotomous distinction between deprived and non-deprived individuals (reflected in the continuity property), (3) individual outcomes are included to capture inequalities within the household because individuals are chosen as units of analysis, (4) well-being is directly affected by non-monetary outcomes which are valuable in themselves (i.e. achieved functionings), and indirectly affected by income (because of the income poverty penalization).

3.1 Data, Dimensions and Indicators

The empirical analysis is based on the CHNS. This database is the result of a program run by the University of North Carolina at Chapel Hill and the Chinese Center for Disease Control and Prevention. Since 1989, the CHNS has collected panel data using a multistage random cluster sampling process. In 2015, the number of individuals surveyed reached 16,622 units covering 12 of the 31 provinces: Beijing, Jiangsu, Liaoning, Shandong, Shanghai (East region), Heilongjiang, Henan, Hubei, Hunan (Centre), Chongqing, Guangxi, and Guizhou (West). The sample is balanced and covers the rural and urban areas of the three macro-regions that constitute Eastern, Central and Western China including three (of the four) directly controlled municipalities (i.e. Beijing, Shanghai and Chongqing).

The CHNS database has been used in several studies and contains information relevant for different fields, including some analyses of multidimensional poverty and well-being (Labar and Bresson, 2011; Qi and Wu, 2015; Nicholas *et al*., 2017). The 2015 dataset allows us to obtain information on several individual and household well-being domains and other characteristics such as gender, age, residence area (rural/urban, East/Centre/West region), and the possession of internal urban passports (hukou). This information is essential in order to analyze disparities between geographic regions and differences between groups of individuals using the ***I-MSI***.

The first step for building a multidimensional indicator is the selection of relevant dimensions. In our case, the choice of dimensions has been guided by previous studies (Clark, 2002; Nussbaum, 2003; Alkire and Foster, 2011; Frediani *et al*., 2019) with particular attention to the Chinese case (Alkire and Shen, 2017) and the components of harmonious development (Li *et al*., 2016). This led to the selection of eight distinct dimensions of well-being: Education, Health, Nutrition, Housing, Sanitation, Assets, Work, and Leisure. The number of dimensions included in composite indexes is critical. If the number is “too low,” the measure may not adequately capture well-being broadly defined; if the number of dimensions is “too high,” the explanatory power of the synthetic value may be reduced (the so-called curse of multidimensionality).

Each of those dimensions was then related to an individual outcome, based on the selection of pertinent variables (Atkinson, 2002; see Appendix B for details). Whenever possible, for the methodological reasons mentioned above, we gave priority to individual level (rather than household-level) variables.

To guarantee comparability between the observed units, we restricted our sample to 13,839 units consisting of adult individuals (18 years old or older) who had completed at least three of the five relevant subsections of the questionnaire (education, physical exams, health, physical activity and jobs).

Table 6 in Appendix C summarizes the eight dimensions included in the ***I-MSI***. The sample contains less than 3 percent of missing data. Missing data were simulated using Multiple Imputation by Chained Equations (MICE) in conjunction with the original data, and clustered according to the three macro-regions and urban–rural location.

Each of the dimensions is transformed to range from “0” (minimum well-being) to “1” (maximum well-being) and is computed from different and non-overlapping groups of variables (see Appendix B for more details).

Table 1 presents the correlation matrix for the eight dimensions selected plus the logarithm of income per capita and a self-reported quality of life index collected in the CHNS survey. Although there are some exceptions, the eight dimensions are in general positively and significantly correlated. They are also positively and significantly correlated with income (except for health, negatively correlated) and with self-reported quality of life (except for nutrition, not significantly correlated).[11] We point out that the positive correlation between the dimensions (and, thus, their average $\boldsymbol{\mu_i}$) and income $\boldsymbol{y_i}$ makes the ***I-MSI*** consistent with the ***MSI*** indexes, as the most vulnerable individuals to heterogeneous outcomes record both generally low $\boldsymbol{\mu_i}$ and $\boldsymbol{y_i}$. Furthermore, it is observed that the standard deviation across the dimension outcomes is negatively correlated with income: in fact, richer individuals tend to have higher and more homogeneous performance in the eight dimensions.

[11]The negative correlation could be due to the severity of pollution and chaotic lifestyles in cities (including stressful and intense jobs). Moreover, negative correlations can hinder omitted variables such as age. Finally, given that health is based on self-reported items (number and severity of various types of diseases), there could be bias if awareness differs between urban and rural dwellers or different socio-economic groups.

Table 1
Correlation Matrix across dimensions and correlation with income and life quality

	Educ.	Health	Nutr.	Hous.	Sanit.	Assets	Work	Lei.	Inc.	L.Q.
Education									0.20***	0.19***
Health	0.13***								-0.03***	0.06***
Nutrition	0.00	0.04***							0.03***	0.01
Housing	-0.14***	-0.05***	0.01						0.02*	0.02**
Sanitation	0.28***	-0.03***	0.02**	-0.08***					0.16***	0.11***
Assets	0.21***	0.02***	-0.00	-0.01	0.10***				0.15***	0.18***
Work	0.37***	0.06***	0.02**	-0.09***	0.15***	0.16***			0.31***	0.13***
Leisure	0.16***	0.10***	-0.01	-0.02***	0.03***	0.05***	0.08***		0.05***	0.08***

Note: *, ** and *** represent significance levels of 10%, 5% and 1% respectively.

The Pareto dominance criterion, adopted in some previous studies to evaluate and rank multidimensional well-being (Labar and Bresson, 2011), is not applicable here, because of the quantity and variety of the dimensions analyzed. Indeed, groups who are disadvantaged in some dimensions perform better in others. For example, rural people are, on average, poorer in education and sanitation but enjoy larger houses. Hence, a synthetic measurement of well-being is needed to compare the levels of well-being in a clearer and more immediate manner.

4 Empirical Results

The empirical results are organized as follows. Section 4.1 reports on the results of the ***I-MSI*** analysis and includes some comparisons with other aggregation methods. Section 4.2 describes inequalities across different regions and groups. Finally, Section 4.3 compares how the ***I-MSI*** correlates with self-reported quality of life.

4.1 The I-MSI Results and the Role of Income

Table 2 reports the results of the ***I-MSI*** for China and for the 12 provinces. The arithmetic and geometric means and the standard ***MSI*** aggregation method are reported for comparison purposes. This table also reports the standard deviation, the minimum value, the 1st quartile, the median, the 3rd quartile and the maximum value.

Table 2
China and provincial multidimensional well-being through the I-MSI

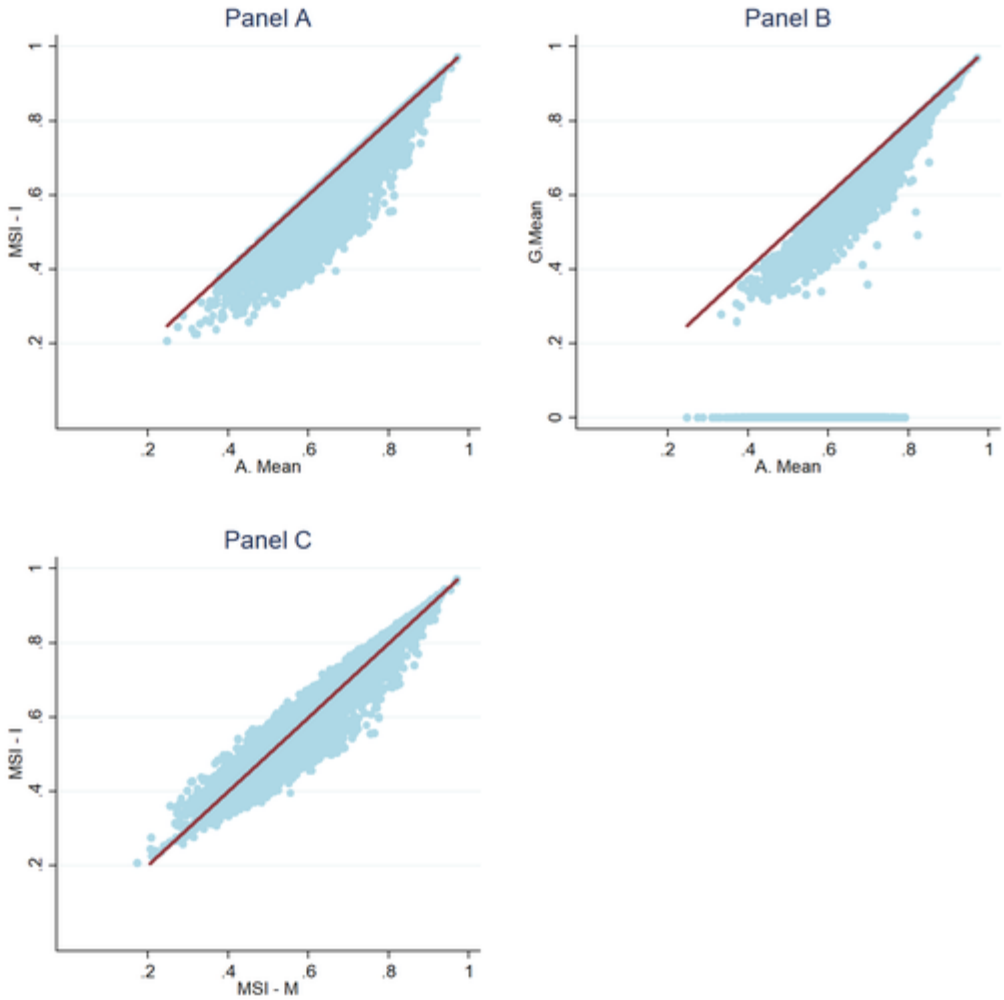
	Obs	Mean	Std. Dev.	Min	I Quart.	II Quart.	III Quart.	Max
Arithmetic Mean	13839	0.70	0.10	0.25	0.64	0.71	0.77	0.97
Geometric Mean	13839	0.59	0.24	0.00	0.55	0.66	0.74	0.97
MSI	13839	0.66	0.12	0.17	0.58	0.67	0.75	0.97
I-MSI	13839	0.66	0.12	0.21	0.58	0.68	0.75	0.97
I-MSI Beijing	1066	0.73	0.09	0.29	0.67	0.75	0.80	0.94
I-MSI Shanghai	1177	0.72	0.08	0.39	0.66	0.73	0.78	0.90
I-MSI Liaoning	1108	0.68	0.10	0.30	0.63	0.70	0.75	0.94
I-MSI Jiangsu	1160	0.67	0.11	0.26	0.60	0.69	0.76	0.90
I-MSI Hubei	1356	0.67	0.12	0.23	0.59	0.69	0.75	0.92
I-MSI Shandong	1109	0.66	0.13	0.24	0.57	0.68	0.76	0.96
I-MSI Guangxi	1315	0.65	0.11	0.30	0.59	0.67	0.73	0.93
I-MSI Heilong.	1005	0.64	0.12	0.24	0.55	0.66	0.74	0.92
I-MSI Guizhou	1164	0.64	0.13	0.27	0.55	0.66	0.74	0.92
I-MSI Hunan	1120	0.64	0.12	0.30	0.55	0.66	0.74	0.97
I-MSI Chongqing	1081	0.63	0.12	0.21	0.54	0.64	0.72	0.90
I-MSI Henan	1178	0.61	0.12	0.25	0.51	0.60	0.70	0.91

The scores emphasize strong disparities among provinces as well as for the overall aggregated results (the provinces are ordered according to their average ***I-MSI*** scores).

As expected, ***I-MSI*** and ***MSI*** are slightly lower than the arithmetic mean due to heterogeneity penalization. In fact, with lower outcomes, especially amongst the poorest quintile, the score is lower because of the overlap among different deprivations. Aggregation through the geometric mean involves the greatest penalization and is characterized by a much higher standard deviation amongst individuals. Both these results reflect the fact that the geometric mean across individual components tends to collapse to zero or close to zero whenever one or more of its components approaches zero. In our case, there are 1,651 individuals (11.9 percent of the sample) whose geometric means collapse to zero. These individuals are a heterogenous group and include extremely poor people (in terms of income and average score across dimensions) as well as people who are deprived in only one or two dimensions and are doing reasonably well in other dimensions. As already mentioned in Section 2, both the ***I-MSI*** and ***MSI*** overcome this problem (which is a particularly serious problem when aggregating micro-data).

Figure 3 compares aggregates scores based on the arithmetic mean with the ***I-MSI*** (Panel A), the geometric mean (Panel B) and a comparison between ***I-MSI*** and ***MSI*** (Panel C). Each point on the graph corresponds to one of the 13,839 observations, while the solid line represents the bisector. Every point would fall on this line in the absence of heterogeneity. These graphs show how the penalization of ***I-MSI*** and the geometric mean behave compared to the arithmetic mean benchmark characterized by perfect substitution across dimensions (no penalization). The most notable difference between the two panels is the presence of several individuals (some with a high arithmetic mean value) whose geometric means collapse to zero in Panel B.

Figure 3 - Penalization of Heterogeneity Through the *I-MSI* (Panel A) and the Geometric Mean (Panel B) and a Comparison With *MSI* (Panel C)



Panel C compares the scores of ***I-MSI*** and ***MSI***. Despite, as mentioned above, the fact that the two aggregation functions have similar penalizations (see the strong positive correlation in Panel C), some relevant differences emerge relative to the distance between f_i and μ_i (i.e. between the transformation of income and the average outcome of dimensions), which depend on the characteristics of the individuals. Specifically, when ***I-MSI*** > ***MSI*** (i.e. $f_i > \mu_i$), ***I-MSI*** penalization is lower because the index considers the instrumental role of their income. These individuals are graphically represented by the dots above the bisector in Panel C. As expected, such cases are more frequent among individuals who are closer to the origin (with low levels of μ_i).

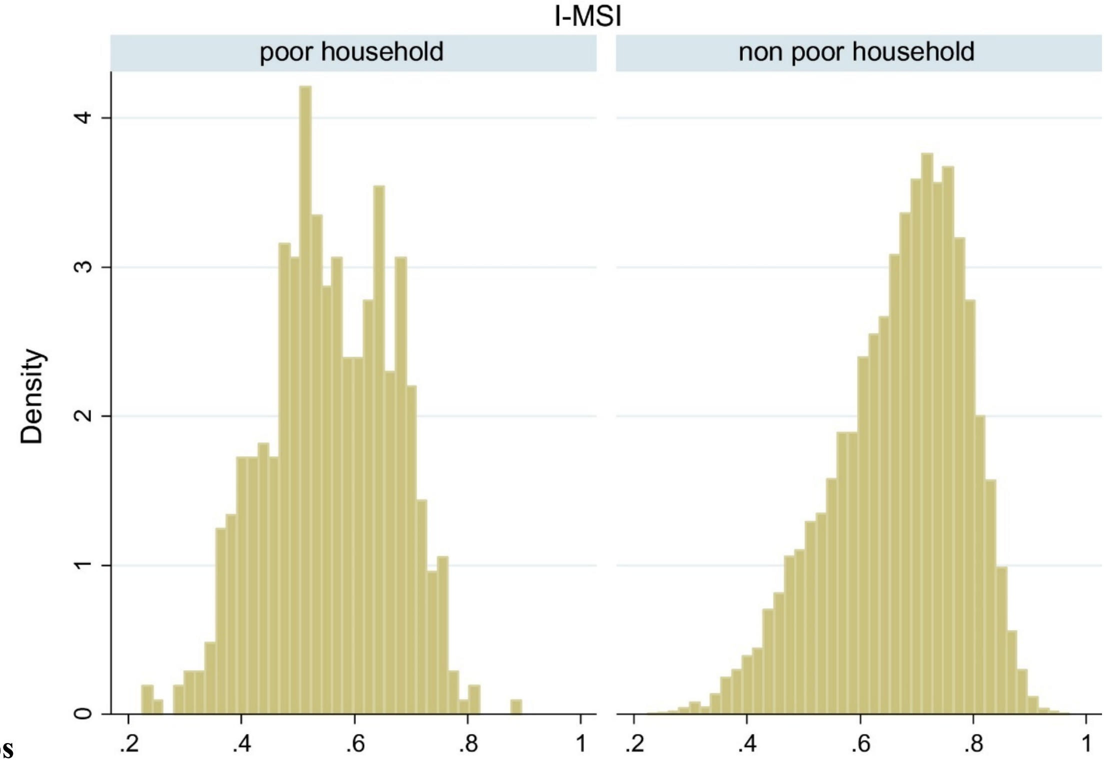
The ***MSI***, ***I-MSI*** and both of the means are positively and significantly correlated with income per capita. As expected, the ***I-MSI*** has the highest correlation with income (see Table 4 in Section 4.3). To capture the links between a unidimensional measure of income poverty and multidimensional indexes, a dummy poor vs non-poor variable was computed for each individual using Chinese urban and rural poverty lines (Bandyopadhyay, 2017).[12] The results correlate with all four measures of multidimensional well-being.[13] As expected, the ***I-MSI*** is most strongly correlated with income poverty. For individuals where ***MSI > I-MSI***, the average income is 6,652 ¥; for the remaining individuals, the average income is 31,739 ¥. If the analysis is restricted to the bottom 4 percent of the sample (the 560 income poorest individuals), ***I-MSI < MSI*** for 95.4 percent of individuals (***I-MSI > MSI*** for only 26 of the poorest individuals, specifically 4.6 percent).

This analysis shows that the ***I-MSI*** is, as expected, strongly related to monetary poverty, although the income dimension is not directly included in the computation of well-being. Nonetheless, income poverty and the ***I-MSI*** do not overlap — theoretically and empirically. Figure 4 confirms this result distinguishing ***I-MSI*** levels of well-being amongst individuals from income poor and income non-poor households. The distribution of the ***I-MSI*** scores amongst non-poor households is skewed toward higher well-being, but income poor and non-poor households are not internally homogeneous once multidimensional individual well-being is considered. For example, there are individuals from income-poor households with high levels of multidimensional well-being. In fact, the well-being dimensions of some individual can be high even in the presence of low income and vice versa. This is also due to the intra-household distribution of resources, which is not captured by income per capita (see next subsection), or to the non-monetary support provided through the informal social protection.

[12]The poverty line of urban areas is defined at provincial level for the purpose of assigning the "DiBao" income subsidy. We, therefore, considered the rural poverty line and the 12 urban poverty lines; one for every province in our dataset, based on the official DiBao thresholds reported for September 2015. The data for the urban poverty lines was accessed on 22 June 2020 via the following link: http://files2.mca.gov.cn/www/201512/20151202084401543.htm.

[13]The correlation coefficient between the multidimensional indexes and the dummy for income poverty is always negative (the higher the well-being, the lower the risk of falling into poverty) and statistically significant at the 1 percent level. The correlation between the multidimensional indexes and income is always positive and statistically significant at the 1 percent level. These correlations are stronger for the ***I-MSI***. Moreover, income-poor individuals report average ***I-MSI*** scores significantly lower than those reported for the *MSI* (0.56 and 0.61 respectively).

Figure 4 **- Distribution of *I-MSI* Within Income-Poor hh (Left) and Non-Poor (Right)**



subgroups

4.2 The I-MSI and Inequalities Between Different Groups

To check whether different groups of individuals have significantly different levels of multidimensional well-being, the sample is divided with respect to geographic characteristics (urban residents with and without an urban hukou and rural dwellers in different regions) and social characteristics (age, gender). The $\overline{\boldsymbol{I-MSI}}$ is the average ***I-MSI*** observed in individuals belonging to different geographical and social sub-groups, as reported in Table 4. A two-group t-test with equal variances is used to check whether the individuals in each group have significantly different well-being scores from the whole sample. The results for the arithmetic mean, the geometric mean and the ***MSI*** are reported for comparisons (Table 3).

Table 3

I-MSI Average Scores and Other Aggregation Methods Grouped by Categories

	Panel A: Geographic Characteristic						
	Urban hukou	Urban no hukou	Rural	East	Centre	West	Total
A. Mean	0.73***	0.69***	0.69***	0.72***	0.68***	0.68***	0.70
G. Mean	0.65***	0.55***	0.57***	0.63***	0.55***	0.54***	0.59
MSI	0.69***	0.64***	0.64***	0.68***	0.64***	0.64***	0.66
I-MSI	0.70***	0.64***	0.65***	0.68***	0.64***	0.64***	0.66

	Panel B: Social Characteristic					
	Age ≤ 30	31-59	Age ≥ 60	Male	Female	Total
A. Mean	0.75***	0.72***	0.64***	0.71***	0.69***	0.70
G. Mean	0.68***	0.64***	0.48***	0.63***	0.56***	0.59
MSI	0.71***	0.69***	0.59***	0.68***	0.64***	0.66
I-MSI	0.71***	0.69***	0.60***	0.68***	0.65***	0.66

Note: *, ** and *** indicate respectively a difference at 10%, 5% and 1% significance levels between overall and within groups averages.

The findings across geographical regions reported in Table 4 Panel A show lower $\overline{\boldsymbol{I-MSI}}$ levels of well-being in rural and inner areas, which are consistent with our expectations as well as previous investigations (Li *et al*., 2013; Alkire and Shen, 2017). These differences are statistically significant at the 1 percent level and robust with respect to the choice of the well-being index.

Table 4 Panel B reports differences across social groups. Such differences are less commonly investigated largely because most traditional measures of development, including income per capita and the Global MPI (Alkire and Shen, 2017), measure well-being at household level thus reducing the prospects for intra-household analysis. However, lower levels of well-being among women and elderly people are expected based on existing literature (Li and Liang, 2007; Chi and Li, 2008; Li *et al*., 2013; Alkire and Shen, 2017). Our results confirm these expectations: women and the elderly report a significantly lower aggregate score irrespectively of the index applied.

Investigating the overlap between social and economic groups in greater depth reveals that some of the inequalities considered above reinforce each other. For example, the gender gap is stronger in rural communities, western provinces and among older age groups in contrast to urban areas, eastern provinces and younger age groups. This finding is consistent with “sticky floor” discrimination theory (roughly the opposite of the “glass ceiling” hypothesis). Chi and Li (2008) found evidence of “sticky floor” discrimination among Chinese women in the economic sphere, indicating that the highest gender discrimination in earnings is found amongst poorer Chinese women, rather than amongst their richer counterparts. Our analysis extends this finding to the multidimensional sphere and indicates that in more backward contexts (associated with lower multidimensional well-being), the gender differential in multidimensional deprivation is higher.

4.3 The I-MSI And Subjective Well-Being

In the previous sections, we showed that the four aggregation criteria are positively correlated with each other; and that all of them are positively correlated with income and negatively correlated with income poverty.

To further advance in our understanding of the ***I-MSI***, we consider the correlation with self-reported perceptions of the life quality (Table 4). The ***I-MSI*** indexes and the other aggregate measures are all significantly correlated with self-reported quality of life. All aggregated indexes are more strongly correlated with the quality of life than with unidimensional indicators of well-being (for a comparison, see Table 4 and Table 1 in Section 3.1), including income. This implies that while every indicator describes aspects of life that matter in subjective evaluations, an index that jointly considers these indicators has greater informative power.

Table 4
Correlation between multidimensional indexes and other well-being proxies

	HH Income		Individual Life Quality	
A. Mean	0.294	***	0.235	***
G. Mean	0.278	***	0.179	***
MSI	0.299	***	0.233	***
I-MSI	0.502	***	0.252	***
Income	1	***	0.157	***

Note: *, ** and *** represent respectively a significance level of correlation of 10%, 5% and 1%

The index reporting the highest correlation with the quality of life is the ***I-MSI*** (Pearson correlation coefficient = 0.252) which is just slightly above the ***MSI*** correlation of 0.233. An F-distribution test confirmed that the correlation between the quality of life and the ***I-MSI*** is significantly higher than other correlations with the quality of life. The geometric mean is significantly less correlated. The correlations between the quality of life and the ***MSI*** or the arithmetic mean are not significantly different. Household income per capita is also well-approximated by the multidimensional indexes (only the 'work' dimension generates a correlation coefficient similar to that associated with income).

5 Conclusions

This paper introduces a new multidimensional index to aggregate different dimensions of well-being: The Income-adjusted Multidimensional Synthesis of Indicators (***I-MSI***). This index represents an attempt to reconcile theoretical and empirical debates on the role of income in the measurement of well-being. Indeed, income remains the core dimension for most economic analyses. It plays a relevant role also in the basic-needs approach as well as in the Capability Approach, however, its main characteristic, i.e. being a means and not an end of human activities, has not yet been translated into a shared statistical methodology.

In the ***I-MSI***, the direct sources of well-being are intrinsically valuable dimensions that exclude income. By rejecting the assumption of perfect substitutability across dimensions, the ***I-MSI*** incorporates the penalization of heterogeneous outcomes. Income becomes a key variable that determines how penalization is computed. In other words, we assume that income richer individuals have instruments to better defend themselves against shocks in any specific dimension and a greater capacity to continue enjoying well-being in remaining dimensions. Technically, this implies computing a specific penalization of heterogeneity for each individual. Moreover, the formulation of the ***I-MSI*** allows for a "smooth" penalization of heterogeneity and avoids a collapse towards a zero well-being index that affects other non-fully compensatory aggregation methods, such as the geometric mean.

The descriptive power of the ***I-MSI*** and its potential advantages can be illustrated through micro-level analysis. As a case study, this paper considers well-being amongst Chinese individuals by focusing on eight dimensions of well-being. The ***I-MSI*** produces a continuous synthetic measure of aggregated well-being for each individual in the sample. This allows us to observe diverse average scores across different areas and population groups. For instance, rural dwellers, communities in western provinces, domestic migrants, women and older people all report lower levels of multidimensional well-being. Moreover, some disadvantages overlap, pointing to intersecting inequalities which distinguish the poverty profiles of individuals at greater risk. These results are statistically significant and robust to the choice of aggregation method.

The results show that the ***I-MSI*** has the highest correlation with self-reported quality of life, indicating that the capacity of our index to aggregate across dimensions is consistent with the overall impression of how individuals included in the survey evaluate their subjective well-being. It is worth reiterating that the self-reported perceptions of well-being among individuals is more significantly correlated with the ***I-MSI*** than any individual components of the index or the results of any other aggregation method.

In sum, the ***MSI*** technique—which is consistent with the theoretical findings of the Capability Approach—is adjusted by the ***I-MSI***, resulting in an empirical framework that accounts for the instrumental role of income.

References

Alkire, S., and Foster, J., "Counting and multidimensional poverty measurement," *Journal of public economics*, *95*(7), 476-487, 2011.

Alkire, S., and Shen, Y., "Exploring multidimensional poverty in China," *Policy Briefings*, 2015.

Anand, S., and Sen, A., "Concepts or human development and poverty! A multidimensional perspective," *United Nations Development Programme, Poverty and human development: Human development papers*, 1-20, 1997.

Atkinson, T., Cantillion, B., Marlier, E., Nolan, B., *Social Indicators: The EU and Social Inclusion*, Oxford University Press, Oxford, 2002.

Bandyopadhyay, S. (Ed.), *Research on Economic Inequality: Poverty, Inequality and Welfare*, Emerald Publishing Limited, 2017.

Biggeri, M., and Mauro, V., "Towards a more 'Sustainable' Human Development Index: Integrating the environment and freedom." *Ecological Indicators*, *91*, 220-231, 2018.

———, Clark, D. A., Ferranini, A., & Mauro, V., "Tracking the SDGs in an 'integrated'manner: A proposal for a new index to capture synergies and trade-offs between and within goals," *World Development*, *122*, 628-647, 2019.

Bin, P., "Dynamic Development of Regional Disparity in Mainland China: An Experimental Study Based on a Multidimensional Index," *Sustainability*, *8*(12), 1287, 2016.

Bossert, W., Chakravarty, S. R., and D'Ambrosio, C., "Multidimensional poverty and material deprivation with discrete data," *Review of Income and Wealth*, *59*(1), 29-43, 2013.

Bourguignon, F., and Chakravarty, S. R., "The measurement of multidimensional poverty," *The Journal of Economic Inequality*, *1*(1), 25-49, 2003.

Burchi, F., Rippin, N., and Montenegro Muñoz, C., "From income poverty to multidimensional poverty: An international comparison,", 2018.

Chakravarty, S. R., and D'Ambrosio, C., "The measurement of social exclusion," *Review of Income and wealth*, *52*(3), 377-398, 2006.

Chen, Z., Yang, G., Offer, A., Zhou, M., Smith, M., Peto, R., Ge, H., Yang, L., and Whitlock, G., "Body mass index and mortality in China: a 15-year prospective study of 220 000 men," *International journal of epidemiology*, *41*(2), 472-481, 2012.

Chi, W., and Li, B., "Glass ceiling or sticky floor? Examining the gender earnings differential across the earnings distribution in urban China, 1987–2004," *Journal of Comparative Economics*, *36*(2), 243-263, 2008.

Clark, D. A., *Visions of development: A study of human values*, Edward Elgar Publishing, 2002.

Combes, P. P., Démurger, S., Li, S., and Wang, J., "Unequal migration and urbanisation gains in China," *Journal of Development Economics*, 2019.

Decancq, K., and Lugo, M. A., "Weights in multidimensional indices of wellbeing: An overview," *Econometric Reviews*, *32*(1), 7-34, 2013.

Deneulin, 2009 MANCA

De Neubourg, C., Chai, J., de Milliano, M. and Plavgo, I., "Step-by-Step Guidelines to the Multiple Overlapping Deprivation Analysis (MODA)", Innocenti Working Papers 695, 2013.

Di Tommaso, M. R., Tassinari, M., Bonnini, S., & Marozzi, M., "Industrial policy and manufacturing targeting in the US: new methodological tools for strategic policy-making," *International Review of Applied Economics*, *31*(5), 681-703, 2017.

Feng, D., Chu, X., and Chen, W., “Study of Comprehensive Evaluation of Poverty Reduction Effect for Chinese Poverty-Stricken Areas-Based on the Data of Chinese 14 Contiguous Poor Areas,” *Applied Economics and Finance*, *3*(3), 37-44, 2016.

Frediani, A.A., Clark, D. A. and Biggeri, M., “Introductory chapter”, in D. A. Clark, M. Biggeri and A.A. Frediani (Eds), *Sen’s Capability Approach, Empowerment and Participation: Concepts, Methods and Applications*, Palgrave MacMillan, 2019

Goodman, D. S., and Segal, G., *China deconstructs: Politics, trade and regionalism*, Routledge, 2002.

Gordon, D., Nandy, S., Pantazis, C., Pemberton, S., and Townsend, P., *Child poverty in the developing world*. Policy Press, 2003.

Joshi, D., “Does China's recent ‘harmonious society’discourse reflect a shift towards human development?,” *Journal of Political Ideologies*, *17*(2), 169-187, 2012.

Klugman, J., Rodríguez, F., and Choi, H. J., “The HDI 2010: new controversies, old critiques,” *The Journal of Economic Inequality*, *9*(2), 249-288, 2011.

Knight, J., and Gunatilaka, R., “Great expectations? The subjective well-being of rural–urban migrants in China,” *World development*, *38*(1), 113-124, 2010.

Kovacevic, M., “Review of HDI critiques and potential improvements,” *Human development research paper*, *33*, 1-44., 2010

Labar, K., and Bresson, F., “A multidimensional analysis of poverty in China from 1991 to 2006,” *China Economic Review*, *22*(4), 646-668, 2011.

Laderchi, C. R., Saith, R., & Stewart, F., “Does it matter that we do not agree on the definition of poverty? A comparison of four approaches,” *Oxford development studies*, *31*(3), 243-274, 2003.

Li, L. W., and Liang, J., “Social exchanges and subjective well-being among older Chinese: Does age make a difference?,” *Psychology and Aging*, *22*(2), 386, 2007.

Li, S., Sato, H., and Sicular, T. (Eds.)., *Rising inequality in China: Challenges to a harmonious society*, Cambridge University Press, 2013.

Li, L., and Wu, X., “Housing price and entrepreneurship in China,” *Journal of Comparative Economics*, *42*(2), 436-449, 2014.

Maasoumi, E., “The measurement and decomposition of multi-dimensional inequality,” *Econometrica: Journal of the Econometric Society*, 991-997, 1986.

———, & Nickelsburg, G., “Multivariate measures of well-being and an analysis of inequality in the Michigan data,” *Journal of Business & Economic Statistics*, *6*(3), 327-334, 1988.

———, & Racine, J. S., “A solution to aggregation and an application to multidimensional ‘well-being’ frontiers,” *Journal of econometrics*, *191*(2), 374-383, 2016.

Mauro, V., Biggeri, M., and Maggino, F., “Measuring and monitoring poverty and well-being: A new approach for the synthesis of multidimensionality,” *Social Indicators Research*, *135*(1), 75-89, 2018.

Mazziotta, M., and Pareto, A., “Methods for constructing composite indices: One for all or all for one,” *Rivista Italiana di Economia Demografia e Statistica*, *67*(2), 67-80, 2013.

Nicholas, A., Ray, R., and Sinha, K., “Differentiating Between Dimensionality and Duration in Multidimensional Measures of Poverty: Methodology with an Application to China,” *Review of Income and Wealth*. 60(1), 48-74, 2017.

Nussbaum, M., “Capabilities as Fundamental Entitlements: Sen and Social Justice,” *Feminist Economics* 9 (2–3): 33–59, 2003.

Pakrashi, D., and Frijters, P., “Migration and discrimination in urban China: a decomposition approach,” *Review of Income and Wealth*, *63*(4), 821-840, 2017.

Pogge, T.,” A critique of the capability approach. *Measuring justice: Primary goods and capabilities,” 17*, 2010.

Qi, D., and Wu, Y., “A multidimensional child poverty index in China,” *Children and Youth Services Review*, *57*, 159-170, 2015.

Ravallion, M., and Chen, S., “China's (uneven) progress against poverty,” Journal of development economics, 82(1), 1-42, 2007.

Ray, R., and Mishra, A., “Multi-dimensional deprivation in the awakening giants: A comparison of China and India on micro data,” *Journal of Asian Economics*, *23*(4), 454-465, 2012.

Sen, A., *Development as Freedom*, New York: Alfred A. Knopf, 1999.

———, “Children and human rights,” *Indian Journal of Human Development*, *1*(2), 235-245, 2007.

Shue, V., and Wong, C. (Eds.)., *Paying for progress in China: Public finance, human welfare and changing patterns of inequality*, Routledge, 2007.

Sicular, T., Ximing, Y., Gustafsson, B., and Shi, L., “The urban–rural income gap and inequality in China,” *Review of Income and Wealth*, *53*(1), 93-126, 2007.

Stiglitz, J., Sen, A. K., and Fitoussi, J. P., “The measurement of economic performance and social progress revisited: reflections and overview,”, 2009.

Trani, J. F., Biggeri, M., and Mauro, V., “The multidimensionality of child poverty: Evidence from Afghanistan”. *Social indicators research*, 112(2), 391-416, 2013.

Ward, P., “Measuring the Level and Inequality of Wealth: An Application to China,” *Review of Income and Wealth*, *60*(4), 613-635, 2014.

You, J., Wang, S., and Roope, L., “Intertemporal deprivation in rural china: income and nutrition,” *The Journal of Economic Inequality*, *16*(1), 61-101, 2018.

Appendix A – Comparing different aggregation methods

In this appendix, we investigate the behaviour of the $I - MSI$ in relation to other common indexes used to aggregate multidimensional indicators with respect to horizontal variability in the outcomes of a single unit. The results of our simulations reinforce the choice of the $I - MSI$ and MSI methods given the theory behind heterogeneity penalisation.

We performed three different simulations on 6 variables that are measured across 20 units. For each simulation, we calculated the $I - MSI$ (and the MSI), the geometric mean, and three other widely used aggregation functions, e.g., the Logistic, Liptak, the Mazziotta Pareto index (M-P) and Fisher functions (see Di Tommaso et al., 2017). The 6 variables used in the analysis are bounded between 0 and 1, and each simulation was set at a different overall average level. In each of the three simulations, all units share the same arithmetic mean but have different levels of variability between outcomes. As a result, the differences in the aggregate score should only reflect the difference between heterogeneity levels. The simulations, reported in Figure 5, were set at three progressive levels of the arithmetic mean (and income), namely 0.2 (low), 0.4 (intermediate) and 0.6 (high). The results show that the Liptak, the Logistic functions, and the geometric mean are very sensitive to changes in cases of high heterogeneity. The Logistic function seems to be slightly more robust, especially for low overall levels, while the $I - MSI$ is very robust for every level of heterogeneity and every average level. For any level of overall outcomes, the general behavior of the

functions is approximately the same. The geometric mean, the M-P and the $I - MSI$ decrease as heterogeneity increases, while the other three functions tend to report higher values for higher heterogeneity. Figure 5 presents the results for low levels of overall outcomes (in the first panel row mean = 0.2 and income = 0.2). Both the geometric mean, the M-P and the $I - MSI$ present index values that are below 0.2, but the geometric mean is very sensitive if units have a high degree of heterogeneity. The other indexes tend to increase as a function of the variability, which is hard to justify theoretically. Adopting means of 0.4 and 0.6 respectively confirms our observations: higher levels of heterogeneity increase the volatility of all indexes, excluding the $I - MSI$. Higher levels of overall outcomes are more easily associated with higher levels of variability. Please note that the horizontal axis in the three figures have different scales.
We have seen that the $I - MSI$ shares some important properties with the MSI and it is therefore bounded between the arithmetic mean (in case of maximum income) and the minimum value of the row of outcomes (in case of minimum income).
The result of this simulation shows that the $I - MSI$ and MSI are better suited for tackling heterogeneity than the other indexes considered above. This feature becomes more significant as the number of dimensions increase, which raises the probability that at least one constituent component, or indicator, will approach zero. In other measures like the geometric mean, this leads to instability or a complete collapse to zero.

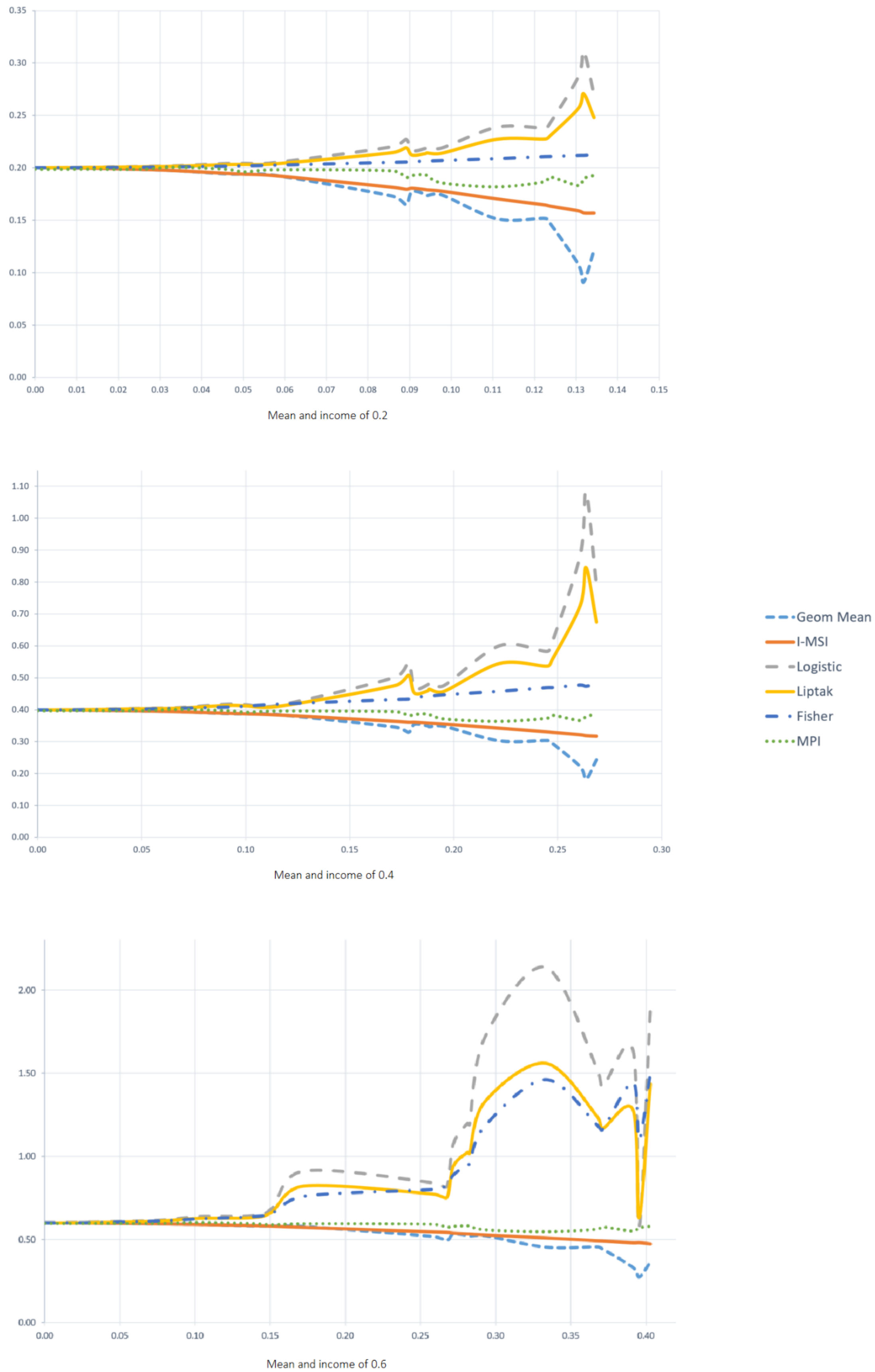


Figure 5. Relative comparison for the 6 transformations for a sample of 20 observations with a mean and income of 0.2, 0.4 and 0.6.

Appendix B – Computation of Dimension Scores

Each of the eight dimensions, as explained in detail below, is computed using designated variables (see Table 5). In total, fifty-five primary data (categorical, continuous and binary) are used grouping some of these data into nineteen variables. Following the HDI, where the reference point for normalization is established before considering the range of the observations, each dimension is constructed to range between 0 and 1, where these boundaries are normatively set. Moreover, different thresholds apply to each dimension, as described hereafter. An advantage of measuring dimensions without the traditional max-min normalization (data driven and possibly biased by outliers), is that this allows us to obtain measures which are comparable across individuals or countries and over time (as far as the variables and computation procedures remain meaningful in different contexts).

Table 5

Selected dimensions and variables grouped accordingly plus hh income

Dimension	Variable	Unit	Code	Mean	Min	Max	Miss.	Total	Miss.%
Education	Education	Years	a11	8.52	0	13	16	13839	0.12
Health	Severity	Categorical	m25	0.32	0	3	981	13839	7.09
	Chronic (20)	Y/N	u12-u24w	0.30	0	8	16	13839	0.12
	Smoking	Y/N	u27	0.23	0	1	998	13839	7.21
Nutrition	Height	Cm	height	161.72	126	193	491	13839	3.55
	Weight	Kg	weight	63.34	25.6	146	510	13839	3.69
Housing	Rooms	Nr/ppl	l17	1.45	0.13	10	280	13839	2.02
	Ownership	Y/N	l200	0.92	0	1	16	13839	0.12
Sanitation	Water	Categorical	l1	0.42	0	3	109	13839	0.79
	Flush	Categorical	l5	0.76	0	3	19	13839	0.14
	Surrounding	Categorical	l6	0.21	0	3	19	13839	0.14
Asset	Durables (8)	Y/N	l31– l140	5.64	0	8	0	13839	0.00
Work	Employment	Y/N	b2	0.48	0	1	3	13839	0.02
	Cause Unem.	Categorical	b2a	3.81	1	9	25	7244	0.00
Leisure	Sleep	Hours/day	u324	7.79	0	24	1101	13839	7.96
	Physical (6)	Minutes/day	u145-u155	49.56	0	1260	972	13839	7.02
	Sedentary (9)	Minutes/day	u339-u508	907.84	0	1260	971	13839	7.02
Income	Income hh	ln (¥)	hhinc_pc	9.31	0	13.94	269	13839	1.94

Note: ‘Chronic’ [diseases] contains information from 20 different dummy variables; ‘Durables’ from 8 dummy variables; ‘Physical’ and ‘Sedentary’ summarize respectively 6 and 9 different variables/activities. ‘Cause Unem.’ is a categorical variable referring to the 7,244 non-working individuals.

The constituent components of the $I - MSI$ are constructed as follows.

Education is based on the linear normalization of a single variable: individual years of education (ranging from 0 and 13).

Health incorporates the number and severity of diseases contracted in the month preceding the interviews (days of inability to work multiplied by the self-reported severity of the illness on a scale from 1 to 3). The number obtained was further downscaled in the case of the presence of permanent diseases and smoking habit.

Nutrition is derived from the Body Mass Index (BMI). The minimum value of this indicator corresponds to a BMI ≤ 13 kg/m² or a BMI ≥ 40 kg/m². The highest well-being corresponds to a BMI equal to 22.5 kg/m² (Chen et al., 2012). Intermediate values are determined linearly.

Housing, which is particularly relevant in many development contexts (Clark, 2002) including China[14] (Li and Wu, 2014), is computed linearly on the number of rooms per capita (up to 3) and house ownership.
Sanitation depends on household facilities, specifically the availability of tap water, flush toilets and the cleanliness of surroundings. These three aspects are caught by categorical variables that were summed-up and normalized in [0; 1].
Assets are estimated at household level, and incorporate the ownership of a color TV, a washing machine, a refrigerator, cooking devices, a bicycle or motorbike, a telephone or mobile phone, and a car or tractor. The dummies for the ownership of all these durables were summed (attaching a double value to cars). A fully deprived household does not own any of these consumer durables.
Work is represented by the employment status of each respondent as well as his or her relatives; distinguishing between employed, students, retirees and unemployed. '0' correspond to no employment nor pension for all household members; 1 to full employment of all household members.
Leisure is based on the sum with proportional weights of the time devoted to sleep (up to 8 hours per day), to "recreational sedentary time" (up to 1 hour per day) and to "recreational non-sedentary time" (up to half an hour per day) activities (Clark, 2002).
Household income per capita is included through the function f_i, described in equation (3). The natural logarithm of income y_i was observed in the interval [5-11]. The values below or above this range were considered as outliers – around 3% below 5 and 5% above 11. Moreover, a was set at 1,427.5 yuan (equivalent to half of the rural poverty line), so that ln(a) > 5 prevents $f_i \leq 0$.

[14] Housing, education and healthcare are commonly known as the new "three big mountains", showing the renewal of China's main concern (the "mountains" of Maoist China were imperialism, feudalism and bureaucrat-capitalism).

Appendix C – Well-being dimensions standardized and income in logarithmic form

Table 6
Well-being dimensions standardized and income in logarithmic form

	Obs.	Mean	Std. Dev.	Min	Max
Education	13839	0.65	0.29	0	1
Health	13839	0.85	0.24	0	1
Nutrition	13836	0.78	0.17	0	1
Housing	13839	0.55	0.20	0.06	1
Sanitation	13839	0.84	0.21	0.11	1
Assets	13839	0.70	0.18	0	1
Work	13839	0.63	0.35	0	1
Leisure	13839	0.58	0.15	0	1
Income hh	13839	9.26	1.97	0	13.94