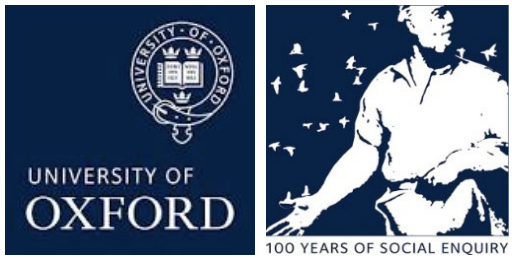


Mixed Method Approach to Multidimensional Poverty Measurement using Bogor City, West Java, Indonesia as a Case Study

Putu Geniki Lavinia Natih

Trinity College

University of Oxford

Trinity Term 2020

Thesis submitted for the degree of Doctor of Philosophy

Approximate word count: 98,800

Abstract

Mixed Method Approach to Multidimensional Poverty Measurement, using Bogor City, West Java, Indonesia as a Case Study

Putu Geniki Lavinia Natih, Trinity College

Submitted for the degree of Doctor of Philosophy, Trinity Term 2020

This thesis moves beyond the limitations of the consumption measure, to exploring the merits of applying a multidimensional framework to the analysis of poverty within the developing world, using Indonesia as a case study. It confirms that how one defines, and measures poverty may significantly affect who and how many are identified as poor and argues for the crucial need for transparency and public discussion on the reasons behind the choices made within poverty measurement.

In support of the above, three existing measures were first examined; i.e. the Consumption measure, the Global MPI and an existing nationally developed multidimensional measure, which I call the Adjusted MPI. Analysis was conducted using one wave of Indonesia's nationally representative National Socioeconomic Survey (Susenas 2013). Weak overlap was found between these measures, indicating that firstly, using consumption as the sole indicator of poverty may not be enough and secondly, localised measures may be needed when measuring multidimensional poverty within a country as diverse as Indonesia.

The above results in mind, this thesis proposed the use of the Delphi method to address the arbitrariness in the way existing multidimensional measures, including those examined above, selected their components; i.e. the dimensions, indicators and weights to include within measurement. The Delphi exercise was conducted with policy makers within one district in Indonesia, i.e. Bogor City, West Java. This exercise found consensus on the importance of five dimensions; i.e. (1) Education, (2) Employment and Assets, (3) Living Standards, (4) Health and Safety and (5) Child Health and Contraception, alongside eighteen indicators to proxy deprivation within these dimensions. Agreement was also achieved on a range of weights indicating the relative importance and possible trade-offs between each of these components. In order to determine the 'exact value' of these weights, opinion data resulting from the Delphi was analysed using an optimisation exercise based on Data Envelopment Analysis and the benefit-of-the-doubt approach to weighting. The weights resulting from this optimisation exercise, dubbed as 'expert' weights, were then utilised to aggregate the dimensions and indicators chosen by Delphi participants to form a 'new' measure of multidimensional poverty using Susenas 2013 data. Two comparative exercises were subsequently conducted. The first exercise evaluated the effects of using these 'expert' weights versus the commonly applied, 'equal nested' weights, within the 'new' measure. Both weighting schemes assigned similar ranks to households and identified relatively similar types of households as poor. The 'expert' scheme, however, identified a significantly higher extent of poverty compared to the 'equal nested' scheme.

The second exercise compared the 'new' measure using 'expert' weights, with the existing measures examined previously within this thesis. Low levels of overlap were also found. The weak overlap found between the Consumption and the 'new' multidimensional measure confirmed the importance of applying both uni and multidimensional approaches when measuring poverty; whilst differences between the globally and nationally constructed multidimensional measures vs. the 'new' district level Delphi-based measure supported the argument that internationally-generated measures should be supplemented with localised measures to support the formation of sound policy which aims to address the myriad welfare priorities that are relevant for specific areas.

Acknowledgements

Concluding my DPhil at Oxford in the midst of the COVID-19 pandemic feels surreal and highlights all that I have been through to get to where I am today. Most importantly, it brings to mind all the continuous support, guidance and kindness I have received throughout this journey. Thinking of this brings hope and warmth in a world where the future is so uncertain, that it feels as if this is all there is to hold on to.

This thesis was completed with the guidance and continuous support of many I am so grateful for. Firstly, my sincerest gratitude goes to my supervisor, Professor Erzsébet Bukodi, for her continuous guidance, constructive comments and willingness to explore issues within the new and sometimes 'quirky' setting of conducting research in Indonesia. Bess was more than a supervisor to me. She helped me through so many things; particularly my grandmother passing away and although it was something very hard for me to talk about at the time, listened to me share about my miscarriage and somewhat challenging pregnancy and birthing process.

I owe a lot also, to Bogor City, its City Government and people. Without their support, this thesis would never have been able to come to fruition. I am wholeheartedly thankful to Dr Bima Arya Sugiarto, the Mayor of Bogor City, who allowed me precious time, to listen to my ideas in the midst of his very busy schedule. Dr Sugiarto also gave me the invaluable experience of going into the field with his team, talking to the people and step-by-step solving the challenges faced by those in need. I ultimately dedicate this thesis' findings to supporting these efforts.

I am entirely grateful to my esteemed examiners, Professor Brian Nolan and Dr David Clark, who gave me the most inspiring and knowledge-rich viva experience. The input they gave me was invaluable. Most importantly, to be appreciated by both these highly respected scholars, gave me so much happiness and further cemented my love for research within the world of welfare and poverty measurement.

I thank Professor Ari Kuncoro and all at the University of Indonesia's Economics Faculty, who gave me so much support. Professor Kuncoro is like a second father to me; his guidance, has been pivotal throughout my academic life.

I am also thankful for the chance given to me by the Indonesia Project to visit the Australian National University on a fully funded visiting scholar's grant. Professor Budy Resosudarmo highlighted that this honour is usually given to those at later stages in their career; I am thus, so grateful that I was given the trust and opportunity to take up this role as a DPhil student. Being able to present my research at ANU gave me crucial contextual input on this thesis. No less valuable were the friendships I was able to build whilst at ANU; I continue to cherish the time I spent with Mbak Ingrid, the gang at Toad Hall and the Coombs building.

My most sincerest gratitude goes to the Jardine-Foundation for making my dream to study at Oxford University possible. It has been a true honour to be a Jardine Scholar. The support that the Jardine Foundation's board of directors has given me surpasses that of a regular scholarship body. All at the Foundation have truly become my family away from home. I cannot thank them enough.

I am truly grateful for the support given to me by Sir Ivor Roberts and Lady Elizabeth Roberts and all at Trinity College. I cannot imagine my Oxford journey without them.

I thank Dr Stuart Basten for suggesting the Delphi method.

I thank friends and researchers at the Oxford Poverty and Human Development Initiative (OPHI) for allowing me to present my research at their weekly seminars and providing me with constructive and insightful comments. Gratitude also goes to the Social Policy Department's Poverty and Social Inequality Research Groups, which were great outlets for me to discuss the issues I faced during my DPhil and very much helped me discover solutions to these challenges.

I am also so grateful to friends within the Indonesian Embassy; His Excellency Ambassador Thayeb, Ibu Lastri, His Excellency Ambassador Sukma and Ibu Hanna. A lot of thanks also goes to Pak Nurul, Mbak Theo, Mas Adit and Mbak Angke from the Indonesian Investment Coordinating Board in London. I thank Professor Subroto and Ibu Ratih for their continuous support throughout my academic journey. I am grateful for my MPhil and DPhil family, Firman, Matthews, Mary Jean, Joe, Yuza, Fara, Nadine, Yesha, Min Sern, Madeline, Tahnee, Camilla, Andrea, Ankira, Floria and Charlotte.

Last but not least, my thanks and love go to my parents, Dr Gede and Dr Sue Natih, my grandparents, my Aunty Terry and Uncle Christopher. My husband, Gde, who has patiently supported me throughout everything. Last but not least, my little Ashvin. We chose his name 'healer of the world', as he has indeed healed me. It seems however, even more fitting, as the world today has never before been in more need of a healing presence.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

'The most important development of poverty research in recent years is certainly the shift of emphasis from a uni- to a multidimensional approach to poverty. Conceptualizing the multiple facets of poverty is, however, not an easy task'

(Kakwani & Silber, 2007, p. 3)

'By 2030, nations should reduce at least by half the proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions'

(United Nations, 2016)

'When poverty is expanded to include other dimensions of human well-being; i.e. not only consumption, but also education and health, for example, the true extent of poverty in Indonesia becomes extremely striking; with two-thirds of the rural population and almost thirty percent of urban households identified as deprived in at least one of these multiple dimensions'

(Milne, 2006, p. 31)

1.1 Overview

Growing consensus is evident within academia and policy on the fact that poverty encompasses multiple aspects of well-being; therefore, to assess it, one needs to look beyond uni-dimensional monetary-based measures. Thorbecke (2007) argues that 'in spite of spectacular methodological advances in the analysis of poverty, a number of conceptual and measurement issues remain to be addressed or clarified' (p. 3). Most of these 'remaining issues', are particularly related to this multidimensionality of poverty. Given that 'just knowing' the amount of money available for a given number of people 'will not take us very far', questions on what information, which criteria and how best to assess 'human thriving' remain to be solved (M. Nussbaum & Sen, 1993).

Kakwani and Silber (2007) argue that 'the most important development within poverty research is this shift from a uni- to multidimensional approach to understanding poverty'; yet, substantial concern has been raised with regard to the way current empirical applications operationalise this approach. These critiques have centred on two main issues, which mirror Amartya Sen (1976a)'s two 'problems' faced when measuring poverty; i.e. (1) the 'identification' of the aspects and criterion of poverty and (2) the 'aggregation' of these multiple aspects to form an index of poverty.

The above in mind, critiques first point to the 'arbitrariness' in the way 'components' are selected within existing multidimensional measures, as often, 'neither the menu of indicators nor the aggregation method utilised are pre-determined from theory and practice' (Martin Ravallion, 2011a, p. 2). Furthermore, on this issue Chiappero-Martinetti and von Jacobi (2012) note 'the lack of discussion' within literature with regard to how decisions on these

components may influence the extent of multidimensional poverty reported, which in turn could carry significant implications for policy (p. 70). Secondly, debates have revolved around the more fundamental question of whether one truly needs a multidimensional index of poverty; as Martin Ravallion (2011b) argues, 'the fact that poverty is multidimensional, does not imply the need for a composite measure which aims to encompass these multifaceted items'. This is primarily due to the fact that, despite the clear policy purpose of these composites, the manner by which multiple attributes are aggregated, may challenge the soundness of such measures rather than strengthen their application. On this issue, Martin Ravallion (2011b) points to the way weights for the multiple types of deprivations are chosen; 'in practice, these weights are set by the analyst, with no obvious reason to suppose that they would be accepted by those one is trying to help through measuring poverty' (p. 11). To forgo this use of weights when empirical evidence on their values is scarce, critiques have thus, advocated for monitoring dimensions separately. By undertaking this, analysts and policy makers, however, lose the ability of examining 'joint deprivations' and 'complete orderings of countries' (F. H. G. Ferreira & M. A. Lugo, 2013). Regardless of which technique is used, however, as touched upon above, 'the ongoing challenge faced by statisticians and researchers is still how to capture important elements of poverty in transparent, reliable, and practical ways' (Morduch, 2006, p. 29).

This thesis proposes a mixed-methods approach to addressing the above debates, through the use of a consensus building technique, known as the Delphi, thereby aiming to further contribute to our understanding of poverty, in particular within the context of the developing world. Findings within this thesis thus, add to growing literature, in particular on the intersection between theoretical arguments and the empirical application of multidimensional approaches to measuring poverty. This thesis argues that the Delphi provides a coherent and transparent framework for improving the current practice of operationalising the somewhat metaphysical arguments underpinning the need to shift the focus of poverty evaluation to include, as A. Sen (2011) dubs it, 'the pluralities of the different features of our lives and concerns' (p. 233). Recognising that poverty is such a 'plural' issue, implies that varying judgments should be taken into account within its measurement and that these choices may significantly affect results (Nolan & Whelan, 1996, p. 2). The Delphi offers a way to synthesize these views by facilitating the collection of qualitative information, whilst also permitting quantitative analysis of whether consensus may be formed with regard to judgments on *which* aspects of welfare to include and *how* they should be combined to reflect the experience of poverty.

Before illustrating how the Delphi may be applied, however, this thesis aims to first 'take stock' and review existing methods of measuring poverty, both uni- and multidimensional. Examining these different measures will facilitate the assessment of whether focusing on uni-dimensional monetary-based measures, such as consumption alone is enough, or if a supplemental approach which takes into account the experience of multiple areas of deprivation is also needed. Furthermore, comparing past multidimensional measures to a 'newly' suggested approach based on Delphi results, seeks to contribute to Chiappero-Martinetti and von Jacobi (2012)'s call for 'discussion' on the implications of 'methodological choices' and 'contextual factors' on the extent of poverty and the socio-economic profiles of those who experience it.

Considering the above, the overarching research question of this thesis is thus: *do monetary-based unidimensional approaches to measuring poverty, particularly those utilising consumption expenditure as a benchmark, versus those applying a more multidimensional framework, lead to 'mismatches' with regard to the extent of poverty and who is identified as poor?* As Morduch (2006) states, 'questions of measurement are not matters of mere description; the way poverty is gauged affects how progress is determined, the implications of which, go beyond any given country at any given moment and are critical for future understandings' (p. 25).

Although not particularly intended to be on poverty within this nation per se, this thesis uses Indonesia, the world's fourth most populous country, as a case study. Indonesia's diverse provinces and districts provide ample testing ground to investigate the above questions and to illustrate the workings of the Delphi method suggested within this thesis. In addition to this, the topic of multidimensional poverty has remained nebulous within literature focused on welfare issues in Indonesia. Very few pieces of empirical research exist within literature aiming to discover 'which' multiple deprivations matter, 'how' extensive the problem is and 'who' are more likely to experience multidimensional poverty in this nation (Sumarto & de Silva, 2014). Among those that do, focus is given to the size of the poor, with empirical findings pointing to the need for considering both uni- and multidimensional approaches to truly understand the extent of poverty in Indonesia (Hanandita & Tampubolon, 2015; Sumarto & de Silva, 2014). As Milne (2006) uncovers, 'non-income poverty is a more serious problem than income poverty in Indonesia'. Alternative measures based on multidimensional deprivations, may thus reveal significantly divergent patterns with regard to *how many* and *who* are considered poor, thus leading to significant consequences for poverty eradication.

The following sections within this chapter are organised as follows: Section 1.2 describes the conceptual and theoretical framework guiding this thesis. The section summarises the key arguments for the importance of considering multidimensional poverty, the methodological challenges of implementing this approach and ways in which current measures have attempted to mitigate these challenges. This discussion frames the main research questions of this thesis, listed within Section 1.3. Section 1.4 outlines the 'mixed method' analytical strategy this thesis proposes, to tackle the research questions noted within Section 1.3. Section 1.5 introduces the case study area. Section 1.6 summarises details on the Indonesian National Socio-Economic Survey (Susenas), the principal data source for all quantitative analysis within this thesis. Finally, Section 1.7 presents a summary of the empirical chapters, which address the research questions outlined within Section 1.3.

1.2 Conceptual Framework

This thesis has two main objectives: (1) to review past measures, uni-dimensional and those encompassing multiple aspects of deprivation in order to identify areas where further development may be of benefit and (2) to suggest a novel way of improving these existing empirical measures through the use of the Delphi method. By examining these alternative measures, this thesis hopes to shed light on the possible miss-matches that may occur with regard to the size and composition of the poor, alongside how poor households are ranked, as a result of the different choices made within measurement. With regard to these choices, as touched upon above, two main challenges exist. Firstly, following Grusky, Kanbur, and Sen (2006), 'consensus is yet to be achieved on *which* dimensions [of poverty] matter, nor on *how* to decide what matters' (emphasis added, p. 12). Secondly, is the issue of how 'weights' attached to these various dimensions should be selected (Decancq & Lugo, 2013; Martin Ravallion, 2011a).

This section aims to summarise ways in which prior research has proposed to mitigate the above challenges and serves to inform, alongside justifying the decision to implement the methodological strategy outlined within Section 1.4. However, before presenting this summary, it is useful to briefly discuss the theories prompting interest in and outlining the importance of multidimensional poverty. The discussion within the sub-sections below thus also acts as background to the research questions listed within Section 1.3 of this chapter.

1.2.1 Uni- vs. Multidimensional Measures

This section aims to explore why reliance on monetary based measures, such as income and consumption expenditure may not be enough, and to guide the forming of expectations

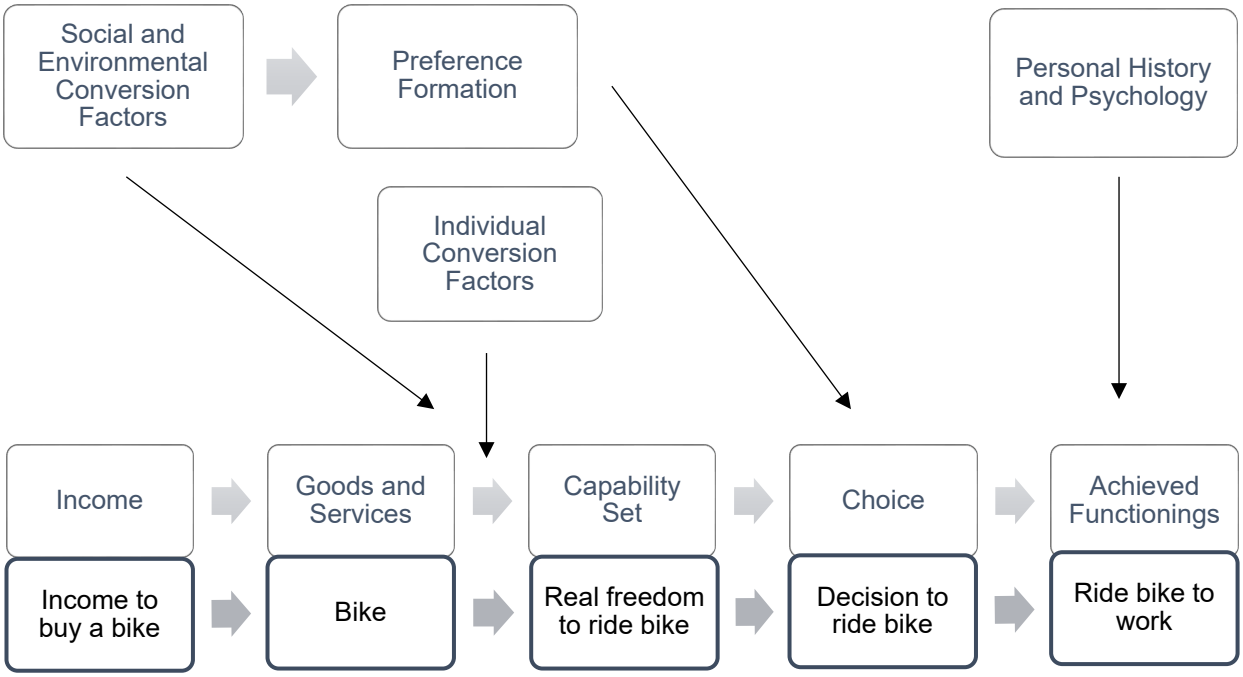
alongside explain the possible mechanisms behind the potential mismatches found when comparing these two forms of measurement.

The debate on the inadequacy of monetary based measures, particularly within the context of the developing world, started as a result of Amartya Sen's (1979a, 2001) seminal work on poverty and on his theoretical framework known as the 'Capability Approach', hereafter the CA. It was in his 'Tanner Lecture on Human Values' at Stanford University in May of 1979 that Sen first presented the radical realisation of the possibility that economics, statistics and policy have been basing measurements of welfare and human values on too simple a criterion. Amartya Sen (1979a) stated that there is evidence that the conversion of 'goods' to 'capabilities' varies from person to person substantially and 'the equality of the former may be far from the equality of the latter' (p. 218). This suggests a 'missing link', connecting 'command over goods' to what Sen dubs as the 'actual freedoms' people are able to enjoy.

Within the CA, Sen proposes an alternative to 'material prosperity' as the goal for development and argues that, development should be seen as a 'process of expanding the real freedoms that people enjoy' (Amartya Sen, 2001, p. 3). In arguing for this approach, Sen does not deny the role 'material' prosperity may have within society, however, considers its use as a means toward achieving freedom, rather than an end in itself. As Aristotle would say, 'wealth is evidently not the good we are seeking, for it is merely useful and for the sake of something else' (C. Rowe & Broadie, 2020). This 'something else', Sen would say, is the various things a person may value and have reason to value or 'functionings'. The freedom to achieve and enjoy these 'functionings' is then determined by a person's 'capabilities' (S. Alkire & Deneulin, 2009, p. 22). Figure 1.1 illustrates Sen's concept of 'capabilities' and 'functionings' and situates them with regard to measures of 'income and resources'.

From Figure 1.1, it can be seen that resources such as bicycles or the income to purchase a bicycle, are only means towards achieving the capability or freedom to actually ride a bike. The capability of the individual is the real freedom to use the bike, which although it presupposes strength or other abilities, fundamentally indicates the real freedom to travel on the bike from home to work, for example. The individual is thus free to decide on whether to translate this 'capability to ride' into the 'functioning' of using the bike to travel. A focus on income alone does not capture this freedom. What if the individual has the bicycle but is not capable of riding because she is weakened by sickness; the ability to own a bicycle thus does not guarantee the capability of riding it. As Nolan and Whelan (1996) explain, 'capabilities reflect the alternative combinations of functionings a person can achieve; functionings thus represent the various things a person manages to do or be in life' (p. 181).

Figure 1. 1 Conceptual Framework Explaining the Relationship between Income/Resources, Capabilities and Functionings



Source: Adapted from Robeyns (2005a) and S. Alkire (2014)

Robeyns (2005a) notes that the relationship between resources and functionings is influenced by three conversion factors, i.e. (1) personal (physical condition, skills, intelligence), (2) social (policy, norms, power relations) and (3) environmental or institutional factors (climate, geography) (p. 99). This author then argues that the CA takes into account human diversity, by its explicit focus on functionings and capabilities, alongside how the three factors may affect an individual's achievements in these domains. On this issue, one may argue that the use of equivalence scales, which take into account the effect of household size and composition on estimates of material resources available for each individual within the household, may somewhat account for Robeyn's 'individual' conversion factors when consumption or income measures are utilised. As equivalence scales are often constructed from consumption behaviour, the effects on welfare, which these scales may have, would also depend on views and societal norms on 'fairness' and how consumption allocations should be made. The interpretation of 'equivalence scales' would then be quite different if policy within said society was dominated by a male dictatorship (Martin Ravallion, 1994, p. 19). This disregard of other possible 'conversion factors' may contribute to the mis-match between reliance on income/consumption alone vs. multidimensional measures stemming from Sen's CA.

In addition to the above, another salient point emerges, i.e. within the CA, focus on functionings and capabilities does not imply that income or resources are unimportant as they may be 'instruments for the enhancement of freedom' (Drèze & Sen, 2002, p. 3). Despite this, however, as noted above 'they are not the ultimate ends of well-being' (Robeyns 2005a, p. 100). To compare with approaches advanced within the developed country setting, this idea is, to an extent, present within Ringen's concepts of 'indirect' vs. 'direct' measures of poverty. Ringen (1988) talks of the importance of the 'deprivation definition' of poverty as it relates to 'how we in fact live'. Although within this definition, poverty is understood as a 'low standard of consumption', similar to the CA, Ringen criticises reliance on income by placing focus on the significant mismatch between 'low income' and more 'direct' measures of deprivation (Whelan, Watson, & Maître, 2019). Laderchi, Saith, and Stewart (2003) note that the focus on deprivations now forms a central aspect of EU social policy, which acknowledges the inherently multidimensional nature of poverty, in particular through the adoption of the Social Exclusion, hereafter SE approach (p. 257). Following significant advances in data collection, measuring 'direct' deprivation within this approach, now moves beyond 'low consumption' to encompass many other aspects of life.

The SE approach was very much influenced by the pioneering work of Townsend (1979a), Mack and Lansley (1985) and Callan, Nolan, and Whelan (1993). It was developed in order to understand and describe the 'processes of marginalisation that can arise in rich countries with comprehensive welfare systems and was a reminder of the multiple faces of deprivation in affluent societies' (Laderchi et al., 2003, p. 257). Being largely based on Townsend (1979a)'s definition of poverty, i.e. individuals, families, groups can be said to be in poverty when they 'lack the resources to participate in activities and do not have the living conditions, which are customary or widely approved in the society to which they belong', the SE approach, focuses 'intrinsically on the structural characteristics of society and the situation of groups' (Guio, Gordon, Najera, & Pomati, 2017; Laderchi et al., 2003, p. 257). More importantly, however, Townsend (1979a)'s definition highlights the importance of the 'relative' nature of poverty. As Nolan and Whelan (1996) note, 'what is considered adequate, what are generally perceived as needs, will change over time and differ across societies'; poverty in that sense, should thus be recognised as a relative concept (p. 10).[1]

[1] Martin Ravallion (1994) points to the 'recognized fact' that the idea of 'absolute' vs. 'relative' poverty seems to act as a dividing point between poverty measurement in the context of the developing and developed worlds; with 'absolute poverty dominating the former, whilst relative poverty being more important in the latter' (p. 29). The vibrant debate between Townsend (1985) and Amartya Sen (1985), highlights this supposed divide. Within this debate, Townsend (1985) criticises Sen's CA and suggests that the approach is purely 'absolute' in its nature. Amartya Sen (1985), however, argues that the idea that poverty has an 'absolute' core, does not mean that 'needs may not vary from society to society or over time'. To make clear his point, Sen (1985) gives the example, 'if a person suffers

Nolan and Whelan (1996) highlight the links between the CA and SE approaches. They argue that Sen sees Townsend's 'being able to participate in the life of the community' as a 'specific capability'; 'non-monetary indicators of deprivation can then be seen, as direct, if crude, measures of success or failure in achieving particular concrete aspects of the broader functionings it entails' (p. 5). Furthermore, looking back at Figure 1.1, within empirical measurement, the SE approach aims to take into account 'preference formations' within its determination of poverty. This is through applying Mack and Lansley (1985)'s 'consensual' view of poverty, which acknowledges the importance of including socially perceived necessities and distinguishing between those who lack necessities from choice and those who cannot afford them, within the measurement of deprivation and poverty. This idea, to an extent, resonates with Sen's particular attention on freedoms that are *valued*. The word 'value' here is of central importance, for there needs to be a clear separation between beings and doings that enhance a person's well-being, from those that are 'evil and trivial' (S. Alkire, 2002, p. 7). The ability to include information on 'preferences' within the developing country sense, is however, challenging, as many surveys do not explicitly include information which aims to reflect this 'choice'.[1]

In a review of the different approaches to measuring poverty noted above, i.e. income and consumption expenditure, the CA and SE approaches, among others, Laderchi et al. (2003), note that the application of these different methods have different implications for policy as, to an extent, they point to different people as being poor; hence, 'definitions and measurement methods do matter for poverty eradication' (p. 243). Robeyns (2006) notes a number of quantitative studies, within the context of both the developing and developed worlds, which have investigated whether multidimensional approaches identify 'the same people' as poor, when compared to those based on income or expenditure. Within the developing country setting, in particular, research by Klasen (2000) compared the use of an expenditure-based measure of poverty with a composite measure of deprivation composed of fourteen indicators, which included a list of 'basic capabilities' as well as 'stated priorities of the population' using data from the 1994 SALDRU household survey conducted in South Africa.[2] It was found that the two measures 'diverged greatly in identifying the poorest and most deprived sections of the population' (Klasen, 2000, p. 33). Furthermore, also within South Africa, although using a different technique to measure the multidimensionality of poverty, Qizilbash (2002) in an analysis of 1996 Census data, applied 'fuzzy set theoretic

from hunger, the diagnosis of this condition cannot then be changed by the fact that others too may be hungry, so that this person may not be, relatively speaking, worse off than the others' (p. 670).

[1] Indonesia's Susenas is a case in point

[2] Klasen's composite measure included 'Education, 'Income', 'Wealth', 'Housing', 'Water', 'Sanitation', 'Energy', 'Employment', 'Nutrition', 'Healthcare' and 'Perceived Wellbeing'.

poverty measures' to examine 'vulnerability in various dimensions of the quality of life'. To define 'human poverty', Qizilbash (2002) included dimensions covering household expenditure, educational attainment, water source, energy source for cooking, employment and an indicator measuring overcrowding within the household.[1] This author found significant distinction between 'financial' and 'human' poverty as measured by these various dimensions. Closer to Indonesia, within an assessment of multidimensional poverty in Vietnam for years 2007, 2008 and 2010, Tran, Alkire, and Klasen (2015) comparing a multidimensional index including dimensions of health, education and standard of living, with household consumption, found that 'those who were monetary poor (and non-poor) are not always multidimensionally poor (or non-poor)' (p. 20-21). Specifically for Indonesia, although as touched upon at the start of this chapter, limited literature exists, Sumarto and de Silva (2014) and Hanandita and Tampubolon (2015), all found that the question of 'who is poor' has different answers when measures reflecting multidimensional poverty are used as opposed to consumption alone [2]

In sum, regardless of the approach taken and the context in which measurement takes place, poverty measures aiming to include 'direct' indicators to reflect the 'limited lives that people actually lead', versus the use of income or consumption expenditure alone, generate a different magnitude of poverty and identify different processes leading to the experience of poverty (Anthony B Atkinson & Marlier, 2010; Layte, Nolan, & Whelan, 2001; Nolan & Whelan, 1996; Stiglitz, Sen, & Fitoussi, 2010b). The importance of considering the multidimensional nature of poverty, alongside monetary measures, is thus, undeniable. The next question is then: how best to empirically measure and evaluate poverty in the multidimensional space? Considering this thesis' particular focus on the developing world, approaches to measuring multidimensional poverty framed within this context are specifically drawn upon.

1.2.2 Challenges in Applying a Multidimensional Framework

As noted at the start of this chapter, Amartya Sen (1976a) states that 'in the measurement of poverty, one must face two distinct problems: (1) identifying the poor among the total

[1] These areas of deprivation were included within measurement as they were argued to be salient for policy. In addition to this, these indicators were found to be 'easy to use' in conjunction with fuzzy measures (Qizilbash, 2002, p. 761).

[2] For the context of the developed world, among others, Balestrino and Sciclone (2001), in a study of the 'functionings-poor' in Italy found that 'a sizeable size of the poor in affluent societies are actually not income poor'. P. Anand and Van Hees (2006) in an analysis of a survey of self-reported functioning outcomes, found that, 'higher income levels are associated with lower capability satisfactions' (p. 276). In addition, results within Callan et al. (1993) and Nolan and Whelan (2010), using Irish data, show that 'employing both income and deprivation criteria rather than income alone can make a substantial difference to both the extent and composition of measured poverty'.

population and (2) constructing an index of poverty using the available information on the poor' (p. 219). Thorbecke (2007) notes the difficulties one has to face when solving these problems within a multidimensional framework; as touched upon within the previous sub-sections, not only does 'identifying' the myriad aspects of wellbeing pose significant challenges, but also, 'most serious' is the 'estimation of interactions between these attributes' when attempting to 'combine' them. On this issue S Alkire and James Foster (2011) note that 'a key direction for research has been the development of a coherent framework for measuring multidimensional poverty that is analogous to techniques developed in unidimensional space' (p. 476).[1] Bourguignon and Chakravarty's (2003) pioneering paper suggested such an approach. They proposed a simple multidimensional extension of the Foster-Greer-Thorbecke (FGT) index which fulfils a series of desirable axioms whilst preserving properties present within the income measure's use of this index.[2,3]

Bourguignon and Chakravarty (2003) note that, 'a simple way of dealing with multidimensional poverty is to assume that various attributes of an individual may be aggregated into a single cardinal index and that poverty may be defined in terms of that index, i.e. an individual can be said to be poor if her index of aggregate well-being falls below some poverty line' (p. 27). These authors, however, argue that this approach would be 'restrictive' in the sense that it would 'amount to considering multidimensional poverty as single dimensional'. Considering this, they argue that multidimensional poverty should be defined as a shortfall from a threshold on *each* dimension of an individual's well-being (emphasis added, Ibid). A 'direct' method to check whether an individual is poor according to the multidimensional framework suggested by Bourguignon and Chakravarty (2003) is then, to assess if she has the subsistence level to ensure non-deprivation within *each* aspect of poverty included in its measurement (p. 28). Anthony B Atkinson (2003) dubs the method of determining the total number of attributes in which an individual falls short of the desired minimum threshold as the 'counting approach' (p. 51).

To summarise whether an individual is poor according to this 'counting' approach, Anthony B Atkinson (2003) distinguishes between two approaches: (1) the 'union' approach, whereby those deprived in *any* attribute are considered poor and (2) the 'intersection' approach, whereby those deprived in *all* attributes are considered poor. The 'intersection approach

[1] Similar to the approach taken within this thesis, S Alkire and James Foster (2011) note that 'the main approaches to multidimensional poverty measurement are axiomatic'. Some have, however, suggested other approaches, such as fuzzy sets and latent variable analysis. I refer to Zadeh (1965), Chiaperro-Martinetti (1994) and Qizilbash (2006), for reviews of the fuzzy set approach, and Kakwani and Silber (2008) for the latent variable approach.

[2] Tsui (2002) provides an in-depth discussion of these axioms.

[3] As will be described within Chapter 2, the Indonesian Central Bureau of Statistics, also utilises the FGT to compute its headline consumption poverty statistics.

thus, somewhat mirrors Bourguignon and Chakravarty (2003)'s idea above, which states that to be 'poor' one has to be deprived in all the necessary attributes thought to describe the experience of poverty. Although providing ease of interpretation, both the 'union' and 'intersection' approaches, however, tend to identify extreme levels of poverty; with the 'union' approach veering towards identifying all the population as poor, particularly when a large number of attributes or dimensions are considered and the 'intersection' approach tending to 'miss' those who experience 'extensive deprivation even if they are not universally deprived' (Sumarto & de Silva, 2014, p. 8).

Aiming to improve on the above approaches S Alkire and James Foster (2011) developed an alternative method which aimed to couple the use of the 'counting approach' with an 'adjusted' FGT measure that 'reflects the breadth, depth and severity of multidimensional poverty' (p. 477). They introduced the use of a 'dual cutoff' approach to identifying the poor. The first cutoff, similar to Bourguignon and Chakravarty (2003)'s 'attribute specific' cutoff, identifies whether an individual is deprived within a specific attribute. The second cutoff then delineates how widely deprived a person must be in order to be considered poor. By applying the 'counting approach', this second cutoff then defines the 'minimum number of attributes' a person needs to be deprived in, in order to be considered multidimensionally poor. The use of this second cutoff thus provides an 'intermediate option', between the 'union' and 'intersection' approaches (S Alkire & James Foster, 2011, p. 479).

S Alkire and James Foster (2011) argue that this approach, known within literature as the Alkire-Foster (A-F) method, works well not only in situations with many dimensions, however, also enables the use of ordinal data within its construction. Whelan, Nolan, and Maître (2014) also note that this approach fulfils desirable axiomatic properties which 'allows one to examine in a structured way the implications of key measurement choices for levels of multidimensional poverty, the dimensional profile and socio-economic processes involved' (p. 184). In addition, the simplicity this method offers has led to its widespread application within policy; most notably is its use within the computation of the Global Multidimensional Poverty Index, hereafter Global MPI.

The Global MPI was first introduced within the 2010 UN Human Development Report. It has since been utilised to monitor multidimensional poverty within more than a hundred developing nations (Sabina Alkire, Usha Kanagaratnam, & Nicolai Suppa, 2019; S. Alkire, Kanagaratnam, & Suppa, 2020). The measure's widespread application and its use of the Alkire-Foster method to aggregate the poor, has prompted this thesis' interest in the measure. As will be explored further within Chapter 2 of this thesis, despite its use of this

axiomatic approach and its policy relevance, the measure has however, received a substantial amount of criticism.

These critiques are primarily aimed at the ‘arbitrary judgment’ basing the decisions behind the selection of its components, i.e. the types of dimensions and indicators it includes, alongside the choice of ‘equal’ weights used to aggregate these multiple indicators (Martin Ravallion, 2011b, p. 245). Indeed Robeyns (2006) notes that the ‘question of which capabilities are relevant to include within a multidimensional measure has provokes a lot of debate in the capability literature’ (p. 355). Sen himself, has ‘always refused to endorse one specific well-defined list’ (Robeyns, 2005a, p. 105). This does not, however discount the importance of said list within the construction of empirical measures; as Amartya Sen (2004a) states, ‘the problem is not with listing, but with insisting on one predetermined canonical list chosen by theorists without any general social discussion or public reasoning’ (p. 77). On this issue, S. Alkire (2007) states that, ‘no single set of domains or combining techniques will always be relevant and one of the important strengths of the CA, the theoretical approach underlying the Global MPI, is that researchers can employ many techniques, selecting those most relevant for each context’ (p. 91). Given the importance of multidimensional poverty and this need to take into account contextual issues within its construction, it is crucial to thus ask: which and how should one select dimensions and indicators, alongside the weights to combine these aspects, within a multidimensional measure?

As will be explored within Chapter 2 of this thesis, the Global MPI argues that its choice of dimensions is based on public consensus as they mirror the three components, i.e. Education, Health and Living Standards, included within the widely accepted Human Development Index (HDI). Martin Ravallion (2011b) however, questions whether the HDI’s dimensions indeed reflect public consensus as ‘there is little sign that debates within the public forum have led to new choices with regard to these components’ (p. 245). To respond to these critiques and to reflect the need for context specific measures, Sabina Alkire et al. (2019) acknowledge the benefit of countries adopting ‘national’ level multidimensional poverty indices, which aim to include aspects that reflect national priorities (p. 6). A number of these ‘national’ MPIs have been constructed, most notable of which are the Mexican Multidimensional Poverty Index (Mexican MPI) and Bhutan's Gross National Happiness Index (GNH). Also relevant for this thesis is a multidimensional measure constructed by a local NGO to reflect policy priorities in Indonesia, which I call the Adjusted MPI.

The Mexico MPI selects dimensions and indicators based on Mexico’s ‘National Policy for Social Development’ as governed within Mexico’s ‘General Social Development Law’. To

supplement the use of this law, however, an independent public institution specifically geared toward evaluating and monitoring social development, Mexico's CONEVAL, was also formed in 2005 (Licona et al., 2010).. CONEVAL undertook extensive participatory processes with well-known experts on poverty measurement in order to support the selection of these dimensions and indicators (Elizalde et al., 2010, pp. 10-11). As for weights, similar to the Global MPI, the Mexican MPI employs equal nested weights, as a way of promoting the ease of interpretation of the index. Also driven by a locally mandated law, Bhutan's GNH index, was formed to reflect values stated within Bhutan's legal code and constitution (Ura, Alkire, Zangmo, & Wangdi, 2012, p. 17). To select indicators within this measure, extensive deliberative and participatory processes were conducted. These processes included stakeholders at various levels ranging from private meetings with government officials and civil servants to Focus Group Discussions (FGDs) with Bhutanese citizens (Ibid, p.13). As for GNH weights, robustness of a range of weights was tested by testing whether district ranks changed, if different weights were applied. Similar to the above two 'national' measures, the Adjusted MPI employed FGDs with local Indonesian experts and academics to select its indicators, however, utilised equal weights to aggregate these components (Budiantoro, 2016; Budiantoro, Martha, & Sagala, 2020).

In sum, from the discussion above, significant progress is evident with regard to attempts at applying a multifaceted framework to measure poverty and well-being. The axiomatic approaches offered within Bourguignon and Chakravarty (2003) and S Alkire and James Foster (2011) in particular, have certainly eased a number of 'difficulties' faced within the formulation of such indices of poverty. The localised measures touched upon above, then provided insight as to, how one may address the challenge of selecting components within these indices. Learning from these advancements, particularly referring to the 'national' measures summarised above, two main areas where improvement may benefit, however, are still evident.

Firstly, these past 'national' measures point toward the key role participatory methods play in ensuring the relevance of the aspects of well-being to include within a measurement and that best describe poverty within each country setting. On this issue, however, Amartya Sen (2004b) talks of 'comprehensive outcomes', i.e. that the focus of a study should not only be on *which* components are important, however, also on *how* they were chosen. A way of transparently reporting the process by which these domains of poverty were selected and whether reliable consensus was achieved, could thus serve to ensure the transparency and robustness of these choices. Secondly, although participatory methods were used to select attributes of well-being, the use of equal weights to combine these aspects, seems to dominate these past measures. Decancq and Lugo (2008) note that 'as any other weighting

scheme, the use of equal weights implies trade-offs between dimensions, which may be considered reasonable or not' (p. 15). Explicit discussion and open debate on the choice of weights is thus, also crucially needed.

These issues in mind, the next sub-section outlines key aspects of participatory methods and acts as a guide to the analytical framework described within Section 1.4 of this chapter.

1.2.2 The Role of Participatory Methods

The use of participatory methods within the measurement of poverty and well-being was pioneered by Chambers (1994) and was developed in order to enable 'local people to share, enhance and analyse their knowledge of life and conditions, to plan and to act' (p. 57). Within this approach, 'contextual methods of analysis are utilised, which attempt to understand poverty dimensions within the social, cultural, economic and political environment of a locality' (D. Booth, 1998; Laderchi et al., 2003).

Laderchi et al. (2003) explain that a range of tools may be used to support the participatory approach, i.e. Focus Group Discussions (FGDs), participatory mapping, wealth and well-being rankings. Considering this, methodological guidelines for participatory analysis are often flexible, unlike other approaches where 'a more rigid framework is involved' (Ibid, p. 261). This is also true for sample selection within these approaches, as achieving a representative sample is generally unfeasible. Sample size then also relates to the 'subjectivity' of data generated by these approaches, which may question the robustness of conclusions arrived upon as a result. Despite these challenges, particularly with regard to the measurement of multidimensional poverty, participatory approaches provide a legitimate way of solving the significant problem of which definition of poverty may be appropriate for a specific setting and consequently, which dimensions to include within a measurement that aims to reflect this definition (S. Alkire, 2007; Laderchi et al., 2003, p. 262).

Considering the above, in order to ensure that participatory processes work well and to reduce possible biases which may occur when utilising this approach within poverty measurement, two main issues need to be addressed: (1) whose voices need to be heard and (2) how to synchronise these views, as disagreements are inevitable given the multitude of dimensions, which may describe the experience of poverty.

With regard to the first issue, Norton and Stephens (1995) suggest 'expanding participatory involvement to various stakeholders' and argue that participation at 'two levels'; i.e. 'primary and institutional levels' is needed in order to arrive at a more thorough understanding of poverty (p. 1-4). These authors define 'primary stakeholders' as the 'intended beneficiaries of development programs', i.e. the poor and marginalized, whereas 'institutional

stakeholders' refers to 'actors in civil society and governments, NGOs, academics and opinion leaders'. With regard to 'who' participates, Cornwall (2008) warns that 'being involved in the process is not equivalent to having a voice'; 'people need to feel able to express themselves without fear or the expectation of not being listened to' (p. 278). Cooke and Kothari (2001) state the fact that 'power relations are evident everywhere' needs to be taken into account when conducting participatory discussions.

The second issue refers to the need to achieve at least some level of *consensus* with regard to the dimensions and indicators that matter for the measurement of poverty, alongside their relative weights. Townsend (1979b) notes that this is crucial to enable evaluation of whether a household or a person falls short from these standards. A. Sen (2011) dubs this challenge of integrating diverse views on important aspects of poverty as the 'danger of non-commensurability' and suggests the use of what he calls 'partial orderings'. Through these partial orderings, 'one may search for areas where congruence of different dimensions occurs and we may confine ourselves, at least for now, to these areas where views intersect' (Amartya Sen & Foster, 1997). Although incomplete, this list of dimensions arrived at through this exercise would still bring progress towards understanding the myriad aspects of poverty that are detrimental to life.

As touched upon at the beginning of this chapter, this thesis argues that the Delphi method provides a participatory avenue for discovering how best to define and measure poverty, alongside mitigating the possibility of power imbalances, which may influence these choices. This is particularly so as the Delphi allows groups of individuals to anonymously provide their opinions through an iterative series of questionnaires. Within this process, individuals are given access to the views of others through presenting them with 'feedback' between the iterative rounds. This 'feedback' consists of anonymous summaries of the outcomes of previous rounds. The Delphi's ability to parsimoniously include a large group of individuals within its process of building consensus, to ensure a degree of anonymity of these participants, alongside enable statistical assessment of whether views do indeed, converge, provides a desirable alternative to the use of FGDs within participatory research (Harold A. Linstone & Turoff, 1977, p. 3). Before delving further into the analytical framework of applying this method, Section 1.3 below outlines the main research questions this thesis aims to tackle.

1.3 Detailed Research Questions

The conceptual framework explored above underlines the importance of measuring poverty within a multidimensional framework. This is even more crucial as possible mismatches between multidimensional and purely monetary-based measures, may lead to significant

policy implications. Given this significance of multidimensional poverty, of equal importance is the potential impact of the choices made within measurement, i.e. how the types of dimensions, indicators and weights utilised, alongside contextual issues, may affect the number and kinds of households identified as poor. Following the discussion above, this section presents the key research questions that this thesis aims to investigate:

i. Firstly, to learn from past measures: Are household ranks, the size and composition of the poor sensitive to these different measures, i.e. consumption vs. existing multidimensional measures? By 'existing measures', this thesis refers to the Global MPI and since the case study area for this thesis is Indonesia, the Adjusted MPI is also replicated.
ii. Through the use of the Delphi method: Can consensus be achieved with regard to which dimensions, indicators and respective weights, to include within a 'new' multidimensional poverty measure?
iii. How should this 'new' measure be formulated, i.e. in dashboard or composite form? Before constructing a 'composite' measure, what would a dashboard measure reveal on poverty issues within the case study area?
iv. Does applying weights resulting from the Delphi exercise within a 'new' composite measure of multidimensional poverty, identify different households as poor, when compared to the use of the commonly applied 'equal' weights?
v. Finally, does the 'newly' constructed composite measure based on Delphi results, lead to differing conclusions with regard to identifying the poor, when compared to the existing measures examined within this thesis?

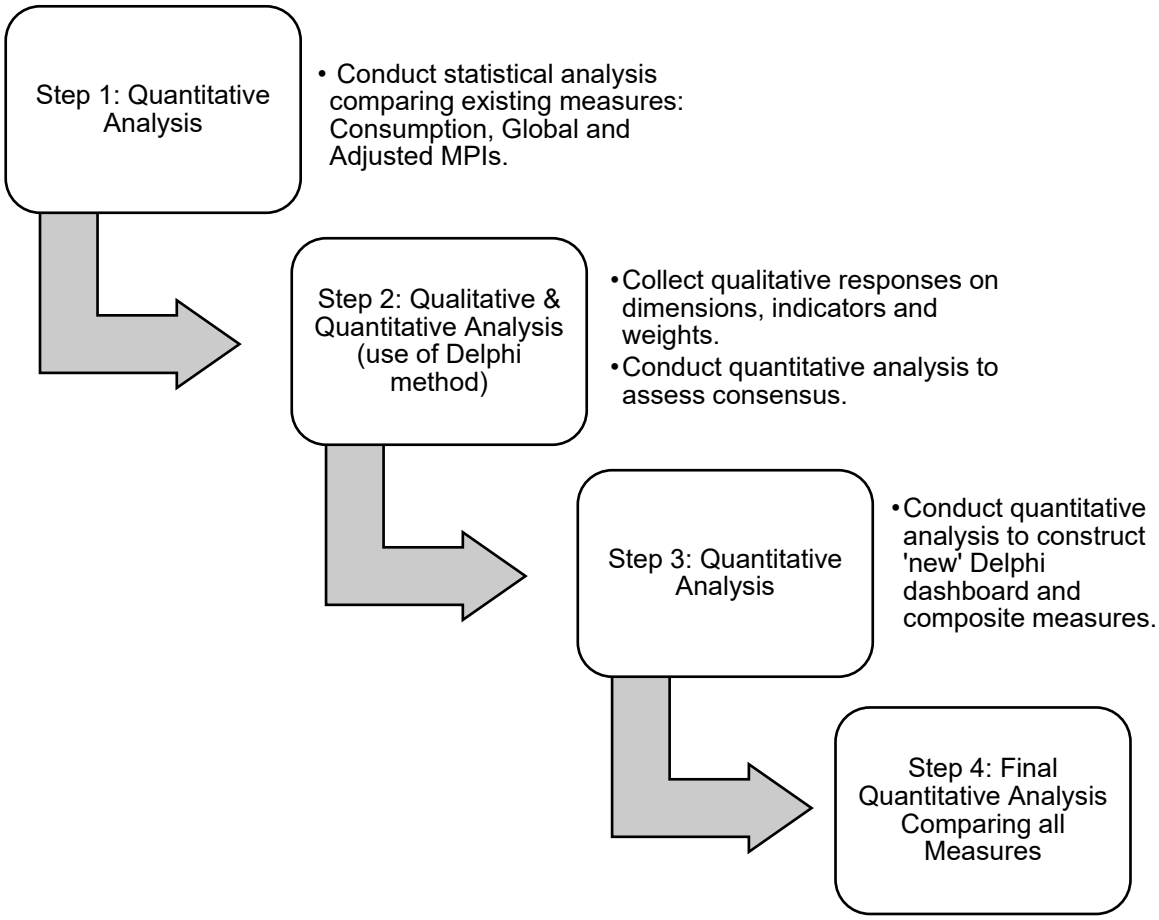
1.4 Analytical Framework

This section aims to outline the methodological steps taken within this thesis. Following Brannen (2005)'s definition of 'mixed-method' research, this thesis employs more than one type of method by coupling both quantitative and qualitative methods, in order to provide comprehensive and clear answers to the research questions listed within the previous section. The steps within this 'mixed-method' approach are illustrated within Figure 1.2.

Before delving into these steps however, a brief note of clarification of key concepts and what constitutes multidimensional poverty within this thesis is warranted. On this issue, this thesis follows the axiomatic method developed within S Alkire and James Foster (2011), i.e. in that multidimensional poverty is measured by deprivation within numerous 'indicators', which are grouped together to form 'dimensions' of poverty. As will be described within Section 1.5, the household is taken as the unit of analysis within this thesis. The 'dual cutoff' system will be employed to define deprivation and poverty within these households; i.e.

cutoffs will first be specified, which set minimum thresholds to define what it constitutes to be deprived within each chosen indicator. A second, 'poverty cutoff', will then be applied, to define the number of weighted indicators, a household needs to be deprived in to be considered multidimensionally poor.

Figure 1. 2 Methodological Steps within This Thesis



Source: Author's summary

From Figure 1.2, in order to address the first research question noted within Section 1.3, the first step of the analytical framework utilises quantitative statistical analysis of existing measures of poverty, i.e. the Consumption, Global and Adjusted MPIs. The Indonesian National Socioeconomic Survey (Susenas) is employed within this step. As Townsend (1979a) argues, 'any attempt to justify a new approach towards the definition and measurement of poverty must begin with previous definitions and evidence' (p. 32). It is thus crucial to learn from these past measures, before suggestions for improvement can be made. Exploring and identifying differences between the Consumption and existing multidimensional measures serves to justify the importance of constructing a 'new' and

'revised' measure of multidimensional poverty; whilst, comparing the Global and Adjusted MPIs aim to highlight whether contextual differences, which may drive the selection of dimensions and indicators within measurement, may lead to significant changes in how measures identify the poor. This paves the way to the implementation of the second step.

Within the second methodological step, given the importance of considering a multidimensional framework and the crucial role participatory methods play in providing solutions to the thorny issues of 'how' and 'which' dimensions, indicators and weights to include within measurement, the Delphi method is suggested. As touched upon within the previous section, through a series of iterative questionnaires, the Delphi exercise conducted within this thesis aims to collect *qualitative* opinions on components to include within a 'new' measure of multidimensional poverty, which aims to be relevant for the particular context where measurement is needed. Although similar to other participatory methods in that representative samples are often unattainable, the Delphi explicitly outlines the parameters used to select participants within its process. These parameters aim to guide and ensure the inclusion of 'experts' within the study, whose knowledge of poverty is crucial to the formation of the final measure.

As for how 'possible disagreements' may be synchronised, the *qualitative* open-ended nature of the Delphi enables the collection of information on these choices, whereas the *quantitative* part of the process enables objective statistical analysis of whether consensus can be achieved. The iterative process of the Delphi enables participants to review how the group has responded and the possibility of revising or strengthening their choices, given this group response. Furthermore, the Delphi's aim of maintaining anonymity is thought to be conducive to 'independent thought', which leads to a gradual formation of group opinions (N. Dalkey & Helmer, 1963, p. 2). Anonymity also aims to mitigate the biases resulting from possible power relations between participants. Furthermore, guidelines set *before* the Delphi takes place, particularly on defining and measuring consensus, aim to ensure objectivity in the study's conclusions (Crisp, Pelletier, Duffield, Adams, & Nagy, 1997). Results from this second step aim to shed light on the second research question listed within Section 1.3, i.e. whether consensus on the types of dimensions, indicators and weights to include, can indeed be achieved.

Within the third step, 'new' multidimensional measures based on Delphi results are constructed. To address the third research question, a literature review on the merits of both 'dashboard' and 'composite' approaches is conducted, alongside analysis of Delphi results on this 'well debated issue' of which 'form of measurement would best describe multidimensional poverty' (Chiappero-Martinetti & von Jacobi, 2012). As this exercise found

use in both forms, this step employs quantitative analysis of pre-existing survey data, i.e. information available within the Susenas, to construct both measures. Firstly, to address the third research question, a dashboard is constructed. Then, to address the fourth research question, a 'composite' measure is formed. For this 'composite' measure, particular attention is given to the possible impact of the choice of weights resulting from the Delphi study, specifically when compared to the use of the commonly applied 'equal weights', on the identification of the poor.

Finally, the fourth methodological step aims to compare the 'new' Delphi generated 'composite' measure with the existing measures examined within the first step. This quantitative step is again conducted using Susenas data to assess whether the different measures examined within this thesis, Consumption, Global, Adjusted MPIs and the 'new' Delphi composite measure, identify different sizes and compositions of the poor, alongside whether households are ranked differently. This step thus provides solutions to the final research question of this thesis.

With the above steps in mind, the next section aims to briefly illustrate the case study area chosen for the Delphi study. Even though this thesis' main aim is to illustrate how the Delphi can be applied as a novel and effective way of mitigating challenges within the construction of a multidimensional measure, rather than being about poverty in Indonesia per se, this short description is warranted as a way of outlining the decisions behind the selection of the case study area.

1.5 Case Study Area

Indonesia is a vast archipelagic nation, with a population of more than 260 million people, whose homes are spread across more than one thousand islands, more than thirty provinces and five hundred districts and municipalities. Collecting data to encompass the views and opinions of all the peoples of Indonesia is thus understandably, a mammoth task, particularly for one researcher with limited resources. Thus, it is only feasible, as Hicks says, to 'select' (Hicks 1983). A decision was thus made to focus on Java as it has the largest number as well as the highest concentration of the nation's poor.[1]

Java itself is vast; it is where nearly sixty percent of Indonesia's population dwells, spread across six provinces, including the capital of Indonesia, DKI Jakarta. Out of this vast population, close to ten percent, i.e. nearly thirteen million people, are consumption poor. Due to this vastness, it was thus only feasible to follow Hick's advice and 'select' once more

[1] For the Voices of the Poor Study on Indonesia, Mukherjee (1999) also stated this reason behind the study's decision to focus on Java.

a particular province and district to focus on.[1] This in mind, the district of Bogor City, within the province of West Java, was selected. There are three main reasons behind this thesis' final focus on the Bogor City: (1) the decentralised nature of how poverty alleviation policies are formed and implemented in Indonesia, (2) the growing urbanisation of the country and (3) the ease of access to the district. A brief overview of West Java is given before delving into details on Bogor City.

Figure 1.3 presents a geographical map of Indonesia and the location of West Java within the nation. It also provides a close-up image of West Java and its districts, including Bogor City. As can be seen within the map of Indonesia, large disparity exists within the nation with regard to the number of poor households across these provinces and districts. Relatively more moderate numbers of poverty can be seen within the Western provinces, whereas significantly higher levels of poverty are evident within Indonesia's Eastern provinces.[2] At close to 11 percent Consumption poor, West Java, relative to the other provinces, somewhat lies between the extremes. When zooming into West Java to look closely at its districts and municipalities, however, a similar pattern of disparity to that of Indonesia seems to emerge; i.e. the province is comprised of districts with relatively low and those with very high levels of poverty. The similar level of disparity evident in West Java and Indonesia is confirmed by the similar Gini ratio recorded for both these areas.[3] Looking closely within West Java, Bogor City's Consumption poverty rate is also shown to be in between the extremes.

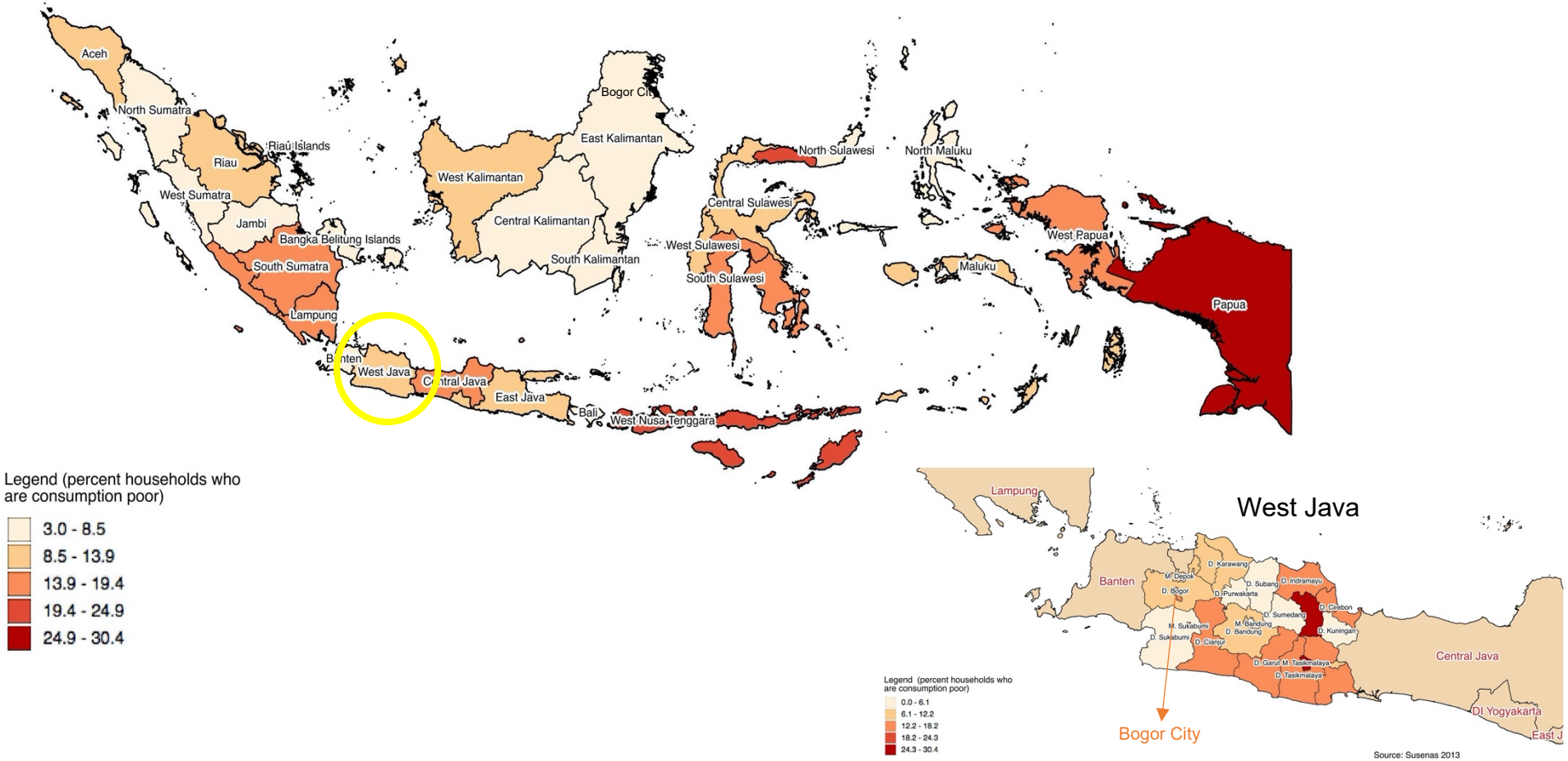
The first reason to justify the selection of a district level case study; i.e. Bogor City, is due to Indonesia's governmental structure and the way it influences how poverty alleviation policies are formed and implemented. Silver, Azis, and Schroeder (2001) note that in 1999, decentralisation within Indonesia took place and local governments were granted considerable autonomy. This ultimately led to a 'fundamental shift from central to local dominance, with districts and municipalities receiving greatly expanded functions' (Brodjonegoro, 2002, 2003, 2004; Holtzappel & Ramstedt, 2009; Niazi, 2012). This 'local dominance' also gave districts and municipalities the right to formulate their own poverty alleviation policies and conduct research to support this effort (Erb, Faucher, & Sulistiyanto, 2013; Hadiz, 2010).

[1] Each province in Indonesia is divided into smaller administrative sectors called districts (kabupatens) if they are predominantly rural and municipalities (kota) if they are predominantly urban (Usman, 2002, p. 1). West Java has twenty-seven smaller administrative sectors; eighteen are rural districts and nine are urban municipalities.

[2] M. Ravallion and Bidani (1994) and Miranti (2010) explore this disparity among Indonesia's provinces in further detail.

[3] The Gini ratio in 2013 for both West Java and Indonesia was 0.41. In March 2019, it was 0.4 for West Java and 0.38 for Indonesia (Central Bureau of Statistics (CBS), 2019).

Figure 1. 3 Extent of consumption poor across Indonesia, with a focus on West Java[(a)]



Note: [(a)] Maps were developed using QGIS; source of spatial data: DIVA-GIS (2012); source of poverty data: Susenas 2013

Source: Author's computations

The aim of this divergence of power is to enable the formulation of policies which match the specific needs of each district. This was thought to be crucial for a nation as diverse as Indonesia. The decentralised nature of where policy is formed in Indonesia, prompted the decision to select a particular district or municipality as a case study area to conduct the Delphi proposed within Section 1.4.

The second reason for selecting Bogor City, i.e. a predominantly urban area, as a case study, refers to the rapid urbanisation of Indonesia as a nation. On this issue, Lewis (2014) points to empirical data illustrating that 'the proportion of the Indonesian population that is urban grew from 17.1% in 1960 to 52.6% in 2009' (p. 194). This points to the importance of urban issues, particularly related to poverty in Indonesia and acts to reinforce the decision made to select an urban case study for this thesis, before additional analysis can be made within other areas of Indonesia when time and resources permit.

On 'who' specifically to include within the study, one notes that in research involving those who are experiencing poverty, often the word 'poverty' cannot be used within discussions as it implies judgement on those participating; participants often believe the term to be 'too loaded' (Hall, Leary, & Greevy, 2014). In many cases, it is necessary to replace the word 'poverty' with other words, such as 'well-being' or 'welfare'. It thus becomes a challenge to disentangle the definitions arrived upon through these opinions; whether they indeed reflect 'poverty' or whether they refer to broader issues concerning welfare. In addition, Clark (2003) also warns of the possible biases arising from 'adaptive preferences', i.e. a condition in which a poor person 'learns to take pleasure in small mercies and desires nothing more', when including opinions of those suffering poverty. Acknowledging that the case study area within this thesis is urban, even though the threat of these 'adaptive preferences' may be small, it may still seriously weaken results. Although far from ideal, in order to focus the research on 'poverty' and before the possibilities of 'adaptive preferences' may be taken into account within results, a *preliminary* focus on policy makers is made within this thesis. This thesis thus aims to act as a first step, aiming to test the methodology proposed, ensure its reliability and effectivity, before further studies including the poor are made. Furthermore, including policy makers who are involved directly in forming poverty alleviation programmes in Bogor City, also aimed to promote discussion and direct action, with regard to the forming of policies based on multidimensional measures.

Considering the above, following Goldstein (2002), gaining access to these 'policy makers' was crucial and indeed, required substantial hard work. This leads to the third reason for the selection of Bogor City as a case study. The City's efforts in recognising that in order to alleviate poverty, multiple dimensions of deprivation need to be addressed, as stated within

its mayoral manifesto, acted as a starting point for this study (Sugiarto, 2016). In addition to this, the City's openness and eagerness to listen to new ideas about how best to assess and support evidence-based poverty alleviation policies, particularly expressed by its Mayor at the time of the study, was crucial in ensuring access and commitment toward the study proposed within this thesis. Fully aware that as Lijphart (1971) states 'a single case study can constitute neither the basis for a valid generalization nor the ground for disproving an established generalization', it is my hope that the analysis and methodological advancements proposed within this thesis may act as a guideline, which may be replicated within different contexts, not only in Indonesia, but also within other developing or developed nations. This in mind, before concluding this chapter, a brief note on the Susenas is presented below.

1.6 Principal Data Source

In addition to data collected from the Delphi exercise, empirical analysis within this thesis is conducted on the Indonesian National Socioeconomic Survey (Susenas).[1] This survey is of relevance as it allows the construction of *all* poverty measures analysed within this thesis, i.e. the Consumption measure, the Global MPI, Adjusted MPI and the 'new' Delphi generated dashboard and composite measures.

The Susenas is a nationally representative survey, which has been conducted by the Indonesian Central Bureau of Statistics (CBS) annually since 1963 (Australian National University (ANU), 2011, p. 3). It has been and continues to be the main data source for computing social and economic indicators on Indonesia. The parts of the Susenas dataset, which I will use, are the Susenas Core (which collects variables on household and individual level characteristics) and the Susenas Consumption Module (which collects data on household consumption). The Susenas Core and Consumption Module 2013 are nationally representative, covering all Indonesia's provinces and districts/municipalities; data is available every quarter and pooled at the end of the year, and is representative at the district/municipality, provincial and national levels (Central Bureau of Statistics, 2013b; H. Hardjo, 2010). [2]

[1] Use of the 2013 SUSENAS (Indonesian National Socio-Economic Survey) within this thesis is sponsored and supported by the Institute for Economic and Social Research, Faculty of Economics and Business, University of Indonesia (LPEM – FEB UI).

[2] In 2013, the Susenas recorded information that is representative for all of Indonesia's 33 provinces and 497 districts. After the data collection period for the 2013 Susenas ended, however, the Indonesian government ratified the existence of a 34th province, i.e. North Kalimantan and its five districts. This new province was formerly included within East Kalimantan.

The surveys employ a stratified sampling methodology to select household samples, using three stratification stages. Within the first stage, sampling areas are divided into urban, rural, slum and non-slum areas. Within these area classifications, sampling areas are selected using a Probability Proportional to Size (PPS) method based on the number of households within each sampling area (Central Bureau of Statistics, 2013b). From this first stage, 30,000 sampling areas are selected randomly from 154,557 potential areas. These 30,000 sampling areas are then allocated randomly and equally into each of the four quarters in that sampling year, giving 7,500 sampling areas to survey every quarter. In the second sampling stage, within every sampling area, one census block is chosen randomly. In the third stratification stage, within every census block, ten households are selected by systematic random sampling (S. H. Hardjo, 2010). Every quarter there are close to 75,000 households that are surveyed, thus, annually, a total of around 300,000 households are surveyed.

The indicators included within the 2013 Susenas Core Survey cover demographic characteristics, health, education, employment, fertility and contraception, housing, and asset ownership (Central Bureau of Statistics, 2013b). Data within the consumption module includes the quantity and rupiah value of consumption for 215 types of food commodities. The Susenas Module survey also includes data on consumption for non-food items, such as housing, clothing, taxes, insurance and religious needs. In total there are 108 types of non-food items included within the survey (Central Bureau of Statistics, 2013b, p. 2).

1.6.1 Limitations of the Susenas

Despite being unprecedented in scale, periodicity and scope, the Susenas has a number of weaknesses. These limitations pertain to three issues: (1) the potential underrepresentation of those living in informal settlements, (2) possible issues with consumption and income data within the survey and (3) possible missing values for indicators utilised to construct the multidimensional poverty indices.

With regard to the first issue, Baker et al. (2013) note that ‘informal marginal settlements are likely to be under-sampled within the Susenas, thus information on poverty is likely to be underrepresented; it is also, not possible to know by how much’ (p. 3). Following Kanagaratnam (2017) this thesis ‘assumed that these numbers are low because informal settlements are smaller in number when compared to slum households with secure tenure’ (p. 29). This fact is confirmed within a World Bank (2016) report stating that a large proportion of slums in Indonesia are categorised as ‘legal slums’ with secure tenure, i.e. the land where households settle ‘belongs to or is rented by residents and the residents’ right to occupy the land is recognised by the local government’ (p. 2-3). The Susenas includes these ‘legal slums’ within its surveys.

On the second issue, Johar, Soewondo, Pujisubekti, Satrio, and Adji (2019) note that it is important to acknowledge that 'expenditure data' in the Susenas is 'the total cost or consumption, comprising out-of-pocket expenses and/or subsidies' (p. 62). This becomes significant *if* this thesis were to 'create income variables based on expenditures recorded within the Susenas'. As Johar et al. (2019) state, if studies proxy income by calculating the sum of expenditures on all items, since the Susenas expenditure data contains subsidies, households may appear to have higher income than they do'. As for the 'income' variable, which is reported in the Susenas, many argue of its unreliability, as it is 'self-reported' income and thus is prone to under-reporting (Martin Ravallion, 1994; Martin Ravallion, Datt, & Walle, 1991). Considering this, consumption data is utilised as it is; an income proxy is not calculated within this thesis. As for the 'income variable' reported within the Susenas, following Piketty and Atkinson (2010), this thesis opts not to use this variable within its analysis; this became a pertinent decision particularly as the Susenas' question on 'earnings over the past month' may lead to 'seasonal variations in income and the moment during the year the survey was conducted may distort estimations' (Piketty & Atkinson, 2010, p. 211).

For the third issue, Chapters 2, 5, 6 and 7 note the treatment of possible missing data within indicators of poverty included in the analysis of this thesis. To summarise, guidelines within S. Alkire and Santos (2010b) and Sabina Alkire et al. (2019) utilised to base the calculation of the Global MPI are followed, i.e. 'that the sample size after the treatment of missing values must be reasonably high, both at national and sub-national levels'. In addition, as proposed within S. Alkire and Santos (2010b), where a significant number of missing values are present 'sensitivity checks aiming to compare the percent of deprived population in each of the other indicators in the group with missing values in certain indicators, with that of the group with observed values in said indicator' are also conducted (p. 28).

1.7 Summary of Chapters

This section provides an overview of the empirical and methodological chapters within this thesis.

1.7.1 Chapter 2: Comparison of Existing Measures

This chapter investigates whether the use of different existing poverty measures, i.e. the Consumption, Global and Adjusted MPIs, leads to different sizes and compositions of the poor, alongside differences in how households are ranked; thus, this chapter addresses the first research question within this thesis. Data from the Susenas was employed within this chapter. Comparing these three measures it was found that each measure identified differences with regard to the extent and types of households identified as poor by these

measures. Households were also assigned significantly different ranks, when each measure was utilised. Reliance on any one of these measures as a single yardstick, ‘the litmus test’, of policy effectiveness may thus be misleading.

1.7.2 Chapter 3: The Delphi Methodological Steps

This chapter describes the analytical framework proposed within Section 1.4 of this chapter in further detail. The Delphi was conducted for the purpose of answering the questions of which dimensions and indicators of poverty, together with their respective weights, should be included within a measure of poverty for the context of Bogor City, the case study area explored within this thesis. Furthermore, the Delphi aimed to achieve a reliable level of consensus with regard to these choices. This chapter provides a detailed illustration of how the Delphi was implemented, i.e. the pilot tests conducted, the construction of the Delphi questionnaire, decisions on the number of iterative rounds to include and the type of feedback to present to participants between each of these rounds, how participants were selected, the sample size of the Delphi and how consensus should be assessed. This chapter also acts as background, paving the way toward discussions on the Delphi results, the subject of Chapter 4 of this thesis.

1.7.3 Chapter 4: Delphi Results

This chapter aims to address the second research question within this thesis: which dimensions, indicators and weights should be included within a contextual measure of multidimensional poverty? Can consensus be achieved on these choices? Firstly, through the Delphi exercise, unanimous agreement was achieved with regard to the multidimensionality of poverty, indicating the need to supplement the Consumption-based measure utilised as the official measure of poverty in Indonesia. Secondly, a degree of consensus was achieved with regard to the choice of five dimensions of poverty to be included within this multidimensional measure; i.e. ‘Education’, ‘Employment and Asset Ownership’, ‘Living Standards’, ‘Health and Safety’ and ‘Child Health and Contraception’ and eighteen indicators to proxy deprivation within these dimensions. Finally, with regard to dimension and indicator weights, disagreements were still evident between participants on the exact values of these weights. Chapter 6 suggests a method to mitigate this level of disagreement in order to construct a ‘composite’ measure of multidimensional poverty. Before this, Chapter 5 illustrates results for a ‘dashboard’ measure.

1.7.4 Chapter 5: A Dashboard Measure of Multidimensional Poverty based on Delphi Results

To address the third research question, this chapter argues for the importance of both 'dashboard' and 'composite' measures when attempting to assess the extent of multidimensional poverty within a given society. This decision was supported by arguments within literature, however, more importantly by the results of the Delphi study. With this in mind, this chapter aims to focus on the creation of a 'dashboard' measure. When forming this 'dashboard', decisions needed to be made with regard to two issues. Firstly, are the 'indicator' cutoffs, i.e. the level of deprivation a household needs to experience to be considered as deprived within a specific indicator. Secondly are the 'dimensional' cutoffs, i.e. the number of deprived indicators a household needs to have to be considered as deprived within a dimension. For 'indicator' cutoffs, existing minimum welfare criteria, set within international and national policy standards, were utilised as guides. For 'dimensional' cutoffs, the widely recognised 'Union' and 'Intersection' approaches were used. In addition to these two approaches, an attempt to discover a robust 'intermediate' dimensional cutoff was also conducted.

The results of this chapter can be summarised within two main points. Firstly, the extent of indicator deprivations in Bogor City mirrors that evident within existing literature discussing urban poverty in developing countries. In addition, the trend evident within Bogor City could also be found when the geographical unit of analysis was widened to include West Java and Indonesia. This strengthens the arguments made, which suggest rapid urbanization in Indonesia (Lewis, 2014). Secondly, dimensional deprivation rates were found to be very sensitive to the dimensional cutoffs chosen. Without sound theoretical and empirical evidence basing the selection of these cutoffs, robust conclusions with regard to which dimension exhibits the highest number of deprived households, could thus, not be made.

1.7.5 Chapter 6: A Composite Measure of Multidimensional Poverty based on Delphi Results

This chapter addresses the fourth research question within this thesis, i.e. whether Delphi generated weights lead to significantly different outcomes when compared to the commonly applied 'equal nested' weighting scheme. Firstly, to address the degree of disagreement with regard to the exact values of the Delphi generated weights, as found within Chapter 4, this chapter suggests the use of a 'linear optimisation' method, rooted within the Data Envelopment Analysis (DEA) and 'benefit-of-the-doubt' approach to weighting. This enabled the formulation of 'expert' weights, which aimed to encapsulate the views of Delphi participants, reflect their policy priorities and to a degree, the levels of trade-offs between the

chosen dimensions and indicators. Two Delphi-based 'new' composite measures were then constructed; the first applying these 'expert' weights to aggregate the dimensions and indicators chosen by Delphi participants, the second using 'equal nested' weights to combine these aspects. Data from Bogor City is used within the analysis of this chapter and three main conclusions were found.

Firstly, household ranks when the two weighting schemes were applied were shown to be highly correlated. If importance is only placed on how households are ranked, given the resource constraints governments face, the use of the more parsimonious 'equal nested' weights, thus, may suffice; this is because both weighting schemes may likely lead to the same household being 'identified' as eligible to receive social assistance. Secondly, despite this high correlation, significant differences were found with regard to the extent of poverty identified by these two weighting schemes. Using the 'expert' weights as a benchmark, the use of the 'equal nested' weights, was found to likely underestimate the level of poverty within Bogor City. Thirdly, with regard to the composition of the poor. It was found that both weighting schemes tended to identify households with relatively similar characteristics, i.e. those headed by older individuals, who are less educated and live within households with additional caring responsibilities toward children, as more likely to be poor.

In sum, the findings of this chapter confirm prior analysis within existing literature, which contrasts the use of opinion-based 'expert' weights generated through a DEA 'benefit-of-the-doubt' approach, with 'equal' weights; i.e. despite difference in the size of the poor identified, similar ranks and compositions of the poor are found when these two schemes are compared. Even though this is the case, the use of 'expert' weights created a sense of ownership of the final composite measure, which in turn, generated significant discussion and concrete action with regard to the application of the 'newly' generated Delphi Composite measure within policy.

1.7.6 Chapter 7: Comparing the Delphi Composite with Existing Measures

This chapter aims to answer the final research question within this thesis on whether the Delphi generated composite measure identifies different sizes and compositions of the poor, alongside differing household ranks, when compared to the existing measures examined within Chapter 2, i.e. the Consumption, Global and Adjusted MPIs. Focus was given within this chapter to comparing the Delphi Composite, which applied 'expert' weights, to these existing measures. Data from Bogor City, West Java and Indonesia, were used within the analysis of this chapter. The main findings of this chapter are summarised within four main points.

Firstly, significant differences were found with regard to the extent of poverty identified by the Delphi generated composite measure, when compared to the three existing measures. Secondly, for the composition of the poor, household head's level of education and additional childcare responsibilities, were shown to be the only consistently significant variables influencing poverty status in *all* the four measures considered within this chapter. These measures thus differed with regard to a number of other household characteristics, such as the gender and marital status of household heads. Thirdly, household ranks according to the four measures also displayed only relatively moderate correlations.

The final finding taps into whether for a nation as diverse as Indonesia, multidimensional measurement should not only take into account national priorities, however, given the dominance of district level poverty alleviation efforts, whether district specific measures are needed. Thus, the comparative exercise between the four measures was conducted using data from three levels of analysis, district (Bogor City), provincial (West Java) and national (Indonesia). This is to test whether the same conclusions found above, also apply when poverty at the different geographical levels are analysed. On this, a number of different conclusions were found. To illustrate, within Bogor City, the Delphi generated measure was found to identify a larger extent of poverty when compared to the Global MPI. This was, however, not evident for West Java and Indonesia, where the Global MPI was found to exhibit a higher poverty headcount, when compared to the Delphi generated measure. Subtle differences were also found when comparing the composition of the poor, between the Delphi generated measure and the two existing multidimensional measures at different geographical levels. These conclusions suggest that internationally blanket type measures, such as the Global MPI or even nationally blanket-type measures, such as the Adjusted MPI, may not suffice to assess poverty within a nation as diverse as Indonesia. Thus, for Indonesia, following the nation's decentralised policy, district level measures, may also be needed.

Chapter 2: A Comparison of Existing Measures of Poverty

> *'Any attempt to justify a new approach towards the definition and measurement of poverty, so that its causes and means of alleviation may be identified, must begin with previous definitions and evidence'*
>
> *(Townsend, 1979a, pp. 32-33)*

2.1 Introduction

In light of the still relevant debate encapsulating the definition and measurement of poverty summarised within Chapter 1 of this thesis and following Townsend (1979a)'s argument for a review of existing methods before 'new' approaches may be proposed, this chapter aims to provide a comparative overview of the use of three existing poverty measures currently employed to measure poverty in Indonesia, i.e. (1) the Consumption poverty measure, (2) the Global Multidimensional Poverty Index, hereafter the Global MPI, and (3) what I call, the Adjusted MPI measure. Discussion on a proposed 'new' methodology, which aims to mitigate the possible weaknesses of these three existing measures, is then the subject of Chapters 3, 4, 5 and 6 of this thesis.

Martin Ravallion (2003) states that differences in definitions and measurement assumptions, may lead to 'dramatically different' positions on the extent of poverty and factors influencing its experience (p. 1). Faced with this dilemma, it seems pertinent to ask, which measure should we believe? For to quote Piachaud (1987)'s well referred to argument, poverty 'carries with it the implication and moral imperative that something should be done about it' (p. 161). Clearly, 'one cannot help the poor without at least, knowing who they are' (World Bank, 2005, p. 10). Considering these arguments, through replicating the three past measures, this chapter aims to answer the following four research questions. Firstly, *do these three measures identify a different size of the poor?* Secondly*, do they assign differing ranks to households?* Third, *do they identify different types of households as poor?* Finally, *do the results of the three research questions above differ when different geographical levels of analysis are considered, i.e. whether district, provincial or national level data is used?* This final question is of particular relevance, as it enables the examination of whether internationally developed measures, such as the Global MPI, suffice, when attempting to measure multidimensional poverty, particularly within a diverse nation, such as Indonesia, or whether more localised measures are needed. In addition to this, addressing the above questions aims to provide a basis to determine whether 'reliance on one particular measure' may 'mislead the extent and nature of poverty', thus, possibly resulting in unreliable and ineffective policy (Nolan & Whelan, 1996, p. 60).

Analysis within this chapter will employ the nationally representative Susenas dataset. One year of analysis is chosen, year 2013, for the reason that complete access to this particular year was available. As touched upon above, three levels of geographical analysis are conducted within this chapter, i.e. (1) national level, using data from Indonesia, (2) provincial level, using data from West Java, the province where Bogor City is located and (3) district level, using data from Bogor City, the case study area chosen for this thesis. The household is utilised as the unit of analysis for all three levels.[1] All statistical computations within this chapter were conducted using Stata 14.

The rest of this chapter is organised as follows. Section 2.2 describes the existing measures examined within this chapter. Section 2.3 presents the results of the four research questions above. Section 2.4 concludes.

2.2 Examining Existing Measures

This section aims to discuss the theoretical underpinnings which base the three existing measures examined within this chapter; the Consumption, Global and Adjusted MPIs. In addition to this, given the importance of Amartya Sen (1979b)'s statement noted within Chapter 1 of this thesis, on the two steps one should take when measuring poverty, a review of how these existing measures 'identify' and 'aggregate' the poor, is also presented. Following this, aspects where further development is needed are then highlighted. Table 2.1 presents a summary of the above details for each measure. I start with a description of the Consumption measure.

2.2.1 Consumption

The Consumption measure is one of the 'most commonly applied' monetary based measures of poverty within the developing world; indeed, for Indonesia it is no exception, for since the 1980s, this measure has been utilised as the 'sole' official measure to assess poverty across all the nation's provinces and districts (Cahyat, Gönner, & Haug, 2007; Martin Ravallion, 1994; 2010, p. 2; Martin Ravallion, Chen, & Sangraula, 2009; World Bank, 1990, 2005).[2]

[1] The Susenas defines 'household' as 'a person or a group of persons who live together within a physical building and who eat from the same kitchen' (Central Bureau of Statistics, 2013b).

[2] In 2013, the Susenas recorded information that is representative for all of Indonesia's 33 provinces and 497 districts. After the data collection period for the 2013 Susenas ended, however, the Indonesian government ratified the existence of a 34th province, i.e. North Kalimantan and its five districts. This new province was formerly included within East Kalimantan.

Table 2. 1 Summary of the Three Existing Measures

Measure	Years in Use	Theoretical Underpinning	Identifying the Poor	Aggregating the Poor	Areas for Further Improvement
Consumption	1984-Present	Basic Needs Approach	Unidimensional: uses one indicator, i.e. 'Consumption Expenditure' of food and non-food items, to identify the poor. Food and non-food items are selected to reflect the consumption pattern of a 'representative population'.	Use of prices for food and non-food items to determine a poverty line and aggregate the measure. Poor if: a household's monthly per capita expenditure is below this poverty line.	Consumption expenditure may not reflect deprivation in other aspects of life. The existence of goods and services, which do not have market prices.
Global MPI	2010-Present	Capability Approach	Multidimensional: uses 3 dimensions and 10 indicators (selected to mirror those included in HDI)	Uses 'equal nested weights' to determine total weighted deprivation within the ten indicators. Poor if: total weighting deprivation shows that household is deprived in 33 percent of the ten weighted indicators.	The use of 'equal' weights to reflect the degrees of importance and trade-offs between dimensions and indicators.
Adjusted MPI	2012-Present	Capability Approach	Multidimensional: uses 3 dimensions and 11 indicators (selected to mirror those included in the Global MPI and to reflect national values)	Uses 'equal nested weights' to determine total weighted deprivation within the ten indicators. Poor if: total weighting deprivation shows that household is deprived in 33 percent of the eleven weighted indicators.	The use of 'equal' weights to reflect the degrees of importance and trade-offs between dimensions and indicators.

Source: Author’s summary

The Government of Indonesia's 'Consumption' measure bases its application on the 'basic needs' approach. With roots going back to Rowntree (1901)'s idea of primary poverty, this approach was brought to the fore by an increased concern for the need to ensure the satisfaction of elementary necessities of a population. Thus, to 'identify' the poor, this approach proposed the inclusion of a broad list of items, which aim to reflect a household's minimum fulfilment of food, clothing, shelter, water, sanitation and elements of services, such as education and health facilities (International Labour Office (ILO), 1976, p. 32; Streeten, 1994; Streeten, Burki, Haq, Hicks, & Stewart, 1981, p. 25). To 'aggregate' the poor, Streeten et al. (1981) note that there has never been unanimous agreement on how this is best done, however, the common practice is to use prices to combine these different items and to create a poverty line specifying the minimum amount of consumption expenditure a household needs to achieve in order to fulfil these basic needs (F. H. Ferreira & M. A. Lugo, 2013, p. 12; Martin Ravallion, 1994, pp. 23-27, 75-76).

Considering this, in Indonesia, the nation's Central Bureau of Statistics, hereafter the CBS, computes two poverty lines, i.e. one determining a minimum threshold for the consumption of food items, the other of non-food items. The final poverty line is then the sum of the values for both these lines (Cahyat et al., 2007; Central Bureau of Statistics Indonesia (CBS), 2014; Iriana, Avenzora, & Abidin, 2012; Martin Ravallion, 1994, p. 75).

To identify which goods and services to include within the above two poverty lines, the CBS utilises the consumption pattern of what it calls its 'reference' population, i.e. those who are thought to be 'vulnerable' to poverty and whose consumption expenditure are 'on' or 'twenty-percent above' the inflation-adjusted 'previous year's' poverty line (Central Bureau of Statistics Indonesia (CBS), 2014; Iriana et al., 2012). This information is collected through an additional survey, i.e. the Survey of the Basic Needs Commodities Package, known as the SPKKD survey (Central Bureau of Statistics Indonesia (CBS), 2014; M. Ravallion & Bidani, 1994, p. 77).[1] Two 'reference populations' are determined to collect two types of 'bundles of goods and services', one relevant for urban areas, the other for rural settings. Finally, to determine the poverty lines, district level commodity prices are utilised to weight the goods and services within these bundles (Central Bureau of Statistics Indonesia (CBS), 2014; Surbakti, 1995).

[1] In practice, the CBS does not collect 'new' information on the 'reference' population every year. To illustrate, the latest SPKKD survey was conducted in 2004. This is also an area of possible weakness of the Consumption measure, i.e. the fact that consumption patterns may change should be taken into account within the choice of goods and services included in the calculation of this measure.

Due to the diversity of the nation, the CBS estimates poverty lines for each district.[1] Poverty headcounts are estimated for these districts using the Foster-Greer-Thorbecke (FGT) headcount measure (Maksum, 2004).[2] Currently, Consumption poverty levels are published every quarter to support the rapid evaluation of government services and policy (Central Bureau of Statistics Indonesia (CBS), 2014).

Considering the above, three main drawbacks of the Indonesian Consumption measure are recognised. Firstly, price data used to compute consumption expenditure, alongside to determine the poverty lines, are often 'incomplete and unreliable', in particular for the non-food items included within the index (M. Ravallion & Bidani, 1994). Furthermore, Sumarto and de Silva (2014) note the possible measurement errors which may occur due to the challenges of collecting information on these prices, alongside data on the quantities of goods and services consumed by households. These authors warn that 'recall errors' often lead to 'downward biases' with regard to the information collected on these aspects, which could then contribute toward the Consumption measure identifying less poverty than that experienced in reality. The second weakness refers to the fact that, although consumption expenditure is estimated by the CBS, at the household level, individual per-capita consumption is utilised to determine the poverty headcount. This is of particular concern, due to the absence of the use of equivalence scales within computations, which may lead to a degree of bias when analysing the demographic profile of poverty (Lanjouw & Ravallion, 1995). Finally, as argued for within Chapter 1 of this thesis, the use of a monetary indicator is often invoked because it may not appropriately proxy other aspects of welfare and poverty (Laderchi et al., 2003, p. 248). On this issue, Maksum (2004) notes that the Indonesian Government's 'consumption expenditure' measure is a tool to only assess 'consumption deprivation'; poverty may thus be more than that measured by this approach.

Considering the above, the following sub-section discusses the steps taken and challenges encountered when replicating the Consumption measure for the purposes of this chapter.

2.2.1.1 Replicating the Consumption Measure

To calculate the poverty headcount for the Consumption measure, two variables are needed: (1) per-capita individual consumption expenditure and (2) district level poverty lines. Per-capita consumption is available within the Susenas' 2013 'Consumption Module. Re-calculating poverty lines using this dataset was, however, not possible, as additional

[1] The Susenas 2013 has information on poverty lines for all of Indonesia's 497 districts. The CBS also computes 'provincial' and 'national' poverty lines. These lines are, however, mostly utilised as bases against to compare district level lines.

[2] J. Foster, Greer, and Thorbecke (1984a) offer further discussion on the FGT method.

information on the proportion of different non-food items consumed by the 'reference population', was not available to the public. This in mind, the pre-specified poverty lines available from the Central Bureau of Statistics (CBS) website, were utilised (Central Bureau of Statistics, 2013a).[1,2]

As proposing the use of different equivalence scales is beyond the scope of this chapter and in order to ensure ease of comparison between the different measures of poverty examined within this thesis, the 'household' is utilised as the unit of analysis. Consumption poverty is then determined by identifying households, whose *per capita daily consumption expenditure for food and non-food items is below that specified by its district poverty line*.

Before ending this sub-section, I note that by utilising pre-specified poverty lines, an exact replication of the poverty numbers reported by the CBS, was not possible. This is most likely due to the fact that, the CBS generates poverty lines every quarter. Publicly available information on these lines was however, limited to those for the March and September quarters. In addition, differences in the proportion of the poor reported, may also be due to the CBS publishing 'individual', as opposed to 'household' level poverty. Even though this was the case, when comparing data published by the CBS and that resulting from the exercise conducted within this chapter, significantly large differences were not found.

2.2.2 The Global MPI

The Global Multidimensional Poverty Index, hereafter, Global MPI, is one of 'the most well-developed and broadly applied MIPs to date' (Martin Ravallion, 2011b, p. 236). The discussion below aims to provide further details on the steps taken within the construction of this measure.

As touched upon within Chapter 1, the Global MPI is based on Sen's Capability Approach (CA), i.e. that the assessment of progress should be based on whether 'the freedoms that people have are enhanced' (Amartya Sen, 2001, p. 4). S. Alkire and Santos (2010a), however, note significant data restrictions faced when constructing multidimensional measures based on the CA, particularly, the limited information, which exist within surveys on 'capability sets'; focus, is thus, often restricted to 'achieved functionings'. In addition, as the exercise of measuring poverty becomes less demanding when attention is placed on the 'shortfalls' of well-being that are experienced, rather than focusing on 'functionings' and 'freedoms' that improve welfare, the Global MPI, thus focuses on such 'un-freedoms'. S.

[1] These poverty lines were also available within the World Bank's Data Archive for Indonesia; i.e. the INDO-DAPOER online data archive (World Bank, 2015). Both data sources, thus reported the same poverty lines.

[2] There is no missing data in the Susenas 2013 for 'consumption expenditure' and the poverty lines.

Anand and Sen (1997) justify this focus on ‘un-freedom’ by stating that ‘agreement on states of unfreedom is easier to obtain when compared to a set of ample freedoms’ (p. 44). S. Alkire and Santos (2010a) follow this argument when developing the Global MPI and constructs the measure based on poverty as a shortfall from well-being.

Considering the above, the Global MPI ‘identifies’ the poor according to their deprivation status, i.e. achieved ‘un-freedoms’, within three dimensions; ‘Health’, ‘Education’ and ‘Living Standards’. Ten indicators are used to proxy deprivation within these dimensions. To enable the exploration of joint distributions of deprivations, the Global MPI aggregates its dimensions and indicators by applying weights; this is in contrast to the Consumption measure’s use of prices (S. Alkire & Santos, 2010a). These weights are assigned equally to each dimension and equally to indicators within each dimension. Table 2.2 lists the Global MPI’s dimensions, indicators and their respective weights.

Referring to Table 2.2, as noted within Chapter 1, the dimensions within the Global MPI, mirror those within the Human Development Index (HDI). For indicators within the measure, a consultation process involving experts across all the three dimensions was conducted to base their selection (S. Alkire & Santos, 2010a, p. 5). To aggregate across these dimensions and indicators, in addition to the use of weights, a ‘dual cutoff’ system is applied. Thresholds for each indicator are first determined, before an overall poverty cutoff is applied. To define deprivation within each indicator, existing internationally recognised minimum standards are referred to. Households that are deprived in an indicator are assigned a score of ‘1’ for that indicator; if they are non-deprived, they are assigned a score of ‘0’. For the poverty cutoff, robustness checks are conducted to assess whether different thresholds lead to significant differences with regard to how nations are ranked according to their poverty levels. Based on these checks, the Global MPI utilises the poverty cutoff of: *an individual is considered poor if he or she lives within a household that is deprived in one third or more of their ten weighted indicator deprivations*.

Finally, as touched upon within Chapter 1, the Global MPI is calculated using the Alkire-Foster (A-F) method, in particular, the M_0 within the A-F family of measures. This method takes into account two components, the simple headcount ratio, i.e. the percent categorised as multidimensionally poor within the population (H) and the average deprivation share among the poor (A). The Global MPI is then the product of ‘H’ and ‘A’ (S. Alkire & Santos, 2010a, p. 3). For the purposes of this chapter however, to ensure comparability between the Global MPI and Consumption measure, which is a headcount measure, inclusion of the intensity of poverty (A) is overlooked.

Table 2. 2 Summary of Components within the Global and Adjusted MPIs (following published guidelines on these measures)

Global MPI					Adjusted MPI				
Dimension	Dimension Weight	Indicator	Indicator Weight	Indicator cutoff (individuals within a household are deprived if:)	Dimension	Dimension Weight	Indicator	Indicator Weight	Indicator cutoff (individuals within a household are deprived if:)
Education	0.333	Years of Schooling	0.167	No household member age 10 years or older has completed six years of schooling.	Education	0.333	Education continuity	0.11	Household has at least one school aged child that is not attending school until they complete high school education.
		Child School Attendance	0.167	Any school-aged child is not attending school up to the age they'd finish Grade 8.			Access to pre-school education	0.11	Household with at least one pre-school child that does not have access to pre-school education.
							Literacy	0.11	Household has at least one family member aged 15-64 years that is unable to read and write.

Table 2.2 (continued)

Global MPI					Adjusted MPI				
Dimension	Dimension Weight	Indicator	Indicator Weight	Indicator cutoff (individuals within a household are deprived if:)	Dimension	Dimension Weight	Indicator	Indicator Weight	Indicator cutoff (individuals within a household are deprived if:)
Health	0.333	Child mortality	0.167	Any child has died in the household.	Health	0.333	Sanitation	0.083	Same as Global MPI.
		Nutrition	0.167	Any adult under 70 years of age or any child for whom there is nutritional information is malnourished.			Clean drinking water	0.083	Same as Global MPI.
							Birth attendant	0.083	A member within the household has experienced a birth process without the help of a trained health worker.
							Under 5 nutritional intake.	0.083	Households with under five children have at least one child who cannot fulfil daily needs of: 70-220 grams of carbohydrates, 15-35 grams of protein, 35-62 grams of fat and 637.5-1600 kkals.

Table 2.2 (continued)

Global MPI					Adjusted MPI				
Dimension	Dimension Weight	Indicator	Indicator Weight	Indicator cutoff (individuals within a household are deprived if:)	Dimension	Dimension Weight	Indicator	Indicator Weight	Indicator cutoff (individuals within a household are deprived if:)
Living Standards	0.333	Electricity	0.056	Household has no electricity.	Living Standards	0.333	Electricity	0.083	Same as Global MPI.
		Improved Sanitation	0.056	Household's sanitation facility is not improved (according to SDG guidelines) or it is improved but shared with other households.			Fuel/Source of energy to cook	0.083	Same as Global MPI.
		Drinking water	0.056	Household's source of drinking water does not meet SDG definitions, or or the distance to this clean source is a 30-minute or more round trip from home.			Roof, floor and wall	0.083	Household without adequate roof, floor and wall.
		Housing	0.056	Household has inadequate housing material in floor, roof or wall.			House ownership	0.083	Household does not own the house they live in.
		Cooking fuel	0.056	Household cooks with dung, wood, charcoal or other solid fuels					
		Asset ownership	0.056	It does not own more than one of the following: radio, telephone, TV, computer, animal cart, bicycle, motorbike or fridge and does not own a car or truck.					

Source: Author’s summary

As noted within sub-section 1.2.2 of Chapter 1, despite the Global MPI representing significant advancements with regard to developing a way of operationalising the concept of multidimensional poverty into empirical measurement, criticism continues to exist on the use of this measure. I revisit these criticisms below.

Firstly, is the issue of the arbitrariness of how dimensions, indicators and weights are selected within the Global MPI. Martin Ravallion (2011b) in particular notes that, whilst starting with an 'ad hoc' choice, may be thought of as a first step toward generating additional public debate, there is 'little sign' that this discussion has occurred (p. 245). The second concern relates to how the Global MPI combines information on various deprivations within these components. On this, Ravallion notes that it has been more than twenty years since the HDI was first published (and ten years since the Global MPI was developed); both measures, however, still utilise equal weights within their measurement. Due to the dangers of basing the selection of weights on arbitrary assumptions, alongside the less demanding data properties which this alternative form of measurement offers, Ravallion argues for the relevance of a 'dashboard', instead of an 'aggregate' index (Oxfamblogs, 2010).[1]

2.2.2.1 Replicating the Global MPI

The Global MPI uses both household and individual level data within its measure. The 'Health' and 'Education' dimensions are measured at the individual level, whereas the 'Standard of Living' dimension is measured at the household level. As they are measured at the 'individual level', indicators within both the 'Health' and 'Education' dimensions, only apply to a sub-section of the population; what Sabina Alkire et al. (2019) call the 'applicable population' (p. 16).[2]

For the individual level indicators within the Global MPI, the deprivation status of individuals within the 'applicable population' in each household, influences the deprivation of the whole household. Santos and Alkire (2011) argue for this assumption by drawing on the Pigouvian and Coasian theories of the 'economics of externalities' (Coase, 1960; Pigou, 1929). Within a household, it is thus, assumed that members enjoy or suffer from shared positive or negative externalities as a result of a member's or members' deprivation status in the Global MPI's

[1] The debates around the issue of 'composite' vs. 'dashboard' measures is further explored within Chapter 5 of this thesis.

[2] Since they were first developed in 2010, guidelines published on the construction of the Global MPI have changed throughout the 10 years of its application, in particular with regard to differences in the stringency of indicator cutoffs. Although this chapter refers to the full set of methodological papers published on the Global MPI, those most recently produced in 2019 and 2020 are specifically utilised within the replication of the Global MPI in this chapter.

'Education' and 'Health' dimensions (Alkire et al., 2016, p. 15). Some argue that these assumptions introduce bias, particularly due to the possible effects of the size of the household on measurement. As attending to this bias is beyond the scope of this thesis, as noted within the Introduction of this chapter, the household, rather than the individual, is considered as the unit of analysis within this chapter.

Considering the above, the below description will particularly focus on the definitions of the 'applicable populations' for each of these 'Heath' and 'Education' indicators, as determined within the Susenas. In addition, this sub-section also highlights how missing values within the Susenas are treated. I start with indicators within the 'Education' dimension.

As listed within Table 2.2, two indicators are utilised to describe deprivation in 'Education'; i.e. 'Years of Schooling' and 'Child School Attendance'. For the 'Years of Schooling' indicator, a household is non-deprived if at least one person in the household has completed *six years* of education (S Alkire, U Kanagaratnam, & N Suppa, 2019; S. Alkire et al., 2020). For Indonesia, the official duration for primary level education is six years and the official primary school completion age is thirteen (UNESCO, 2017). The Global MPI methodological papers allow for one year of late entry or late completion, thus, within this chapter, a household is non-deprived if it has at least one individual within it who is 'aged 14 or older' who has completed six years of schooling.

The 'School Attendance' indicator only applies to households which have members who are of school age. Households that do not have school aged children are automatically considered as non-deprived in this indicator (Alkire et al., 2016). For this indicator, an age criterion is also put in place, i.e. a household is deprived if the household has any school-aged child that is not attending school until *the age they would finish grade 8* (emphasis added, Alkire et al. 2016, p. 8). For Indonesia, this age range is seven to fifteen years old, marking the start of primary school until the age a child should have completed Grade 8 (UNESCO, 2017).[1] No missing data are present for both the 'Education' indicators.

Within the 'Health' dimension, two indicators are used: 'Child Mortality' and 'Nutrition'. For 'Child Mortality', the 'applicable population' consists of households with 'eligible women'. Within the Susenas, these are 'women who are over ten years of age and have ever married' (Central

[1] For the computation of the Glpbal MPI within this chapter, as with the 'Years of schooling' indicator, one year of late entry and completion is applied (Alkire et al., 2016, p. 8; United Nations Development Programme (UNDP), 2015).

Bureau of Statistics, 2013b). Households which do not include these women are automatically considered as non-deprived in this indicator. There are ten households, out of a total sample of nearly 300,000 households for Indonesia, that have missing data in this indicator.[1] Following guidelines within S Alkire et al. (2019), households that lack data on any indicator are not included within the final analytical sample.

For the 'Nutrition' indicator, data on malnutrition, such as BMI for adults and weight-for-age for children, are not available for all households, within the Susenas. Following S Alkire et al. (2019), if a survey is missing any of the Global MPI's indicators, that indicator cannot be used in the final computation of the measure. Considering this, 'Nutrition' is thus excluded within the computation of the Global MPI in this chapter.[2] As this is the case, following Global MPI guidelines, the weight for 'Child Mortality' is adjusted so that the total dimensional weight for 'Health' continues to be 0.33. By conducting this, for this chapter's computations, the indicator weight for 'Child mortality' is thus 0.33 (S. Alkire & Santos, 2010a, p. 26).

Within the 'Living Standards' dimension, available information in the Susenas presented challenges for the computation of deprivation in two indicators, i.e. 'Drinking Water' and 'Ownership of Assets'. For 'Drinking Water', the cutoff for this indicator applied within this chapter, slightly differs from the official guidelines published for the Global MPI noted within Table 2.2. This is due to the fact that the Susenas does not collect data on 'distance to clean water source', however the 'distance of a household's water source to waste disposal areas' (Central Bureau of Statistics, 2013b). Following guidelines within S Alkire et al. (2019), 'if a survey report uses other definitions of clean drinking water, which differs from the official Global MPI indicator cutoff, the definition stated within said survey should be followed' (S Alkire et al., 2019, p. 8). This in mind, within this chapter, the Susenas' criteria is followed, i.e*. 'households are deprived in drinking water if their source of drinking water is not 'improved' according to SDG criteria and the distance between this source and waste disposal areas is less than 10 meters'* (Central Bureau of Statistics (CBS), 2020). A number of households within the Susenas, however, report missing values with regard to this 'distance from waste disposal' variable. Again, following S Alkire et al. (2019) guidelines, these households are not included in the final analytical sample.

[1] The Appendix to Chapter 7 presents a full list of indicators with missing values.

[2] Data on 'Nutrition' is also unavailable within the Demographic and Health Surveys (DHS) used by the Oxford Poverty and Human Development Initiative (OPHI) to compute their Indonesian Global MPI country reports (Oxford Poverty and Human Development Initiative (OPHI), 2017).

For the 'Assets' indicator, the types of assets collected within the Susenas slightly differ from those listed within the Global MPI's methodological guidelines. Referring to Table 2.2, the Susenas collects data on all these assets, except 'ownership of a radio and animal cart' (Central Bureau of Statistics, 2013b). Following guidelines within S Alkire et al. (2019) these particular assets are thus not included within the analysis in this chapter.[1]

In sum, taking into account the missing values within 'Child Mortality' and 'Drinking Water', the total number of households within the Susenas 2013 sample possessing complete data for the computation of the Global MPI leads to a retained sample size of close to ninety percent of the original sample for Indonesia. This is also the case if analysis is conducted for West Java and Bogor City. The Global MPI computational guidelines stipulate that in order to mitigate problems of representativeness, a country must have a national retained sample of at least 85 percent of the original sample and every sub-national region must have a retained sample of at least 75 percent of the original sample (S Alkire et al., 2019, p. 15). As shown above, the retained sample within the Susenas 2013 fulfils these criterions.

2.2.3 The Adjusted MPI (MPI for Indonesia)

The Adjusted MPI (MPI for Indonesia) was developed by a local Indonesian think-tank in 2015, under the guidance of the Oxford Poverty and Human Development Initiative (OPHI) with funding from the Ford Foundation (Budiantoro, 2016; Budiantoro et al., 2020; Oxford Poverty and Human Development Initiative (OPHI), 2016a; Prakarsa, 2016).[2] Although this measure is not utilised as an official measure by the Government of Indonesia, it is of interest to the writer of this thesis, as it offered a way of testing the effects of incorporating country specific indicators resulting from consultations with, for the most part, local academics, on the extent and composition of the poor when compared to the Global MPI's more internationally relevant indicators (Budiantoro, 2016; Budiantoro et al., 2020; Prakarsa, 2016).[3]

[1] Unlike the missing 'Nutrition' data, which are excluded within OPHI's official country level reports on Indonesia, the DHS dataset utilised within these reports includes 'Ownership of radio and animal cart (Indonesia Demographic and Health Survey 2017, 2018)'. Considering this, calculations of the Global MPI within this chapter may thus differ from those reported by OPHI.

[2] To note, this measure is not officially dubbed as the 'Adjusted MPI', Official publications refers to it as 'MPI for Indonesia'. I employ this label to indicate how this measure 'adjusts' the Global MPI, as it mirrors this international measure in almost every way, except the indicators it includes, which aims to reflect Indonesian values. This is also to distinguish this measure from the 'new' measure proposed within Chapter 6 of this thesis.

[3] Guidelines for the Adjusted MPI lists three Indonesian academics and two statisticians working within the Indonesian Central Bureau of Statistics (CBS), who were involved within discussions and committee meetings aimed at selecting indicators for this measure (Prakarsa, 2015).

As described within Table 2.2, the Adjusted MPI mirrors the Global MPI in almost all its aspects (both measures utilise the same three dimensions and the same 'equal nested' weighting scheme), except for a number of country specific indicators and the cutoffs this measure uses to define deprivation within them. This sub-section places particular focus on these indicators.

I will start with the 'Education' indicators, which differ from those included in the Global MPI. The first indicator, 'Education Continuity', as described within Table 2.2, aims to identify those within the household who are of school age until the age at which they are expected to complete high school education (Budiantoro et al., 2020).[1] This indicator somewhat resembles the Global MPI's 'Children's Education' indicator, however, utilises a more stringent cutoff. Although not specifically stated within official guidelines for the Adjusted MPI, this criterion may seek to reflect the Government of Indonesia's renewed commitment toward achieving universal access to education up until the 'high school' level (Republic of Indonesia Ministry of Education and Culture, 2016).

The second 'Education' indicator; i.e. 'Access to Pre-school Education', also refers to children's education within the household and defines deprived households as those, which 'have any child aged three to six who does not have access to pre-school education'. The inclusion of this indicator, may also reflect current policy priorities set by the government concerning the importance of early childhood education (Republic of Indonesia Ministry of Education and Culture, 2014a, 2014b, 2018).[2] As noted within Table 2.2, the Global MPI excludes 'Access to Pre-school Education'. Finally, the third indicator, 'Literacy', somewhat refers to 'Adult's Education', as it records skills of 'reading and writing' for those aged 15-64 within the household. The household is considered deprived in this indicator if 'at least one individual of this age range cannot read or write'. The Global MPI also excludes this indicator within its measurement.

Within the 'Health' dimension, two indicators not present within the Global MPI, are included. Firstly, 'Birth Attendant', similar to the Global MPI's 'Child Mortality' indicator, refers to 'households with women who are over ten years of age and have ever married'. In order for a household to be *non-deprived* within this indicator *all* such 'eligible women' within the household

[1] Taking into account the official school entry age in Indonesia and the Global MPI's guidelines to consider possible 'late entry and late completion' the 'eligible population' for this indicator is thus, those 'aged eight to twenty' (UNESCO, 2017). Households, which are deprived in 'Education Continuity' are thus those with *at least one* chid *aged eight to twenty who are not attending school until they complete their high school education* (Budiantoro et al., 2020).

[2] Sub-section 5.3.1.1 of Chapter 5 further discusses these policies.

should have access to health services provided by doctors or midwives when they give birth (Budiantoro et al., 2020; Central Bureau of Statistics, 2013b). Secondly, to base the inclusion of the 'Under Five Nutritional Intake' indicator, the Adjusted MPI argues for the use of 'micronutrient' data to replace missing data on 'weight-for-age' within the Susenas.

For the 'Living Standards' dimension, both the Adjusted and Global MPIs include an indicator assessing the adequacy of the dwelling a household is living in; difference is however present with regard to the cutoffs these measures use for this indicator. For the Global MPI, a house is inadequate if is uses rudimentary material for any of these three components, i.e. its roof, floor or wall (Sabina Alkire et al., 2019; S. Alkire et al., 2020). Whereas within the Adjusted MPI, a more stringent cutoff is applied, i.e.a household is deprived if its roof, floor *and* wall are rudimentary (Budiantoro et al., 2020; Prakarsa, 2016). Finally, the Adjusted MPI does not include the Global MPI's 'Ownership of Small Assets', however, includes an indicator, which assesses whether a household owns the house in which it lives.[1]

The above in mind, a final note is given on indicator weights within the Adjusted MPI. Although, both the Global and Adjusted MPIs apply 'equal nested weights' which implies that weights for the three dimensions within both measures are the same, the way indicators are grouped within these dimensions has led to differences in how (similar) indicators are weighed within these two measures. To illustrate, the indicators of 'Sanitation' and 'Drinking Water' are included within the Adjusted MPI's 'Health' dimension, whereas for the Global MPI, they are included in the 'Living Standards' dimension. As a result, these indicators are given different weights, i.e. 'Sanitation' and 'Drinking Water' are given weights of 0.083 by the Adjusted MPI, however 0.056 by the Global MPI. These indicator weights, thus, to an extent, seem to be dependent on two factors: the number of indicators within each dimension and which dimension an indicator is assigned to. This fact highlights a key area within the Adjusted MPI, which may benefit from further improvement, i.e., its use of the somewhat arbitrary 'equal nested' weighting scheme; as it is unclear whether these weights, reflect the degrees of importance or trade-offs between dimensions and indicators, or whether they are simply a result of the two factors noted above.

Another area of weakness within this measure is that although the selection of indicators within it was based on participatory discussion amongst a number of Indonesian academics, it is

[1] The reasons behind this difference in cutoff and the inclusion of 'House Ownership' are unclear. However, these indicators may aim to tap into the Government of Indonesia's increased recognition of the importance of land tenure to support poverty alleviation (Badan Perencanaan Pembangunan Nasional (BAPPENAS), 2018). Sub-section 5.3.5.1 of Chapter 5 further discuss these policies.

unclear how these opinions were formed and analysed; in particular, transparency is needed with regard to information on: how these 'poverty experts' were selected? How were ethical issues, such as 'dominant voices' handled? How was 'consensus' measured? Was 'consensus' on the choice of indicators achieved?

2.2.3.1 Replicating the Adjusted MPI

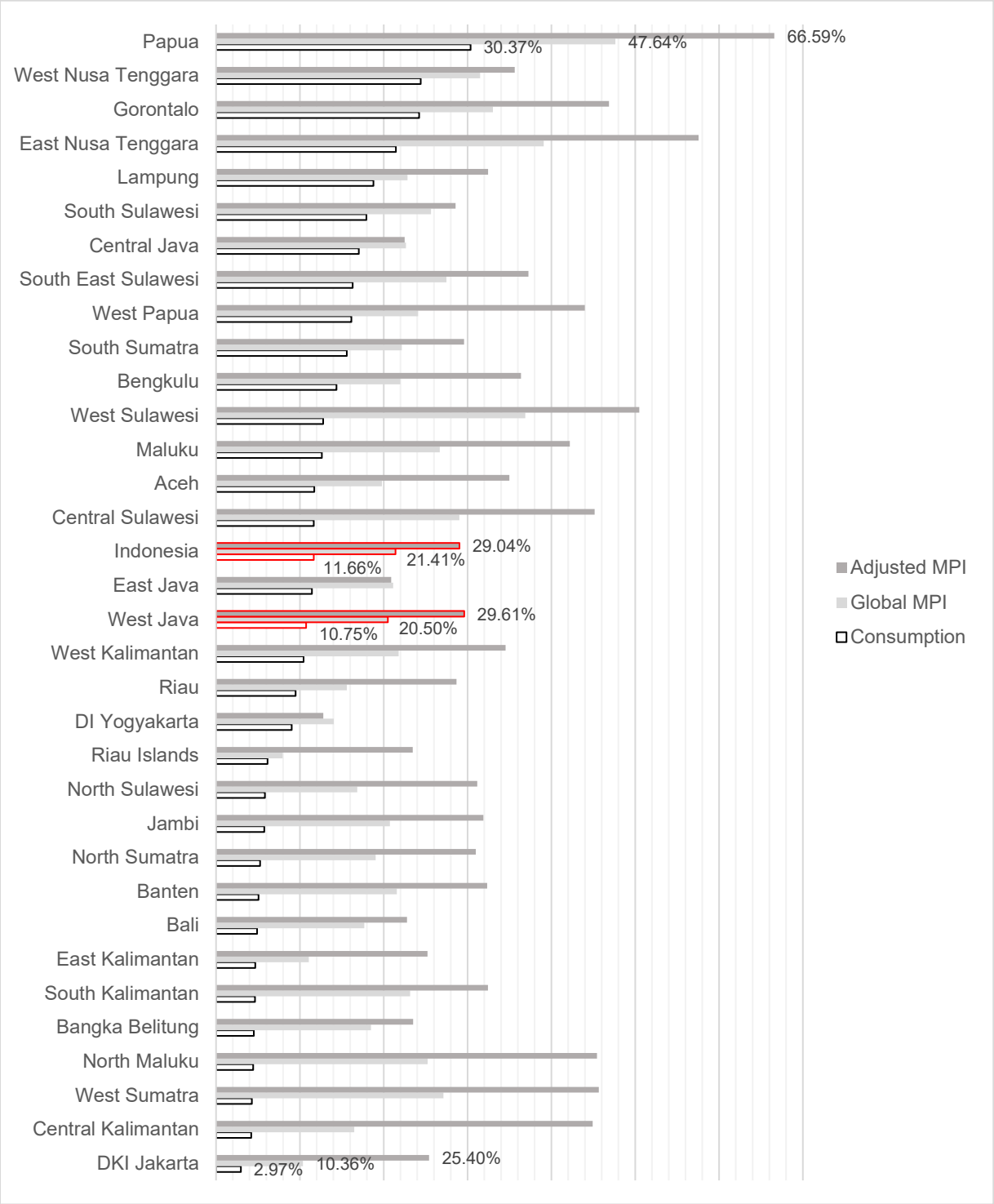
Although the computation of the Adjusted MPI within published literature also utilised the Susenas, I note two issues, which emerge when using this dataset. The first issue concerns the 'Under 5 Nutritional Intake' indicator. Limitations exist with regard to how micronutrient data is collected within the Susenas. Available data is in the form of 'total' household micro-nutritional intake; thus, in order to determine 'under five' nutritional intake, the only solution is to calculate household 'per capita' nutritional intake and apply these levels to those aged zero to five within the household. Biases, particularly those relating to household size, may thus, occur when using this information to measure deprivation in this indicator. Nevertheless, in order to follow the steps taken within the calculation of the Adjusted MPI, these biases were overlooked within this chapter. The second issue concerns missing values within the 'Drinking Water' indicator. Similar to that encountered within the computation of the Global MPI, a number of households did not report information on 'distance of their water source to waste disposal facilities'. Thus, these households are excluded from the final analysis. Despite this, a total retained sample of around 90 percent for all levels of analysis, i.e., national, provincial and district, continued to be available for the computation of the Adjusted MPI.

2.3 Results

2.3.1 Size of the Poor

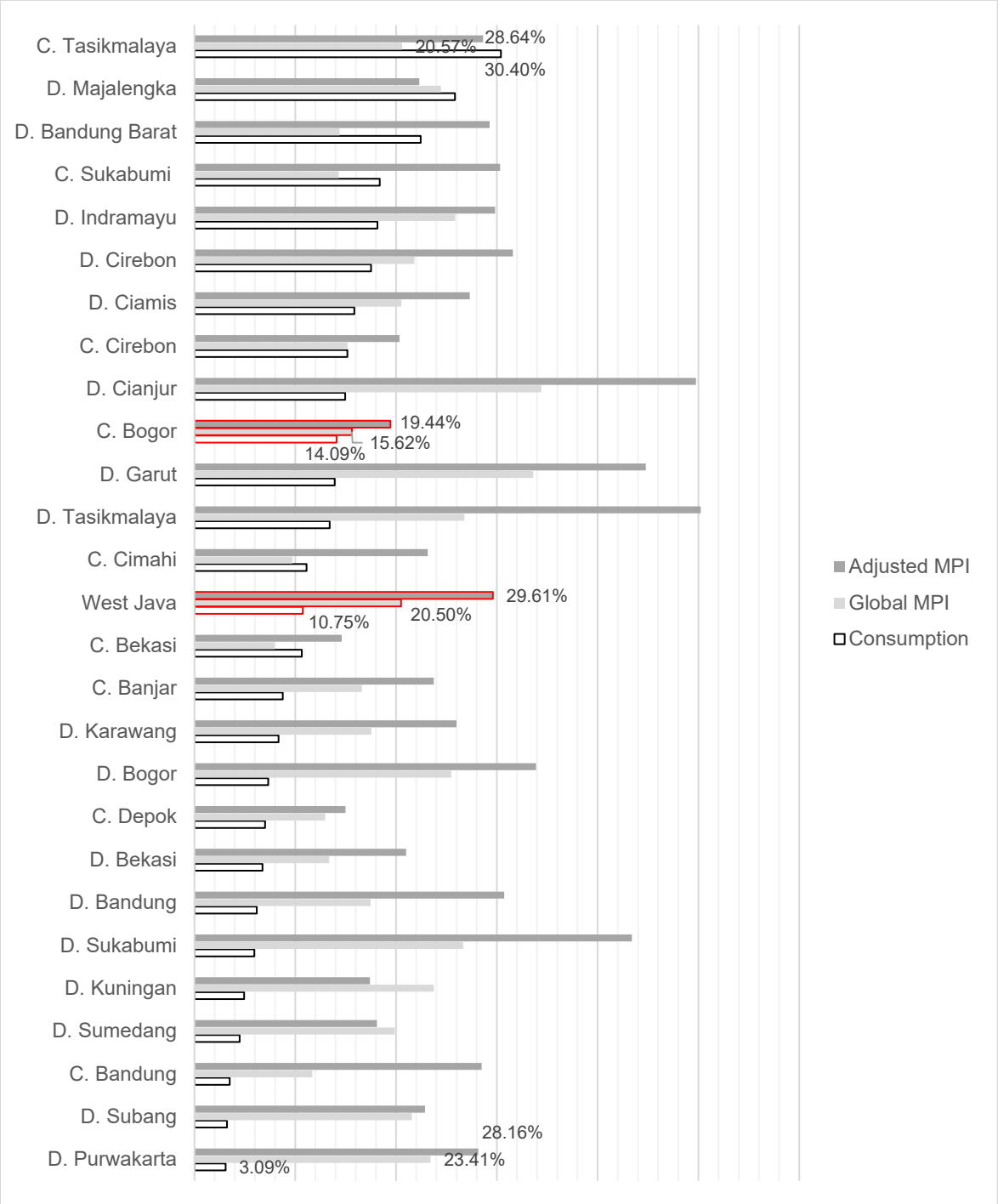
This sub-section aims to address the first research question within this chapter, i.e. whether the three measures identify a different extent of poverty. In answering this question using data from the three levels of analysis, district, provincial and national, this sub-section also seeks to attend to the fourth research question, i.e. whether conclusions on the extent of poverty apply regardless of the geographical level of analysis examined. Figures 2.1-2.2 illustrate results for these questions. The discussion below provides a short overview of these findings, before further analysis is given in Chapter 7, where these three existing measures are compared with the 'newly' developed measure based on Delphi results, proposed within this thesis.

Figure 2. 1 Proportion of poor households for Indonesia and its provinces according to the Consumption measure, Global and Adjusted MPIs [(a)] [(b)]



Note: [(a)] Population weights available in the Susenas 2013 are used to compute these percentages;
[(b)] Data is sorted from highest to lowest level of consumption poverty
Source: Author's computations (using Susenas 2013)

Figure 2. 2 Proportion of poor households for West Java and its districts according to the Consumption measure, Global and Adjusted MPIs [(a) (b) (c)]



Note: [(a)] Population weights available within the Susenas 2013 are used to compute these percentages;
[(b)] Data is sorted from highest to lowest level of consumption poverty;
[(c)] D: District (predominantly rural), C: City (predominantly urban)
Source: Author's computations (using Susenas 2013)

Figure 2.1 describes the percentage of poor households evident when the three measures were used to measure poverty within Indonesia and its provinces (percentages for Indonesia and West Java are highlighted in red). Figure 2.2 then displays the extent of poverty for West Java and its districts, including Bogor City (highlighted in red). Results from both these figures suggest four main points.

Firstly, when examining Figure 2.1, as was touched upon within Chapter 1, significant disparity exists across Indonesia's provinces with regard to the extent of poverty; this is true, regardless of the poverty measure utilised. Indonesia's eastern-most provinces, particularly the provinces of Papua, West and East Nusa Tenggara, West Papua and Maluku, are shown to exhibit relatively higher levels of poverty, according to all three measures. Whereas within Indonesia's western-most provinces, i.e. Riau, North Sumatra, West Sumatra, DKI Jakarta, DI Yogyakarta, poverty seems to be less prevalent.[1] This eastern and western divide is well recognised within literature describing the extent of Consumption poverty within Indonesia (Kanagaratnam, 2017; Miranti, 2010; Pradhan, Suryahadi, Sumarto, & Pritchett, 2000; A. Suryahadi, Sumarto, & Pritchett, 2003). It is interesting to note, that this 'divide' continues to exist when the multidimensional measures are used to assess poverty across Indonesia's provinces. When looking at Figure 2.2, disparity between districts within West Java also exists; with Tasikmalaya City exhibiting the highest number of *Consumption* poor (at 30 percent), Purwakarta District the lowest (at 3 percent) and Bogor City somewhat in the middle of the distribution (at 14 percent poor households).

Secondly, when comparing the Consumption vs. multidimensional measures, referring first to Figure 2.1, for all of Indonesia's provinces, including West Java, and for Indonesia overall, Consumption poverty is shown to underestimate the extent of poverty when compared to the two multidimensional measures. This result strengthens Amartya Sen (1979a)'s well established theoretical argument of the existence of a possible 'missing link' between monetary-based measures of poverty and those aiming to reflect more 'direct' indicators of deprivation. This finding also confirms existing empirical findings summarised within Chapter 1 on the lack of overlap between uni-dimensional Consumption, which aims to proxy deprivation in income, versus multidimensional measures aiming to reflect deprivation in other aspects of life (Laderchi

[1] Within Figure 2.1, poverty levels are ordered from the province with the highest level of Consumption poverty, i.e. Papua, Indonesia's eastern-most province, to the province with the lowest, i.e. DKI Jakarta, the capital of Indonesia. Compared to these two extremes, as noted within Chapter 1, West Java's level of Consumption poverty, seems to be located in the middle of this distribution.

et al., 2003).[1] When examining district level data within Figure 2.2, for almost all districts, the Consumption measure is shown to identify less households as poor when compared to the multidimensional measures. This is, indeed, the case for Bogor City (highlighted in red). For some districts, i.e. Tasikmalaya City and Majalengka District, however, Consumption is shown to identify higher numbers of poor households compared to both the Global and Adjusted MPIs.[2] This result, in turn, highlights Ravallion's argument on the continuing relevance of Consumption as an indicator of poverty, which should be considered alongside other multidimensional aspects of deprivation, to enable assessment of households' responsiveness to economic shocks (Martin Ravallion, 2011a, p. 8).[3]

Thirdly, when comparing the two multidimensional measures, first within Figure 2.1, for almost all provinces, the Adjusted MPI, which is based on nationally specific indicators, identifies more poverty, when compared to the Global MPI.[4] This indicates that the deprivation criteria set within the Global MPI, may generally have been met within most of Indonesia's provinces, however, this may not be the case for the more 'country specific' issues, which the Adjusted MPI aims to capture.[5] When examining Figure 2.2, a similar situation is found, i.e. for almost all districts within West Java, the Adjusted MPI is found to identify a larger extent of poverty compared to the Global MPI. This is indeed the case for Bogor City. Existing literature comparing the use of global indices vs. locally developed measures, particularly in the context of the developing world, is limited. Nevertheless, coming back to the conceptual framework described within Chapter 1, this result points to the importance of Townsend (1979a)'s idea of 'relative poverty', alongside Amartya Sen (1985)'s acknowledgement of the need to recognise that 'needs may vary from society to society' (p. 669). Thus, to address the fourth research question, localised measures may truly be needed.

---

[1] The proportion of households which are poor in all three measures in Indonesia is 2.06 percent, in West Java it is 1.77 percent and in Bogor City it is 0.63 percent.

[2] Djamaluddin (2014), within an analysis of Susenas 2010 data, also found relatively high levels of Consumption poor households within these two districts. This author found that a particularly significant determinant of poverty within these two districts was household's ownership of assets and land tenure. Bhargava and Ravallion (1993), Martin Ravallion and Jalan (1998) and Jalan and Ravallion (2000), among others, talk of the correlation between 'consumption' and measures of household 'wealth and assets'. This may offer some explanation as to why Consumption poverty is relatively high within these districts.

[3] As will be described within Chapters 3, 4, 5 and 6, Delphi participants also argued for the importance of 'Household Consumption Level' as an indicator of poverty.

[4] An exception to this is present within DI Yogyakarta, a province located in Central Java, where the Global MPI reports higher poverty when compared to the Adjusted MPI. Analysis as to why this is the case is beyond the scope of this thesis. Further research is thus, needed on this issue.

[5] Section 7.3.2 of Chapter 7 further explores this issue.

Finally, despite the differences highlighted above, it seems that in general, provinces that are considered as the 'poorest', remain similar, regardless of the poverty measure. This is less so when district poverty levels are examined. To illustrate, in Figure 2.2, Majalengka City has the highest number of Consumption poor, however, does not have the highest level if the Global or Adjusted MPI are used. Differences in how, in particular, *districts* are ranked, may lead to significant policy implications. As noted within Chapter 1, Indonesia's government is decentralised at the district level and disbursements of funds from the central to district governments depend very much on the relative positions of these districts with regard to their poverty levels (Brodjonegoro, 2002, 2003, 2004).[1]

2.3.2 Household Ranks

This sub-section seeks to answer the second research question within this chapter, i.e whether household ranks differ when the three measures are utilised. Within this section, household ranks according to the Consumption measure are determined by their levels of consumption-per-capita, whilst for the multidimensional measures, ranks are determined by a household's total weighted deprivation scores across the multiple indicators included within the measures. Table 2.3 presents these results.

Table 2. 3 Spearman Rank Correlation Coefficients between Household Rank as Determined by the Three Measures [(a) (b)(c)]

	Bogor City			West Java			Indonesia		
	Consumption	G MPI	A MPI	Consumption	G MPI	A MPI	Consumption	G MPI	A MPI
Consumption	1			1			1		
Global MPI	0.28	1		0.32	1		0.35	1	
Adjusted MPI	0.42	0.38	1	0.42	0.4	1	0.43	0.5	1

Note: [(a)] All coefficients are significant at 1% level; [(b)] Ranks for the Consumption measure were assigned at descending order ; [(c)] GMPI: Global MPI, AMPI: Adjusted MPI

Source: Author's computations

Relatively low levels of correlation (an average of 0.38 across the three geographical levels of analysis) can be found between the ranks assigned by the three measures. As discussed in Chapter 1, differences in how households are ranked, may lead to opposing conclusions with regard to how social assistance packages are allocated; this is important as targeting policies

[1] Sub-section 7.3.1 of Chapter 7 explores this further.

play a key role within poverty alleviation efforts within Indonesia (Alatas, Banerjee, Hanna, Olken, & Tobias, 2012; A. Booth, 2003).

To also respond to research question four, results within Table 2.3 show that correlation between Global MPI and Adjusted MPI household ranks are only at modest levels, however, is relatively higher when data from Indonesia is considered, compared to when data from Bogor City and West Java are utilised. Following results within sub-section 2.3.1, this suggests that the Global MPI's indicators may not capture specific poverty issues that are important for Indonesia; this is even more the case, when concerns within Indonesia's provinces and districts are considered.[1]

2.3.3 Composition of the poor

This section aims to answer this chapter's third and fourth research questions: do the three different measures identify different types of households as poor? Do conclusions on the composition of the poor identified by the three measures, hold, when the different geographical levels of analysis are considered? To address these questions, multivariate logistic regressions are conducted to analyse the effect of different socio-economic variables on whether a household is poor according to the different measures, across the three geographical levels. This regression framework aims to construct what M. Ravallion and Bidani (1994) dub as 'poverty profiles', for each of the measures examined. Before discussing results, a brief overview of the socio-economic characteristics included within this analysis is presented below.

2.3.3.1 Socio-economic Characteristics Examined

Table 2.4 details the socio-economic characteristics examined within this sub-section. It also lists the hypothesised relationship between each of these variables with a household's probability of experiencing poverty. These hypotheses are however, based for the most part, on previous research on Consumption poverty, as limited empirical evidence exists with regard to the poverty profile of those experiencing Global and Adjusted MPI poverty, particularly in Indonesia. Literature on the characteristics of the urban poor in Indonesia, as the case study area within this thesis, i.e. Bogor City, is predominantly urban, are also drawn upon.

For the first characteristic, 'gender', having a female household-head is hypothesised to increase the chances of experiencing poverty (Central Bureau of Statistics Indonesia (CBS), 2014; Millar & Glendinning, 2009). This hypothesis is confirmed, among others, by Pritchett,

[1] Sub-section 7.3.2 of Chapter 7 explores this further.

Suryahadi, and Sumarto (2000), who found that 'while female headed households have the same level of expenditures, they are more vulnerable to shocks and hence more at risk of poverty' (p. 21). Qualitative research on Indonesia, particularly that conducted within the 'Voices of the Poor' reports, further strengthened this finding, as they concluded that 'women in all communities across Indonesia's provinces, stress the economic and social vulnerabilities of female-headed households' (Mukherjee, 1999; Narayan & Petesch, 2002).

Table 2. 4 Socio-economic characteristics of household and hypothesised relationship with household's probability of being Consumption poor

Variable	Hypothesised Relationship with Consumption Poverty
Age of household head	Inverted U
Gender of household head (Female)	+
Marital status of household head (Married)	-
Level of education of household head	-
Employment sector of household head (employment within other sectors, other than the Extractive/Agricultural Sector)	-
Household composition	
Number of adults in household	-
Number of children in household	+

Source: Author's summary

For the second characteristic, the 'age' of the household-head, an 'inverted U relationship' is hypothesised, i.e. poverty decreases as a household head reaches middle age and increases as he or she approaches old age (Alisjahbana & Manning, 2006; John C. Anyanwu, 2014; Datt & Jolliffe, 2005; Gang, Sen, & Yun, 2002). Household-head's 'marital status', i.e. being 'married' as opposed to being 'single, divorced or widowed', is then hypothesised to decrease the chances of the household experiencing poverty. This follows a number of previous empirical works on Indonesia which found that households with 'married heads' tend to have more workers within them, this in turn increases the household's potential of generating greater income, thus, reducing its chances of experiencing Consumption poverty (Alisjahbana & Manning, 2006; Dartanto & Nurkholis, 2013).

For the next characteristic, 'education level of household head', the probability of experiencing poverty is hypothesised to decrease as the household-head moves from being lowly educated to completing higher levels of education (primary, secondary, high school and higher) (Central Bureau of Statistics Indonesia (CBS), 2014; Millar & Glendinning, 2009; Pritchett et al., 2000; Asep Suryahadi & Sumarto, 2010). For the next characteristic, following findings within Lanjouw and Murgai (2009), among others, as the household head's 'employment sector' moves from being employed within the agricultural sector, to that within the industrial, manufacturing, services and the government sectors, the likelihood of experiencing poverty is hypothesised to decrease.[1]

For the final characteristic, if only household size, rather than measures indicating household composition, are considered, following Lanjouw and Ravallion (1995) and Deaton and Paxson (1998), a positive relationship between size and poverty is often found, i.e. 'people living in larger households are typically poorer' and that 'household scale economies, which are assumed to make larger households better off, do not occur'. Using 'total household size' thus, renders us unable to uncover the different impact the 'number of adults' and 'number of children', may have on poverty. As argued for within Dartanto and Nurkholis (2013), the 'number of adults' could indicate the number of 'working individuals' within a household thus, increases in this variable may lead to decreases in the probability of being poor. As for the impact of the 'number of children' within the household, Jan Priebe (2010) argues that the probability of being poor increases, as the number of children within the household rises. This is due to the possibilities of the household incurring additional child costs, which in turn, act to increase the household's vulnerability toward poverty.[2]

2.3.3.2 Results

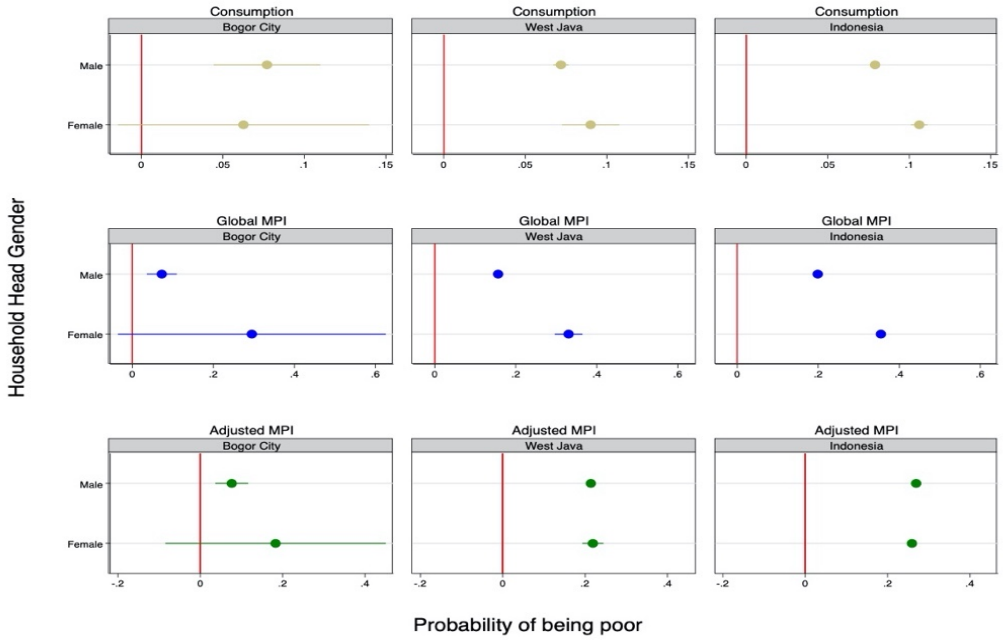
Figures 2.3-2.7 below graphically illustrate the predicted probabilities for the set of multivariate logistic regressions described at the start of this section.[3]

[1] Although Bogor City is predominantly urban, some households are still employed within the agricultural and extractive sectors (Central Bureau of Statistics Bogor City, 2018).

[2] In addition to the above, for regressions using data from Indonesia and West Java, a variable controlling for urban/rural effects is included. Following Asep Suryahadi, Suryadarma, and Sumarto (2006) and Asep Suryahadi and Sumarto (2010) findings on determinants of Consumption poverty, living within a rural area is hypothesised to increase the probability of being poor.

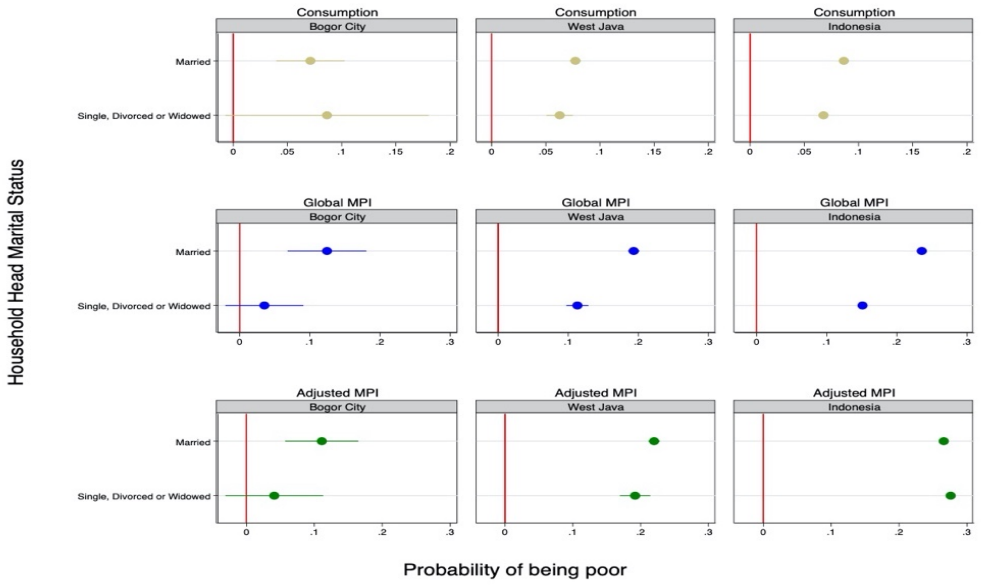
[3] These predicted probabilities were calculated for each socio-economic determinant, whilst keeping the other determinants at their mean values. The log odds for these multivariate logistic regressions are presented within Table 7.9 in the Appendix to Chapter 7 of this thesis. Conscious of space limitations, the

Figure 2. 3 Predicted probabilities of household poverty status by gender of household head (with 95 percent confidence intervals)



Source: Author's computations

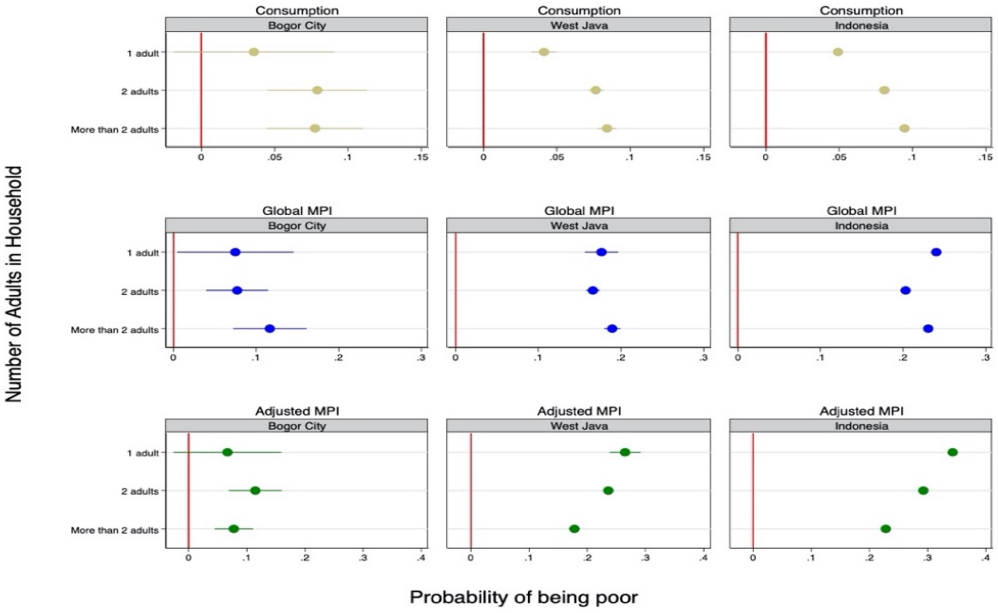
Figure 2. 4 Predicted probabilities of household poverty status by marital status of household head (with 95 percent confidence intervals)



Source: Author's computations

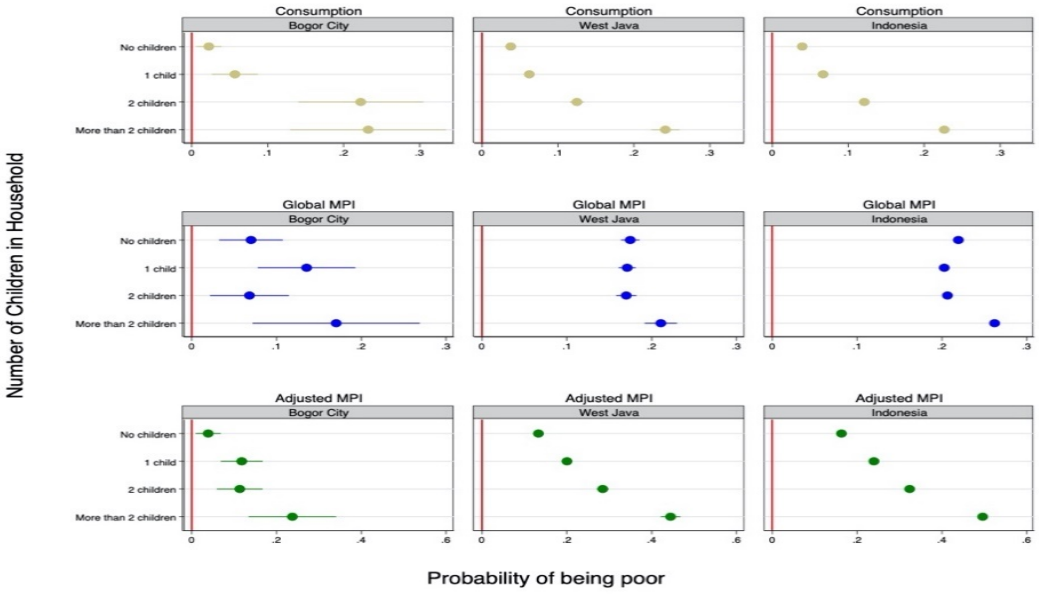
impact of household heads' 'work sector' is not discussed further within this sub-section as the effects of this variable were found to be relatively insignificant for Bogor City and West Java.

Figure 2. 5 Predicted probabilities of household poverty status by the number of adults in the household (with 95 percent confidence intervals)



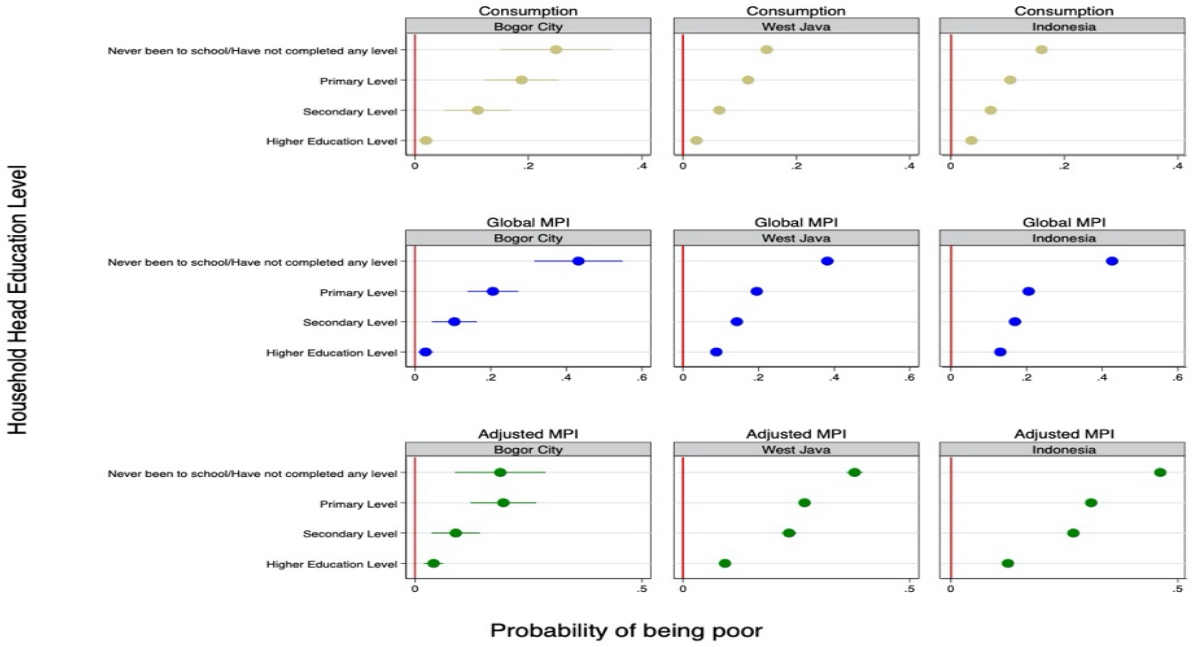
Source: Author's computations

Figure 2. 6 Predicted probabilities of household poverty status by the number of children in the household (with 95 percent confidence intervals)



Source: Author's computations

Figure 2. 7 Predicted probabilities of household poverty status by household head's education level (with 95 percent confidence intervals)



Source: Author's computations

The results below aim to briefly focus on areas where significant differences were found, whilst also noting important similarities between the measures, before a more thorough review of results is presented within Chapter 7.

I start with characteristics where the measures particularly differ. Figure 2.3 illustrates the relationship between household head's 'gender' and the probability of the said household experiencing poverty according to the three measures. Firstly, across the different geographical levels, only subtle differences are shown to emerge, however, within Bogor City, the confidence intervals are wider, suggesting the statistical insignificance of this characteristic. Relatively more pronounced differences are, however, evident when examining across the measures of poverty; these are with regard to the pattern and size of the effect of this household characteristic on the probability of a household experiencing poverty. This in mind, Figure 2.3 suggests that a household's probability of being multidimensionally poor, particularly for the Global MPI, increases for households with 'female' heads. This is in line with the hypothesis noted within Table 2.4. Furthermore, although few pieces of literature exist which examine characteristics of

the multidimensionally poor, Rogan (2016), within an analysis of multidimensional poverty in South Africa also found similar conclusions, i.e. that 'women and those living within female headed households suffer the most poverty' (p. 1003). On reasons behind why this is the case, Rogan (2016) suggests that 'disadvantages in the labour market for women, greater responsibility for the care of children, in particular reflected within the child mortality indicator', may have led to this 'gender gap'. As touched upon within sub-section 2.3.3.1, qualitative analysis within Mukherjee (1999), Mukherjee (2006) and Narayan and Petesch (2002), found that women face these similar challenges in Indonesia. Interestingly, such a relationship is not the case for Consumption poverty when data from Bogor City is examined, where 'male' headed households are shown to be more likely to experience poverty (although as noted above, the confidence intervals for Bogor City suggest the insignificance of this variable). Nevertheless, at the provincial and national levels, in line with the hypothesis in Table 2.4, Consumption is also shown to identify 'female' headed households as more likely to be poor.

Figure 2.3 also illustrates differences with regard to the magnitude of the effect of 'gender'. Particularly when comparing the multidimensional measures, 'gender' seems to influence the probability of being Adjusted MPI poor less when compared to how this characteristic affects the Global MPI poor. This is even more so when analysis moves to include more information available at the provincial and national levels, where the distance between the predicted probabilities of being Adjusted MPI poor, become closer, regardless of whether the household is headed by a 'male' or a 'female'. In fact, at the national level, although only a subtle difference is evident, the Adjusted MPI tends to identify a higher probability of being poor within 'male' headed households.[1] The types of indicators included within the Adjusted MPI compared to those within the Global MPI may have influenced this result. To briefly illustrate, the Adjusted MPI's exclusion of the Global MPI's 'Child Mortality' indicator, measuring whether women within the household have ever experienced the death of a child, may have led this measure to be less sensitive toward deprivations concerning women within the household, particularly when national level data is considered.[2]

[1] Espinoza-Delgado and Klasen (2018) found similar evidence when examining the influence of 'gender' on multidimensional poverty in Nicaragua. These authors found that 'males' and 'females' are almost equally likely to be multidimensionally poor; this is even though, the multidimensional index they used included indicators specifically aimed at capturing the vulnerability of women within the household.
[2] This is the case, even though, the Adjusted MPI, includes an indicator assessing maternal health, i.e. 'Birth Attendant'. For Indonesia as a whole, the level of 'Child Mortality' is higher (at 18 percent), when compared to levels within Bogor City (at 15 percent). Deprivation within 'Birth Attendant' (at 6 percent in Indonesia) is also significantly lower than deprivation in 'Child Mortality'.

Figure 2.4 then examines the relationship between household head's 'marital status' and the probability of being poor. As illustrated within this figure, this relationship varies across the different poverty measures. In Bogor City, the relationship between 'marital status' and a household probability of being Consumption poor, seems to be in line with the hypothesis within Table 2.4, i.e. households with 'married' heads tend to experience less poverty. The contrary is shown to be true for the multidimensional measures; households headed by 'married' individuals tend to experience more multidimensional poverty in Bogor City. However, at the provincial level, all three measures are shown to identify households with 'married' heads as more likely to be poor. On this issue, for the context of Indonesia, as noted in sub-section 2.3.3.1, existing research has for the most part, found that households with 'married' couples tended to experience less poverty. These pieces of research however, utilised earlier versions of the Susenas, i.e. from years 2002 to 2005. More recently, changes as to how 'marital status' affects poverty, thus, may have taken place. Although limited literature which utilises more recent data exists, this is confirmed within Rustiadi and Nasution (2017) who using 2012 Susenas data found that similar to results within Figure 2.4, households with 'married heads', tended to experience more Consumption poverty.[1] There is limited information as to why this may be the case for Indonesia. John C Anyanwu (2005) and John C. Anyanwu (2014), however, argue that within the context of African economies, this 'reduction' in the probability of experiencing poverty within single headed households, may be due to an improvement in the economic status of divorced or separated men, whilst, for women, some studies have shown that moving back to live with parents, alongside possible welfare and child support transfers, may reduce the probability of experiencing poverty.

A slight difference in conclusion is however, found when examining the effects of 'marital status' on the Adjusted MPI at the national level, where 'married' households, are shown to experience less poverty. The types of indicators included within the Adjusted MPI may again, play a role behind this result. Considering this, an analysis of the proportion of 'married' or 'single, divorced or widowed' headed households, which are poor and deprived in each of the Adjusted MPI's indicators, could uncover which indicator may be 'driving' high levels of deprivation within specific types of households. Although beyond the scope of this chapter, preliminary analysis on this issue finds that a significantly large proportion of Adjusted MPI poor households, which

[1] In addition to this, Chen, Leu, and Wang (2019), using data from Taiwan, also found that 'married' individuals tended to be more susceptible to experiencing *multidimensional* poverty. These authors used a version of the Global MPI to measure poverty within their analysis, however added an indicator assessing deprivation in 'Employment' within their measure.

have ‘married’ heads, are deprived in the indicators of ‘Education Continuity’ and ‘Under 5 Nutritional Intake’. As will be discussed further within Chapter 7, deprivation rates within both these indicators were found to be significantly high (71 percent of households were found to be deprived in ‘Education Continuity’ in Indonesia and 36 percent in ‘Under 5 Nutrition’).

Figure 2.5 shows that the ‘number of adults’ within the household increases the chances of a household being Consumption and Global MPI poor. This is particularly the case when data from West Java and Indonesia are examined. Lanjouw and Ravallion (1995) and Deaton and Paxson (1998)’s argument that a positive relationship between ‘household size’ and poverty, seems to exist, when the Consumption and Global MPI are used to measure poverty.[1] Coming back to the hypotheses within Table 2.4, a similar result to that found within Dartanto and Nurkholis (2013), is thus, only evident for the Adjusted MPI, i.e. increases in the ‘number of adults’ in the household, leads to decreases in the said household’s probability of experiencing poverty. In addition to Dartanto and Nurkholis (2013)’s hypothesis that ‘more adults’ may reflect the additional household members who are able to contribute toward increases in income and resources within the household, looking at the types of indicators within the Adjusted MPI, having older individuals within the household (and less children), may play a role in reducing the likeliness of experiencing deprivation within the measure’s child-specific indicators, particularly those of ‘Access to Pre-school Education’, ‘Under 5 Nutritional Intake’ and ‘Education Continuity’.[2] As noted above, deprivation within these indicators are relatively high. [3]

Figures 2.6-2.7 illustrates characteristics where the measures are relatively similar. Firstly, Figure 2.6 shows that ‘the number of children’ within the household, increases the chances of being poor for all three measures. As noted within sub-section 2.3.3.1, this may indicate the additional ‘costs’ that children bring into the household. Secondly, although differences in the pattern of the effect of this characteristic exist, increases in the education level of the household head are shown to decrease the chances of households experiencing poverty for all measures,

[1] Albeit not attempting to separate the effects of the ‘number of adults’ vs. the ‘number of children’ in the household, Chen et al. (2019) also found similar results in Taiwan, i.e. that ‘household size’ was significantly positively associated with multidimensional poverty.

[2] As noted within Table 2.2, the ‘applicable population’ for ‘Access to Pre-School Education’ are those aged three to six, for ‘Under 5 Nutritional Intake’ are those who are under five years old, and for ‘Education Continuity’ are those aged eight to twenty.

[3] Another area of difference evident within Table 7.9 of the Appendix to Chapter 7 detailing the log odds resulting from the regressions within this sub-section is that for all measures, except the Global MPI, an inverted-U relationship is, for the most part, evident when the impact of ‘age’ of household head on the probability of being poor, is examined. In other words, for the Consumption and Adjusted MPI, poverty decreases as a household head reaches a certain age and increases as he or she grows older. This is not the case for the Global MPI.

regardless of geographical level. For the Consumption measure, a steady decrease in the probability of being poor is shown as the household head moves from having ‘no qualifications’ to being educated at a ‘high school level or higher’. In Bogor City, the impact of having ‘no qualifications’ on increasing the probability of being poor, is shown to be higher for the Global MPI, as opposed to the other measures. In Bogor City, for the Adjusted MPI, the impact of having ‘no qualifications’ vs. having a ‘primary level certificate’, seems to be similar; thus, for this measure, higher levels of education, i.e. at ‘secondary’ and ‘high school or higher’, play a bigger role in lowering the probability of being poor. Referring back to the types of indicators included in the Adjusted MPI, this may particularly be due to the stringency of the cutoff used to define deprivation within the measure’s ‘Education Continuity’ indicator, which specifies the need for children to be attending school ‘at least until high school level’.

In sum, a number of differences were found between the types of households identified as poor by the three measures. These differences were particularly evident when comparing the consumption versus the multidimensional measures; more so when the Adjusted MPI was examined. The Adjusted MPI was, in particular, also shown to exhibit differences in conclusions (although for the most part, subtle) with regard to which household characteristic contributes most to the experience of poverty, when the different geographical levels of analysis were considered. These findings firstly point toward the importance of considering both uni- and multidimensional measures when monitoring poverty; and secondly, indicate the need to consider context specific poverty measures as different characteristics were shown to influence the different multidimensional measures.

2.4 Conclusion

As M. Ravallion and Bidani (1994) state, ‘when practices in empirical work have a bearing on policy choices, they deserve especially close scrutiny’ (p. 75). This in mind, this chapter presented a review of three existing poverty measures, i.e. the Consumption, Global and Adjusted MPIs, to firstly, highlight areas within which further improvement may be of benefit and secondly, to illustrate possible differences between how these measures identify the size and composition of the poor, alongside, how households are ranked according to their deprivation level as determined by these three past measures. Analysis within this chapter was conducted using Susenas 2013 data at three geographical levels, national (Indonesia), provincial (West Java) and district (Bogor City) levels, in order to examine whether conclusions with regard to how the measures identify the poor, differ when information from these different levels is analysed. Four main conclusions can be drawn from the analysis within this chapter.

Firstly, this chapter found that using the three different measures led to significant differences with regard to the extent of poverty, with the Consumption measure identifying less poverty, compared to that identified by the multidimensional measures. The Adjusted MPI in particular, identified nearly three times more poverty in Indonesia, compared to the Consumption measure. This confirms existing evidence within literature, summarised within Chapter 1 of this thesis, on the 'miss-match' between these two forms of measurement. In particular, Stiglitz et al. (2010b) note that 'expenditures' required to maintain consumption levels could be viewed as 'a sort of intermediate input'. Thus, to truly monitor poverty, consideration of multidimensional measures, which include indicators aiming to assess deprivation in the myriad factors affecting well-being, is also needed.

Secondly, when examining differences in household rank as identified by the three measures, only moderate correlation (an average of 0.38) was evident between these measures. This step of the exercise was important, for as noted within Chapter 1, targeting policies play a significant role in Indonesia's efforts to mitigate poverty. These differences in ranks, indicate that differing conclusions will arise, with regard to which households, deserve to receive crucial social assistance. Again, the need to consider both forms of measurement before sound policy can be made, is highlighted by this result.

Third, when examining the types of households each measure identified as most likely to experience poverty, differences also emerged. This was found to be particularly so, when comparing the Consumption and the multidimensional measures using data from Bogor City. To illustrate, using data from Bogor City, the Consumption measure identified households which were headed by 'males', who were 'single, divorced, or widowed' and living in larger households, as more likely to experience poverty; whereas the multidimensional measures, identified households headed by 'females', who were 'married' and living in larger households, as having a higher probability of being poor. As analysis moved to include more information at the level of West Java, these differences, however, became less pronounced, i.e. both Consumption and multidimensional measures tended to identify more 'female' and 'married' headed households as poor. Looking at the multidimensional measures, at national level, 'gender' and 'marital status' were shown to influence less the probability of being Adjusted MPI poor, compared to the effects of these characteristics on being Global MPI poor. Despite these differences, two household characteristics, i.e. 'number of children' in the household and the 'education level' of the household head, were found to influence the experience of poverty in similar ways for all measures. Although the size and pattern of the effect differed, increases in

the ‘number of children’ was shown to increase the probability of being poor for all measures, regardless of the geographical level of analysis. For ‘Education’, again, even though increases in ‘education level’ led to a different pattern of the effect, across the three measures, the likelihood of being poor decreased as household heads moved from having ‘no education’ to attaining ‘higher levels’ of education.

Finally, as touched upon above, differences were found, particularly with regard to the types of households identified as poor, when different geographical levels of analysis were utilised. Particularly for the multidimensional measures, this highlights the fact that dimensions, indicators and weights, included within an internationally relevant measure, such as the Global MPI, may not reflect the different policy priorities and contextual needs, of measures aiming to be nationally relevant. Indeed, Sabina Alkire et al. (2019) themselves argue for the relevance of ‘national MPIs’, to supplement the Global MPI. As argued for within Chapter 1, however, within a nation as diverse as Indonesia, where poverty alleviation policies are decentralised at the district level, a thorough exploration of the implications of taking into account ‘district level’ contextual issues, becomes important. This in turn, strengthens the argument made within Chapter 1 to utilise a district level case study within this thesis.

In sum, I quote Townsend (1979a)’s classical statement: ‘the fact that how poverty is defined and measured has implications for policy, should be recognised at the outset, particularly since, social sciences have increasing influence upon decision-makers, both in providing information and implicitly or explicitly legitimating action’ (p. 31). The findings within this chapter legitimised the need for multidimensional measures, not as substitutes, but as complements to that of Consumption. Given this apparent need, scrutiny as to how these multi-faceted measures are constructed, particularly how dimensions, indicators and weights are selected within them, is needed. This issue becomes even more crucial, as shown within this chapter; although both the Global and Adjusted MPIs included similar dimensions and utilised the same weighting scheme and differed for the most part, only on the types of indicators they included, differences, particularly with regard to the extent of poverty identified by these two measures, were found. Considering this, a reliable and robust methodological step, aiming to reduce the degree of arbitrariness with regard to how components within multidimensional measures are selected, is needed. The Delphi suggested within the following Chapter aims to be such an approach.

Chapter 3: Methodological Steps of Incorporating the Delphi Method into Poverty Measurement

Consensus has not been reached on which dimensions [of poverty] matter, nor even on how to decide what matters'

(Grusky et al., 2006, p. 12)

3.1 Introduction

A poverty measure provides 'a basis from which to evaluate the effectiveness of welfare states' and acts as 'a powerful instrument' which may be utilised to 'focus the attention of policy makers on living conditions of the poor' (Mack & Lansley, 1985, p. 9; Martin Ravallion, 2001, p. 2). Chapter 2 examined the use of three existing poverty measures, i.e. (1) the Consumption measure, (2) the Global MPI, and (3) Adjusted MPI. When comparing these measures, it was found that each identified both a different size and composition of the poor. Reliance on any one of these measures as a single yardstick, 'the litmus test' of policy effectiveness, may thus be misleading, for each measure not only identified a different extent of poverty, but also categorised different types of households as poor. Given this mismatch found, one may posit that the fundamental issue with relying solely on these past measures is 'not simply one of measurement', it is, however, whether these measures, 'in fact, tell us what we want to know when we set out to measure poverty' (Nolan & Whelan, 1996, p. 61).

As was argued within the preceding chapters, there exists widespread agreement on the importance of acknowledging that poverty is a multifaceted phenomenon. This has led to important theoretical and empirical advancements, which aim to 'shift' the emphasis of theory and policy, from uni to multidimensional approaches to poverty measurement.[1] Despite these progressions, however, substantial debate still exists on the issue of how to 'select' the multiple facets of poverty to be included within such measures (Kakwani & Silber, 2007, p. xv). In particular, as discussed within Chapters 1 and 2, vehement critique remains on the arbitrary way 'weights' for these multidimensional aspects are chosen.

Considering the above, this current chapter, Chapter 3, suggests the use of the Delphi method as a robust and transparent technique of soliciting a reliable level of consensus with regard to the choice of dimensions, indicators and weights, to include within a measure of

[1] The Global MPI and the Alkire-Foster method utilised within its construction, discussed in Chapter 2, are examples of such significant advancements.

multidimensional poverty. Furthermore, the Delphi aims to address the need for what Amartya Sen (2004b) dubs 'comprehensive outcomes' as it is 'not enough' to only focus on 'cumulative outcomes', i.e. *which* dimensions and indicators matter, but rather, discussion on *how* they were chosen, in other words, the 'comprehensive process' by which they were selected, is also pivotal when aiming to construct sound empirical measures (emphasis added, Sen (2004b)). This chapter serves to illustrate the methodological steps taken to conduct the proposed Delphi study, whilst its results are the subject of Chapter 4.

The Delphi method, as defined by Harold A Linstone and Turoff (1975), is a research approach used to gain a level of agreement on crucial issues through a series of iterative questionnaires, where information and results are fed back to participants between each iterative step (p. 10-11). The Delphi technique has three main features: (1) anonymous response, (2) iteration and controlled feedback and (3) statistical group response (Norman Crolee Dalkey, Brown, & Cochran, 1969). Considering these three aspects and in order to construct a step-by-step guide to the Delphi applied within this thesis, this chapter aims to address three main questions. Referring to the first feature, the anonymity of participants, this chapter starts by addressing the question of: *Who will be included and how should participants within the Delphi be selected?* Given the second feature, i.e. the method's iterative nature, the second question seeks to outline: *How many iterative questionnaire-rounds should take place and how should results from previous rounds be 'fed' back to participants?* Finally, considering the third aspect: *How can 'group response' resulting from the Delphi's iterative process be measured and assessed in order to determine whether a level of 'consensus' is achieved with regard to the choice of dimensions, indicators and weights, to be included within poverty measurement?*

As described within Chapter 1, the Delphi in this thesis is conducted within one district-level case study area in Indonesia; i.e. Bogor City, West Java. The above questions are thus answered with reference to this case study area. Even though this is the case, the overall aim of this chapter and the following chapters within this thesis, is not to be on poverty in Bogor City, per se. It is rather, to illustrate how a novel method, i.e. the Delphi, may be incorporated within the construction of empirical measures of multidimensional poverty. Furthermore, the methodological steps outlined within this chapter aim to contribute to existing literature, as, despite being applied within numerous fields, to the best of the author's knowledge, the Delphi method has yet to be utilised within research specifically concerned with the selection of components of a multidimensional poverty measure. It is hoped that the steps outlined in this chapter may act as a guide for studies aiming to replicate this technique within different settings.

The rest of this chapter is organised as follows. Section 3.2 provides a literature review describing the Delphi's history, definition and epistemological stance. Section 3.3 summarises past research addressing the challenges and methodological considerations when conducting a Delphi study. Given these existing guidelines, Section 3.4 describes the Delphi protocol taken within this thesis and the steps utilised to address this chapter's three research questions. Section 3.5 ends this chapter with a list of potential limitations of the Delphi and how the methodological choices taken within this chapter seek to mitigate them.

3.2 Literature Review

3.2.1 Definition of the Delphi

Developed by the RAND Institute at the beginning of the Cold War in the 1950s, the Delphi method, was first implemented to 'forecast the impact of technology on warfare' (Custer, Scarcella, & Stewart, 1999; N. Dalkey & Helmer, 1963). Since it was first implemented, however, the use of this method within the social sciences has increasingly been recognised (de Villiers, de Villiers, & Kent, 2005; J. Jones & Hunter, 1995; Keeney, Hasson, & McKenna, 2011; Keeney, McKenna, & Hasson, 2010; Landeta, 2006; Lewis-Beck, Bryman, & Liao, 2003; Wang et al., 2003). Although this is the case, as stated within the Introduction of this chapter, previous literature on the use of the Delphi within poverty measurement has been limited. Existing studies, which I will draw upon within the subsequent sections, have however, applied this method within the evaluation of the 'quality of life' and many have utilised it to support evidence-based policy formation (Norman C Dalkey, 2002; Norman C Dalkey, Rourke, Lewis, & Snyder, 1972; Turoff, 1970, 1975).

The Delphi may broadly be defined as 'a method for structuring group communication so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem' (Harold A. Linstone & Turoff, 1977, p. 3). This 'structured communication' is achieved through allowing individual members of the group access to the contributions and knowledge of other group members. This 'access' is granted through the presentation of 'feedback' between an iterative round of questionnaires. Based on this 'feedback', individual participants are then given the ability to 'assess' the judgment of the group and the possibility to 'revise' their responses. This process is most commonly conducted anonymously, i.e. individuals are not required to state ownership of their particular responses and group feedback is presented in a way that is unspecified to particular individuals.

Norman C Dalkey and Rourke (1971) argue that the Delphi is often found to be a favourable method for 'issues where the best available information is the judgments of knowledgeable

individuals' (p. 1). Coming back to the purpose of this thesis, the exercise of 'defining and measuring poverty' comprises such issues, for ideally, measurement 'must take note of the prescriptions made by members of the community' (Amartya Sen, 1979b, p. 285).[1] Norman C Dalkey and Rourke (1971), Harold A Linstone and Turoff (1975) and Keeney et al. (2011), among others, state that in gathering these 'views and judgments', the Delphi serves the purposes of: (1) enabling the incorporation of participants with diverse areas of expertise, (2) facilitating the inclusion of a substantial number of experts, which often, may be too large for face to face discussions, and (3) given the ethical issues related to participant time and resources, acts as an efficient platform when regular meetings are not feasible. This chapter aims to illustrate how the Delphi may serve to fulfil these three purposes as a technique to support the construction of multidimensional measures of poverty.

3.2.3 Epistemological Stance

Chapter 1 of this thesis notes the importance of 'participatory methods' as a way of understanding and uncovering the diversity of deprivations, which may afflict a household experiencing poverty (Chambers, 2013). This in mind, before further discussing the strengths of the Delphi compared to other participatory and consensus building methods, clarification first needs to be made on its ontological and epistemological stance. Can the Delphi indeed, be considered qualitatively participatory or does it rather, belong to the realm of quantitative science? Keeney et al. (2010) argue that the answer to this question is dependent on how the method is applied in practice, as a Delphi study may incorporate both qualitative and quantitative approaches (p. 18). This reasoning is further explored below.

Since part of the data collected within the Delphi is 'quantitative', Day and Bobeva (2005) posit that the Delphi follows a positivist paradigm, i.e. it assumes that 'the researcher is objective' as he or she is supported by the use of 'statistical measures to assess consensus'. Hanafin (2004) and Turoff (1975), among a number of other authors, however position the Delphi technique within the interpretative paradigm, particularly 'social constructivism', viewing it as 'subjective and qualitative in nature as it is a process of collecting individual feedback about group opinion with opportunities for participants to change their position based on this feedback'. In this case, the Delphi is thought to 'validate truth through human experience, as it is directly linked to the context-dependent nature of its participants' knowledge' (Crotty, 1998; Keeney et al., 2010, p. 19; Strauss & Zeigler, 1975a, 1975b). Yet another interpretation of this method's epistemological stance, as the Delphi is

[1] Townsend (1979a) presented a similar argument in that, a 'good definition of poverty' should 'not conceal the part played by value judgements', the use of which, 'does not mean that such a definition cannot be objective' (p. 38).

often utilised to derive quantitative data which are estimated through qualitative approaches, Critcher and Gladstone (1998) and Bowles (1999), situate this method as having 'hybrid' epistemological status. These authors argue that it is this 'hybrid' status, that has allowed the method to be applied within different settings and to support various research goals.

Considering the above, this chapter follows the 'hybrid' understanding of the Delphi. As proposed in Section 1.4 of Chapter 1, the Delphi within this thesis, aims to offer a 'mixed method' approach to mitigating the challenge of selecting dimensions, indicators and weights, within poverty measurement. The method thus, seeks to be *qualitatively participatory*, in that it aims to ensure the participation of local experts in defining poverty and selecting areas of deprivation which matter the most for its measurement. When gathering these opinions, the views of all those involved within the Delphi, are valued. The level of anonymity offered by the method is utilised to support this freedom of expression. The *quantitative* nature of the Delphi within this thesis, on the other hand, arrives, in particular, from the steps taken to collect data on 'weights' for chosen dimensions and indicators of poverty. This quantitative data enables the calculation of central tendency measures and standard deviations; crucial statistics to have, which as will be illustrated within the following sections, support the assessment of whether consensus with regard to these choices, may be achieved. I argue that this 'hybrid' nature acts to strengthen the use of the method within poverty measurement; offering the ability to incorporate qualitative judgments alongside the statistical advantages of using quantitative data. The next section further explores the strengths of the Delphi compared to other commonly applied consensus seeking methods.

3.2.4 Delphi vs. Alternative Methods

3.2.4.1 Delphi vs. the Nominal Group Technique

The Nominal Group Technique, hereafter, NGT, is a 'structured group decision-making technique' used for the generation of a 'vast array of alternatives' regarding a specific issue or concern (A. L. Delbecq, Van de Ven, & Guftafson, 1974; Schuler). For the purposes of this chapter, the 'issue', would be the 'definition of poverty and its multiple dimensions and indicators'. The NGT method generally involves nine to twelve 'relevant' experts as 'panellists' and consists of 'two rounds in which panellists rate, discuss, and then re-rate a series of items or questions' (J. Jones & Hunter, 1995, p. 377). Within the NGT, individuals first, silently and individually record their ideas; this step is followed by a 'round-robin' procedure where each group member presents their ideas to the group. After all ideas are

presented, ‘the meeting concludes with a silent independent voting on priorities’ (Van de Ven & Delbecq, 1974, p. 606).

Requiring participants to ‘individually and openly’ present ideas, often proves to be a disadvantage of the NGT method. Particularly for a study on sensitive issues, such as poverty, ‘power imbalances’ may ‘derail discussion’ (S. Alkire, 2007; Cooke & Kothari, 2001). Furthermore, N. Dalkey and Helmer (1963) explain, ‘direct confrontation’ may induce ‘hasty formations of preconceived notions’, generate an ‘inclination to close one’s mind to novel ideas’ and give rise to a tendency to either, ‘defend a stance once taken’ or ‘be swayed by the persuasive opinions of others’ (p. 2). The degree of anonymity offered by the Delphi mitigates these problems, thus, seeking to support the method in synthetizing the views of individual participants to validly reflect the views of the group (Norman Crolee Dalkey et al., 1969, p. 14). Furthermore, on a more technical note, the NGT could prove to be time and cost prohibitive, as providing space for all participants to present their ideas could be time consuming (Delp, Thesen, Motiwalla, & Seshardi, 1977).

3.2.4.2 Delphi vs. the Consensus Conference

The ‘consensus conference’ involves: firstly, ‘inviting a purposive sample of individuals to a conference venue’ and ‘requiring these individuals to present on the importance of the issue at hand’ (Keeney et al., 2010, p. 16). These individuals are then divided into groups to discuss the ‘pros and cons’ of the issue. The exercise ends with ‘a plenary session’ where delegates are prompted to reveal their preferences. Given this process, within the ‘consensus conference’, ‘strong-willed individuals or groups’ may ‘dictate the direction of the discussion’, thus this method is subject to the similar weaknesses found within the NGT method (Keeney et al., 2010, pp. 16-17). As argued previously, the Delphi attempts to avoid ‘power imbalances’ by providing a level of anonymity to its participants.

3.2.4.3 Delphi vs. the Focus Group Discussion

When compared to Focus Group Discussions, hereafter FGDs, the Delphi method also offers the advantage of preventing ‘professional status and high position from forcing judgments into certain directions’ (Keeney et al., 2010, p. 16). The anonymous and iterative discussion process within the Delphi aims to ensure that estimates derived from it reflect ‘rational judgment’, which minimises the influence of certain opinion leaders (Weaver, 1971, p. 267). A further advantage of the Delphi compared to purely qualitative data collection methods, such as FGDs, as explained within sub-section 3.2.3 above, is that, the Delphi’s hybrid nature allows the assessment of qualitative judgment in a relatively objective manner by drawing upon statistical inference methods (Faulkner & Valerio, 1995, p. 168).

With the strengths of the method in mind, the next section presents a review of literature concerning key aspects of a Delphi study. Following Hasson, Keeney, and McKenna (2000), careful consideration with regard to these aspects needs to be made prior to entering the field to ensure unbiased outcomes of the study.

3.3 Review of Existing Methodological Guidelines

This section aims to illustrate existing methodological considerations, which aim to guide the application of the Delphi within this thesis. I start with a review of recommendations on sample selection.

3.3.1 Sample Selection and Size

Keeney et al. (2010) point to the fact that the Delphi does not always use a random sample which is representative of the target population, rather, the method employs the 'purposeful selection of participants', where each is assumed to be 'expert' in the issue of interest. Furthermore, Pietersma, de Vries, and van den Akker-van (2014) state that 'although a Delphi-procedure does not require representative sampling, it does require the cautious selection of panel members who are information- and experience-rich'. As this is the case, these authors note that 'the selection of experts is critical to the success of the Delphi' (Ibid, p. 1544). This inclusion of 'experts', however, leads to the dilemmas of how to define what is meant by 'expert' and how 'expertise' should be assessed.

McKenna (1994) defines 'experts' as a 'group of informed individuals'; or as Lemmer (1998) puts it, 'persons who have knowledge about a specific subject'. For the analysis of poverty, Norton and Stephens (1995) suggest expanding the participatory involvement of various 'experts' to include 'institutional stakeholders, policy makers, academics, NGOs, and most importantly, the 'intended beneficiaries of development programs', i.e. those experiencing poverty themselves. Given this diverse array of possible participants, decisions as to 'who' to include within a Delphi present a significant challenge (Crisp et al., 1997; Duffield, 1993).

Norman Crolee Dalkey et al. (1969) aim to mitigate the above challenge by setting concrete criteria with regard to the definition of 'expertise'. They suggest the use of parameters, which may be utilised to assess *how much* an individual knows about the subject matter examined (Norman Crolee Dalkey et al., 1969, p. 76). Assessing levels of 'knowledge', however, may prove to be far from straightforward. To mitigate this, Keeney et al. (2010) suggest including experts who have 'clear vested interests in the outcome of the study', similar to that suggested by Norton and Stephens (1995); this may include those who are 'key stakeholders and decision makers' with regard to a specific issue (p. 47). Alternatively,

Sitlington and Coetzer (2013) suggest the inclusion of individuals with adequate ‘years of experience’ and who ‘hold key positions’ within the study’s area of interest.

Another issue to consider is how to recruit these ‘experts’ and how many should be included within the Delphi study. Keeney et al. (2010) state that ‘representativeness within a Delphi is not based upon panel size, however, determined through the assessed qualities of the expert panel’; thus, ‘more often than not, non-probability sampling techniques, including purposeful, convenience, criterion or snowballing sampling’ are used (p. 26).

Finally, on the ideal sample size for a Delphi study, Keeney, Hasson, and McKenna (2001) point to the fact that ‘there is little agreement’ within literature on the optimum size of the expert panel. Through an extensive systematic review of existing literature, however, Okoli and Pawlowski (2004) recommend the inclusion of ten to eighteen participants within Delphi studies. This is supported by findings within A. L. Delbecq et al. (1974), de Villiers et al. (2005) and Keeney et al. (2011), which warn that ‘large panels can be difficult to manage’, furthermore, ‘increasing the group size to beyond thirty has seldom been found to improve results’.

3.3.2 Phases of a Delphi Study

The usual practice of implementing the Delphi, known as the ‘classical’ approach to applying this method, involves four distinct phases. The first phase aims to be an exploration of the subject under discussion. Within this phase, Delphi participants are encouraged to brainstorm and respond to open-ended questions to ‘allow complete freedom to generate ideas in order to identify all issues to be addressed in subsequent rounds’ (Gibson, 1998; Hasson et al., 2000, p. 1101; Harold A. Linstone & Turoff, 1977, p. 5; Sitlington & Coetzer, 2013, p. 3).

The second phase of a Delph then involves ‘grouping items generated within the first phase into clusters according to the (judged) similarity between them’; i.e. ‘the items within a cluster should aim to be more similar to one another than to items belonging to a different cluster’ (Norman C Dalkey, 2002, p. 383; Hasson et al., 2000, p. 1101). This requirement to analyse and conduct pairwise comparisons for an extensive list of items, has prompted many studies to forgo an ‘exploratory and open-ended’ approach to the first phase of the Delphi. As Hasson et al. (2000) argue, the ‘freedom to generate ideas’ within the first phase often leads to large amounts of data, which may be unmanageable (Hasson et al., 2000, p. 1101). In response to this, Hsu and Sandford (2007) state that ‘it is both an acceptable and common modification of the Delphi process to use a structured questionnaire in the first phase of the Delphi’ (p. 2). Furthermore, in defence of foregoing an ‘exploratory approach’ to the Delphi’s

first round, Pietersma et al. (2014) note that the use of a predefined list of items within the first round provides a Delphi study with a 'solid base', whereas, 'presenting participants with no list may often be too cognitively demanding, resulting in a selection of domains that often only includes those that come to mind most easily' (p. 1553). Kerlinger (1973) notes that the use of this 'modified' Delphi becomes further appropriate when 'basic information concerning the target issue is available and usable'.[1]

Given the items listed within 'phase one' and how they are grouped in 'phase two', the 'third phase' of a Delphi consists of 'iterative exercises', which require 'participants to rank items or groups of items from most important to least important' (Harold A. Linstone & Turoff, 1977, p. 5). Within this phase, participants are also given the chance to review their responses through the use of a series of structured and iterative questionnaires, commonly referred to as Delphi 'rounds' (Norman C Dalkey & Rourke, 1971; Hasson et al., 2000). Ranks within this 'third phase' are commonly expressed through the use of abstract scales that represent relative priorities. Methods of 'psychological scaling' such as 'Likert-type' scaling techniques and 'simple ranking' are often employed (Scheibe, Skutsch, & Schofer, 2002, p. 267). The use of 'Likert-type scales' and 'simple ranks', however, may present significant weaknesses; particularly for the purposes of the Delphi proposed within this thesis, where rather than 'ranks', participants are requested to assign 'weights' to dimensions and indicators of poverty. On this, Scheibe et al. (2002) argue that when the primary reasons for using a Delphi include the determination not only of *which* items are considered most important, but also the *degree* to which each item is preferred over the other possibilities, an 'interval scale' that expresses this degree of 'item importance' should be obtained (p. 267).

Considering the above, the 'fourth phase' of a Delphi acts as a 'final evaluation', which occurs when all previously gathered information has been initially analysed and results have been fed back for re-consideration by participants through the iterative process within 'phase

[1] Other than the 'classical' and 'modified' Delphi, additional 'types' of empirical applications of the Delphi method exist. Two, which may be particularly relevant for this thesis, are the 'decision' and 'policy' Delphi. The 'decision' Delphi aims to 'create reality' through the 'deliberate inclusion of participants based on their position in the decision making process' (Rauch, 1979, p. 163). Within a 'decision Delphi', reality is generated by the Delphi itself as what is sought is not for a 'situation to be correctly understood and described', the essential thing is rather what 'the decision makers define as real' (Rauch, 1979, pp. 163-164). As for the 'policy Delphi', the iterative process within this method focuses on 'ongoing discussion' and seeks 'relevant arguments with regard to policy formation'; the process, thus, is not aimed to be a mechanism of 'defining' and 'reaching consensus' on issues (Keeney et al., 2010; Kuusi, 1999; Rauch, 1979). Considering these alternative 'types', the 'modified Delphi' continues to be most appropriate for the purposes of this thesis, as its aim, i.e. to gain a reliable level of 'consensus' on relevant issues, goes hand-in-hand with this thesis' goal of seeking reliable consensus on components of a multidimensional poverty measure (Keeney et al., 2010, p. 5).

three' (Harold A Linstone & Turoff, 2002, p. 6). This 'final phase' involves the assessment of the 'validity' of the resulting judgment of the entire group, which is typically measured in terms of the 'degree of consensus' among the experts included in the study (Harold A Linstone & Turoff, 2002, p. 22). This in mind, the next sub-sections explore empirical guidance on the optimum number of iterative rounds to employ, alongside statistical tools to assess whether consensus is achieved.

3.3.3 Number of Iterative Rounds and Type of Feedback between Rounds

Keeney et al. (2010) note that in order to provide the opportunity for participants to revise their earlier responses, the Delphi is required to consist of at least two rounds; no other strict guidelines however, exist with regard to the correct or optimum number of rounds to conduct (p. 81). This 'ideal number' of rounds could depend on the time available or as noted within sub-section 3.3.2 above, may depend on whether the study was started with a broad question or if opinions were prompted based on a pre-existing list of items (Ibid, p. 82).

On this issue, McKenna (1994) suggests that 'diminishing returns' to the Delphi process may be achieved by administering the number of rounds which 'minimises reductions in the amount of new information collected'. In addition, Starkweather, Gelwicks, and Newcomer (1975) note the importance of factoring in participant 'fatigue' as 'it bears on whether conclusions are valid or whether they have resulted from over or under consensus' (p. 38). Whitman (1990) concurs with this warning and notes that 'the length of a Delphi places many demands towards its participants', which can directly 'affect their motivation and choices identified within the study'. Considering this, the point of 'diminishing returns' is thus often determined as: the number of rounds in which reductions in response rate due to fatigue is minimised.

Starkweather et al. (1975) found that participants usually 'become fascinated by the Delphi process at rounds 2 and 3; but just as typically, they may agree to anything to avoid recycling the survey beyond rounds 3 or 4; thus, risking the identification of a false consensus if the number of rounds is extended beyond three' (p. 38). Harold A Linstone and Turoff (1975) and Dietz (1987) also note that 'three rounds' within a Delphi study, seem to be typical in the published literature. On this issue, Keeney et al. (2010) argue that a decision needs to be made between gaining consensus on all items and the importance of maintaining high response rates; thus, 'researchers commonly accept that some items are

not going to gain consensus after three rounds and will stop the exercise at this point' (p. 82).[1]

With the above guidelines in mind, the next issue is the 'type of feedback' to be presented to participants between rounds. G. Rowe, Wright, and McColl (2005) argue that in conjunction with participant anonymity, it is generally assumed that the Delphi improves judgmental accuracy, because of the feedback it provides between rounds (p. 378). These authors note that the 'lack of cues' as to the identity of the sources of the feedback allows participants to 'focus upon the content of that feedback rather than be distracted by extraneous social information' (G. Rowe et al., 2005). Furthermore, feedback between Delphi rounds, 'informs participants of the current status of their collective opinion and helps them identify items that they may have missed or thought unimportant' (Goodman, 1987; McKenna, 1994). Within the Delphi, it is at this point that a participant may change or modify their opinion, in light of the judgment of others in the study (Couper, 1984; McKenna, 1994, p. 9; Munier & Ronde, 2001). Despite a lack of concrete guidance, recommendations on this issue do exist within literature on the types of feedback to utilise within a Delphi.

Crisp et al. (1997) state that one of the most common forms of feedback consists of measures of central tendency (mean, median), which may or may not be accompanied by a measure of dispersion (standard deviation). On this, Murphy et al. (1998), Armstrong (2001) and Theodore Jay Gordon (2003) and von der Gracht (2012), among others, argue for the use of the median rather than the mean and standard deviation, for the reason that medians are generally more robust to outliers. In addition to these measures of central tendency and dispersion, Norman Crolee Dalkey et al. (1969) suggest requesting participants who are at the extremes of the distribution to provide 'explanations' as to their decisions; these 'explanations' could also be included within feedback between rounds (p. 58). Caution, however, needs to be exercised with regard to the amount of information presented to participants between rounds. Including feedback beyond simple statistical reports may lead to rapid 'diminishing returns', as piling information on participants may increase the chances of the 'biasing effects of fatigue' (Norman Crolee Dalkey et al., 1969, p. 58). This thus, leads to another consideration, i.e. how much time should be allocated between the Delphi's iterative rounds.

Waldron (1971) notes that 'timing between rounds' is an essential issue, where 'the quality of the Delphi increases as the time between filling in a questionnaire and the next one being sent, becomes shorter'. Keeney et al. (2010) and Sandrey and Bulger (2008) warn that 'time

[1] Keeney et al. (2010) further argue that 'sometimes items, which do not gain consensus after three rounds, are as important in terms of findings as those items that do gain consensus' (p. 82).

delays' between rounds may affect 'motivation and choices' identified by participants. In addition, time delays and the long temporal time commitment required of participants, could lead to high attrition rates (Donohoe & Needham, 2009, p. 421). On this, Walker and Selfe (1996) and Sumsion (1998) state that, 'in order to maintain rigour, a 70% minimum response rate should be achieved' (p. 679).

3.3.4 Defining Consensus

Following Keeney et al. (2010), it is important to define what a reliable consensus level is, *before* embarking on a Delphi study, rather than determining it post hoc. On this issue, Keeney et al. (2010) note that there are different types of criteria for describing when consensus is reached, nevertheless, two of the most commonly used are the 'percentage level' and 'statistical' approaches (p. 45).

Within the 'percentage level' approach, consensus is defined when a certain percentage of participants' votes falls within a prescribed range (Williams & Webb, 1994). Yang (2003) labels this 'prescribed range' as the 'stopping criterion' (p. 3). Guidelines as to 'accepted standards' for this 'percentage range' are however contentious; the use of this criteria thus, tends to be somewhat arbitrary (Keeney et al., 2001; McKenna, 1994; Ulschak, 1983). On this issue, Fink, Kosecoff, Chassin, and Brook (1984) state that 'there are no firm rules for establishing when consensus is reached'; the only clear realisation with regard to this issue is that 'the stricter the criteria, the harder it will be to achieve consensus' (p. 982). Due to this challenge, Dajani, Sincoff, and Talley (1979), Crisp et al. (1997) and Scheibe et al. (2002), criticise the use of the 'percentage level' approach and suggest the use of the 'statistical' approach, alongside measures of the 'stability of responses' over successive Delphi rounds, as ways to assess whether consensus has occurred.

For the 'statistical' approach, Murray and Jarman (1987) and Hasson et al. (2000) suggest the use of measures of central tendency, i.e. mean, median and mode, to illustrate the final collective judgment of the group. Andre L Delbecq, Van de Ven, and Gustafson (1975) and Dietz (1987) suggest calculating these statistics from results of the 'final' Delphi round, which they argue represent the 'aggregate' judgment of Delphi respondents. The reason behind this is that, as the 'final round' came about through 'the deliberation and consideration of feedback' from preceding rounds, results from this round may arguably serve as a 'reliable summary' of the opinion of the group. Dietz (1987), among others, stresses the importance of using 'medians' from the 'final round', as they are less sensitive to extreme opinions.

In addition to the above, Crisp et al. (1997) argue for the assessment of the 'the stability of responses' as an indicator of consensus. This 'stability' refers to the consistency of

responses between successive rounds of a Delphi. Dajani et al. (1979) note that ‘while stability does not necessarily imply a given level of agreement, it is only when a stable answer is reached that the analysis of the level of agreement should be attempted’ (p. 84).

Literature on the Delphi recognises two types of ‘stability’; i.e. (1) the stability of ‘individual’ responses and (2) the stability of ‘group response’ between rounds. Dajani et al. (1979) state that as ‘it is highly unlikely that one will occur without the other, the examination of group stability should be prioritised in order to assess consensus’ (p. 84). These authors argue that consensus is ‘meaningless’, if ‘group stability’ has not been reached beforehand. Scheibe et al. (2002) support this conclusion by positing that since the interest of the Delphi method lies in the ‘opinion of the group rather than the individual, assessing group stability should thus be a necessary criterion for consensus’.

To measure ‘group stability’, English and Kernan (1976), Dajani et al. (1979) and Yang (2003) suggest checking for changes in the Coefficient of Variation of responses, hereafter CV, between two successive rounds to examine the existence of consensus. The CV is a ‘standardized measure of dispersion and may act as a useful tool within the comparisons of distributions’ (von der Gracht, 2012, p. 1530). It is a ‘dimensionless number and is calculated as the standard deviation divided by the mean’ (von der Gracht, 2012, p. 1530). English and Kernan (1976) and von der Gracht (2012) state that, ‘various empirical pieces of literature have used this measure to assess consensus within Delphi studies, as it allows for a direct comparison of statements from succeeding rounds’. This argument is strengthened by Shah and Kalaian (2009)’s study, which compared several parametric statistical methods commonly used to analyse Delphi results and found that examining the CV was, ‘the best procedure to assess reliability in Delphi studies’ (p. 226). In addition to this, their study found that a difference of ‘0.5 and below’ between the CVs of successive rounds indicates a good degree of consensus and ‘no need for additional rounds’.

Besides the above, to measure ‘group stability’, G. Rowe and Wright (1999) suggest the use of the variance between rounds, where less variance is understood to mean greater consensus. In addition to variance about the mean, measures of variability around the median indicating the level of ‘conformity to the group median’, was also suggested (Armstrong, 2001; Theodore Jay Gordon, 2003; Matthew R Hallowell & Gambatese, 2009; Jordan & Javernick-Will, 2013; Murphy et al., 1998). A common measure to utilise is the ‘Median Absolute Deviation from the Median’ or ‘MAD’, which may act as a robust alternative to the standard deviation, particularly when the possibility of outliers is evident (Leys, Ley, Klein, Bernard, & Licata, 2013).

3.3.5 Reflections on Research Ethics within a Delphi Study

As in all research involving human participants, Delphi studies should strive to take the utmost care to ensure ‘respect for human dignity’ (Keeney et al., 2010, p. 105). This in mind, written explanations and guides to the Delphi proposed should be presented to each potential participant before the study takes place. Keeney et al. (2010) suggest that participants receive this information a minimum of two-weeks before the study takes place to provide ample time for them to decide whether to take part. In addition, balance should be achieved between the need for reliable and unbiased information, versus respect for participants’ time and commitment. Hitch and Murgatroyd (1983) suggest maintaining contact and reminding participants that ‘success depends on completed questionnaires and a quick return of answers’ (p. 415). Beretta (1996) however, notes that this approach to maintaining response ‘could cause respondents to feel forced’. A balanced and respectful approach is thus key.

Another crucial issue within a Delphi study is that of anonymity and confidentiality. Keeney et al. (2010) note that, dependent on how the Delphi is conducted, complete anonymity may not be achievable. This may be the case when researchers need to link each participant with their response in order to provide them with a summary of their individual responses. Keeney et al. (2010) suggest allocating a ‘unique code’ to each participant, keeping all data that links the participant to this ‘code’ within a password protected device and restricting access to this information to only the researcher (p. 106). Complete anonymity becomes a further challenge when participants within the Delphi study may know, or at least, recognise each other. G. Rowe and Wright (2001) dub the condition of ‘recognising one-another, yet maintaining the anonymity of individual answers’, as ensuring ‘quasi-anonymity’. Even though unable to achieve ‘complete’ anonymity, Keeney et al. (2010) note that maintaining this ‘quasi anonymity’ throughout the Delphi, would allow the method to continue to ‘facilitate the openness and truthfulness of participants, without the fear of offending others’. Regardless of the level of anonymity achieved however, it is ‘imperative that the issue of anonymity is disclosed to each participant before the study takes place’ (Keeney et al., 2010, p. 107).

Alongside the above, possible power imbalances may occur, between participants and the researcher herself. On this issue, Keeney et al. (2010) point to the importance of being aware that researchers ‘should not make judgements on what the content of subsequent Delphi rounds should be, however must only facilitate the process for participants’ (p. 109). Researchers within a Delphi should be cautious and careful that their actions do not influence decisions and push participants toward consensus (Crisp et al., 1997).

3.4 The Bogor City Delphi Protocol

With the guidance described within the previous section in mind, this section outlines the steps taken for the Delphi conducted within this thesis and is divided into three main parts; i.e. 3.4.1 Steps taken before the Delphi, 3.4.2 Steps taken during the Delphi and (3) Steps taken following the Delphi.

3.4.1 Pre-Delphi Steps

3.4.1.1 Pre-pilot and Pilot Tests

Keeney et al. (2010) states that 'effective planning' is key to the success of a Delphi study. This in mind, the Delphi conducted within this thesis involved eighteen months of planning, which culminated in two pilot tests conducted before entering the field. Firstly, a pre-pilot test was carried out with post-graduate students at the University of Oxford. Secondly, a pilot test with policy makers was held in Jakarta, Indonesia. Both pilots aimed to assess the workings of the proposed Delphi questionnaire, whether it could be easily understood and whether the online method through which it would be administered provided an adequate platform for participants to express their views. The pilot tests also sought to evaluate the effectivity of the type of feedback presented between Delphi rounds, the impact of the length of the Delphi on participant fatigue and whether the proposed method to measure consensus was suitably acceptable. This sub-section will focus on how results from these pilots helped shape the final Delphi study proposed in this thesis.

The pre-pilot Delphi was conducted in May 2016 at Trinity College, the University of Oxford. Eleven post-graduate students who were fluent in Bahasa Indonesia were included in order to facilitate the assessment of the wording of the questionnaire. As will be described further in sub-section 3.4.1.2 of this chapter, the final Delphi within this thesis utilised Bahasa Indonesia for its language of instruction in order to facilitate deeper understanding for its proposed participants. Results of this pre-pilot test indicated that the use of an online questionnaire, facilitated by the Qualtrics platform, enabled a straightforward and accessible way of collecting opinions, alongside analysing and presenting feedback between rounds. This strengthened the advantages of administering what Keeney et al. (2010) dub as an e-Delphi.[1] On use of this approach, Keeney et al. (2010) however note that, the researcher

[1] By utilising this online platform, the Delphi within this thesis aims to be an e-Delphi, rather than what T. Gordon and Pease (2006) and Theodore J Gordon (2009) dub as a 'real-time' Delphi, which is 'round-less'; i.e. rather than using iterative rounds, within a real-time Delphi, results from previous rounds are presented to participants immediately as they input and save their answers.

should ensure that participants are computer literate, have easy access to devices which support an online questionnaire and a secure internet connection.

As will be discussed further within this chapter, participants within this thesis' Delphi consisted of policy makers. Considering this, a second pilot test was conducted in July 2016 with twelve policy makers from Indonesia's central government. Results from this second pilot confirmed the utility of using an online questionnaire platform. Furthermore, concerning the 'type' of feedback between rounds, whereas students within the pre-pilot were happy with analysing detailed statistical reports, policy makers within the pilot, indicated preference toward summary graphs. An additional important consideration was that participant 'fatigue' was indeed found to be a significant factor when designing a Delphi.

In addition to the above, both pilots indicated that results from the final Delphi round may act as a reliable summary of the group's response. Furthermore, measures of the 'stability of group responses' between rounds, i.e. the convergence of answers, were considered as important statistics to assess whether a reliable level of consensus is evident.

3.4.1.2 Sample Selection and Size

Table 3.1 presents a summary on the methodological guidelines utilised within the Delphi in this thesis. These guidelines aim to address the three questions stated within the Introduction of this chapter, i.e. (1) who and how many 'experts' to include within the Delphi, (2) how many iterative rounds to apply and the type of feedback to present between these rounds and (3) how consensus should be measured. I start with the criteria used to guide decisions on sample selection and size.

Following recommendations discussed in sub-section 3.3.1 of this chapter, on the selection of 'experts', Keeney et al. (2010)'s reasoning is followed, i.e. individuals with a vested interest in the outcomes, who are decision makers and key stakeholders, were selected to be included in this thesis' Delphi (p. 47). As discussed in Chapter 1, considering limited time and resources, focus is given to policy makers. 'Expert' within this thesis thus, refers to senior level policy makers, who held key positions and were still in active office within the case study area at the time of the Delphi.

Once identified, the next step to consider was how to recruit these 'experts' and how many should be recruited. Following MacNealy (1999), 'experts' within this thesis' Delphi were selected through 'purposeful sampling'. This sampling methodology was supplemented by an 'inclusion protocol', which consisted of a list of characteristics essential for a participant to fulfil in order to be included in the study (Keeney et al., 2010, p. 47; Sitlington & Coetzer, 2013).

Table 3. 1 Methodological Aspects of the Bogor City Delphi

Methodological Aspect	Guidelines Followed
Sample Size	Minimum: 30 policy makers.
Sample Selection	Expert criteria: Senior level policy maker or: Has substantial years of experience on poverty and social exclusion, which could be supplemented by significant 'perceived expertise' on this issue
Delphi phases	Modified Delphi (forgo explorative first phase)
Questionnaire	Pre-specified indicators of poverty available in Susenas. Use of online questionnaires.
Number of Rounds	Three iterative rounds
Type of Feedback	Each participant was given summary statistics of group response from previous rounds (mean, median, mode, standard deviation, MAD). Participants given 20 minutes before each round to individually review feedback. Researcher presents the feedback to participants before each round and allocates 20 minutes for possible questions from participants.
How Consensus is Measured	Measures of stability between rounds: Standard deviation, MAD, CV (difference of < 0.5 between the CV of successive rounds) Final group response: median of final round results.

Source: Author's summary

Referring to the criteria proposed within Keeney et al. (2010) and Skulmoski, Hartman, and Krahn (2007), these characteristics required that participants: (1) share a common interest in the area of poverty alleviation, (2) possess the necessary amount of practical experience, (3) have the capacity and willingness to contribute and (4) exhibit good written and communication skills. In addition to this, participants' academic qualifications were also recorded as a possible proxy of 'expertise' (Keeney et al., 2010, p. 47; Skulmoski et al., 2007). Furthermore, participants were also selected with the aim of representing possible different points of view regarding the crucial areas of deprivation that need to be addressed when measuring poverty in their district (Kaynak, Bloom, & Leibold, 1994). Considering this, the sample for this thesis' Delphi included senior officials from *all* sectors of the Bogor City Government, with each sector sending one or two representatives to attend the study. These sectors represented the myriad welfare priority areas, which the government focuses on to maintain the well-being of its people. Following Sinéad Hanafin et al. (2007), care was taken to include an equal number of participants from each of these government sectors.[1]

[1] A detailed list of the sectors included within this thesis' Delphi is available within the Appendix to this chapter.

As to the number of experts, guidelines set within Keeney et al. (2011) were followed; i.e. that the inclusion of 'more than thirty' participants is often difficult to manage. With the number of local government sectors in mind however, slightly more than thirty participants, i.e. thirty-six, were included within the study. The recruitment process for these participants was as follows.

Two months before the Delphi took place, working together with the district's Mayoral Office, a list of potential participants, i.e. Heads of Sectors and Heads of Sub-districts, was compiled. Invitation letters were then sent. To adhere to the ethical considerations listed within sub-section 3.3.5, within this invitation a written explanation of the study was presented to each participant. This was in the form of a short booklet attached to the formal invitation presenting the aims of the study, what the study involves and a detailed schedule. A description of how the Delphi method works, what was expected of participants and details on ethical clearance of the study were also presented. These invitations and booklets were sent to participants three weeks before the study.

During the study, the goal of achieving at least quasi-anonymity was strictly maintained. Complete anonymity was less possible as the study involved senior level policy makers, who were all working within the same government, thus, the potential that participants may know or at least recognise one another, was apparent. In order to maintain quasi-anonymity, participants were not required to record their names on their responses. Consent was requested before participants took part in the study and before photographs, video and voice recordings of participants were made. Data and results from the study were kept safe and secure within laptops and hard drives that were password protected. A CUREC1 application was approved and ethical clearance from the University of Oxford Central University Research Ethics Committee was obtained.

3.4.1.3 Delphi Phases and Questionnaire

The Delphi conducted within this thesis follows the design of Hsu and Sandford (2007), Rossouw, Hacker, and de Vries (2011) and Sitlington and Coetzer (2013)'s 'modified' Delphi. This in mind, referring back to sub-section 3.3.2, the Delphi within this thesis *forgoes* the first 'brainstorming' phase of a 'classical' Delphi, by presenting participants with a structured questionnaire containing a pre-determined list of indicators of poverty to choose from.

As summarised within Table 3.1, these pre-specified indicators were available within the nationally representative Susenas survey. The Susenas offers a substantial range of variables concerned with household welfare in Indonesia; thus, following Kerlinger (1973),

forgoing the 'brainstorming' phase of the Delphi is 'supported' as basic information on possible individual and household level deprivations which may contribute to poverty was available through this dataset. Furthermore, alongside the importance of reducing participant fatigue and considering ethical issues on participant time-use, the use of existing variables within the Susenas which were concrete in their nature, sought to mitigate the possibility of 'differences in interpretations' with regard to how participants defined these indicators. Sinéad Hanafin et al. (2007) dub this, 'the challenge of mediating different stakeholder understandings' (p. 99). It was crucial for each indicator included within the study, to represent concrete conditions describing a household's socio-economic characteristics, particularly in the hopes of maintaining the comparability and robustness of how each participant chose and assigned weights to each of these indicators. The use of Susenas variables also enabled the direct application of the Delphi results to the construction of concrete empirical measures of multidimensional poverty, which will be the subject of Chapters 5 and 6 of this thesis.

Finally, with regard to how opinions on 'weights' were collected, following von der Gracht (2012), participants were required to utilise an interval scale to express their views.[1] This in mind, participants were requested to assign 'weights' ranging from zero to one to each dimension and indicator (with "0" indicating 'no importance', and "1" indicating highest importance'), where the sum of dimension weights had to equal "1" and the sum of indicator weights within dimensions also had to equal "1.[2] A built-in calculator was included within the questionnaire to remind participants of this point.

3.4.1.4 Number of Rounds, Type of Feedback and Consensus Criteria

Referring to the guidelines explored within sub-section 3.3.3 of this chapter, the Delphi conducted within this thesis will follow the common practice of limiting the study to three iterative rounds. As for type of feedback between rounds, as listed within Table 3.1, measures of central tendency, i.e. means, medians, modes, standard deviations and median absolute deviation from the median (MADs), were presented to participants in table and graphical displays. Participants' attention was particularly drawn toward the importance of the 'median results' of each round.

[1] Sin Hanafin et al. (2007) note the advantage of allowing participants to assign zero importance to a dimension or indicator, compared to the use of Likert scales and simple ranks, which may lead to 'all items being considered as important for the mere fact of filling in the form forces some rank ordering' (p. 87).

[2] Reasons behind this are discussed within sub-section 6.3.3 of Chapter 6.

The Delphi within this thesis also included an exercise within its iterative rounds, which requested that participants state 'reasons' behind possible disagreements with the 'feedback' presented to them between rounds. This exercise was conducted with the main purpose of providing time and space for participants to contemplate their answers and anonymously write arguments for or against the feedback given. This information was, however, not presented to participants within their feedback sheets in order to simplify the feedback they received and minimise the effects of participant fatigue.

To analyse consensus, 'group stability' of responses was first examined; if stability occurs, results from the third, i.e. final Delphi round, are considered as the final results of the study. Particularly for dimension and indicator weights, following Dietz (1987), Armstrong (2001), and Matthew R. Hallowell and Gambatese (2010), the median of the outcomes from the third Delphi round are utilised to reflect these final results. As listed within Table 3.1, 'group stability' will be assessed by examining the differences in the Coefficient of Variation (CV) between successive rounds. Guidance within Dajani et al. (1979), English and Kernan (1976), Shah and Kalaian (2009) and Yang (2003) is followed in that a 'difference of 0.5 and below' between the CVs of Round 2 and Round 3 indicate a reliable level of consensus. As discussed within sub-section 3.3.4 of this chapter, for measures of dispersion, particular focus will be given to the median absolute deviation from the median (MAD).

3.4.2 Implementation of the Delphi

The Bogor City Delphi was conducted during two working days in August 2016. For both days, the study was conducted within the Mayoral City Hall of Bogor City. Conducting the Delphi within the City Hall significantly reduced travel time and costs for participants as it was located within the centre of Bogor City. The City Hall also offered a safe and secure area to conduct the survey.

Similar to the Delphi conducted within the pre-pilot and pilot, questionnaires for both days were developed using Qualtrics. As argued for within Witkin and Altschuld (1995) and confirmed by the pilot tests described within sub-section 3.4.1.1, the use of Qualtrics' electronic questionnaires helped ensure the speed of transmission of the questionnaires, maintained respondent anonymity and supported the development of rapid feedback between rounds. Each participant received their own computer where a link to the Delphi questionnaire was made available within their desktops.[1]

[1] Funding for this Delphi study was generously provided by the Jardine Foundation. In addition, the Bogor City Government also gave the study the generous use of its Green Technology IT Conference Room located within the Bogor City Hall throughout the two-day study.

Conducting the Delphi during two consecutive days made it possible to establish full commitment of participants, uphold the ethical aim of the study of valuing participants' time and significantly diminished the possible challenges faced when a large level of attrition arises. As Whitman (1990) argues, 'the length of the Delphi depends on the commitment of the individuals participating'. Extending the Delphi to span longer than two days was not feasible, as participants had to forgo their working days in order to fully commit to this study. Figure 3.1 describes the steps taken within these two days.

3.4.2.1 Delphi Day 1

As described within Figure 3.1, given the decision to apply a 'modified' Delphi and to utilise pre-specified indicators available within the Susenas, Day 1 of the study primarily acted as the 'grouping of indicators into clusters' phase and was divided into three steps.

Firstly, before the 'clustering' exercise, participants were asked to respond to a demographic and 'perceived expertise' questionnaire which aimed to assess participants' level of expertise.[1] Within the second step, participants were then requested to respond to the open-ended question of 'What does poverty mean in Bogor City? What is the definition of poverty that best describes the condition of being poor in Bogor City?' Participants were reminded that they should respond to this question as 'representatives' of their community. Following this, based on these definitions, participants were requested to answer the open-ended question of whether an aggregate or dashboard measure of poverty was more relevant to measure poverty within their community; if 'aggregate', 'how should the dimensions and indicators be grouped to reflect the condition of poverty?' If 'dashboard', 'how many dimensions should indicators be clustered into?'

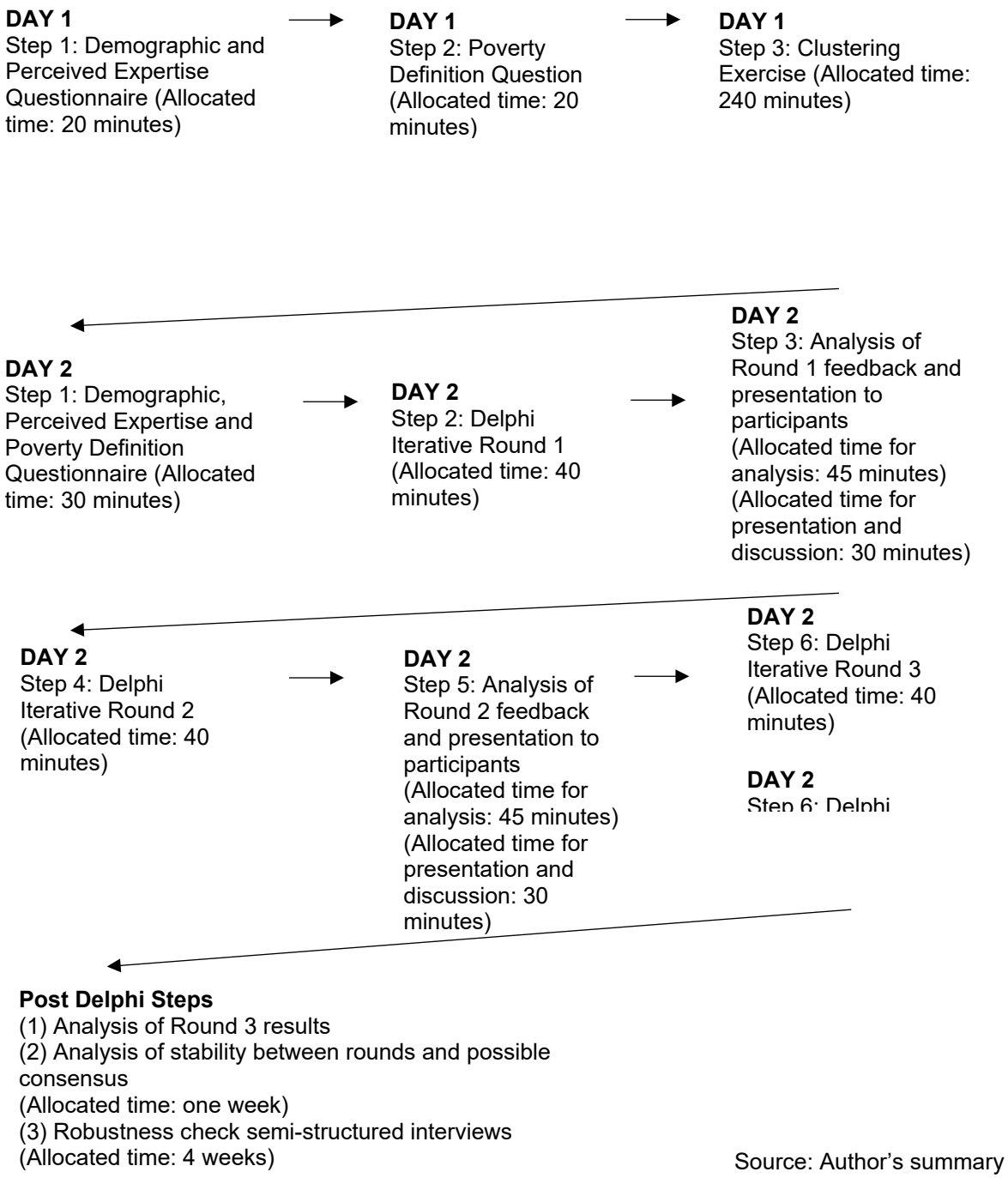
Finally, within the third step, the 'clustering step', following guidelines set within Norman C Dalkey (2002) and Norman C Dalkey et al. (1972), participants were requested to rate the similarity *o*f all possible pairs of indicators available within the Susenas.[2,3] For this exercise, each participant received a computer generated list of pairs of indicators. To reduce possible biases which may occur if a particular pair of indicators were continuously given to participants first or last, within the 'clustering questionnaire', random orderings of these pairs was presented to participants.

[1] These demographic and 'perceived expertise' questionnaires were designed to replicate those developed within Scheibe et al. (2002). Section 4.2 of Chapter 4 discusses these questionnaires further.

[2] Although using a sample of students as participants, Norman C Dalkey (2002) provided an example of the application of the Delphi in solicitating aspects of the 'quality of life'.

[3] As will be discussed further within Chapter 5, the clustering exercise also aimed to fulfil Martin Ravallion (2011b)'s argument for grouping indicators into dimensions.

Figure 3. 2 The Delphi Implementation Process



Source: Author's summary

The Susenas provided a list of 28 pre-specified indicators. A complete list of these indicators is provided within Table 4.6 of Chapter 4. For the purpose of this study, the possible 'pairs' of indicators were generated using Stata 14. The formula for the number of independent pairwise comparisons was $\frac{k(k-1)}{2}$, where k is the number of indicators (Nordstokke &

Stelnicki, 2014). Considering this and given the 28 pre-existing indicators available within the Susenas, participants were thus required to conduct 378 pair-wise comparisons.

To measure ‘similarity’ between indicators, participants were asked to assign a ‘proximity score’ to each pair of indicators. This score was described by a number within the “0” to “4” scale; with “0” indicating complete similarity (indicators are practically the same), “1” indicating that they are closely related, “2” moderately related, “3” slightly related, and “4” indicating that they are completely unrelated. This exercise was conducted within a four-hour time slot.

Following Norman C Dalkey (2002), the resulting proximity scores were ‘transformed into a group matrix and processed with an agglomerative hierarchical clustering routine’ (p. 386). The resulting ‘clusters of indicators’ were considered as ‘dimensions’. Robustness checks of the resulting clusters were conducted through testing for internal cluster validation and relative cluster validation. Measures of ‘cluster compactness’, ‘silhouette coefficients’ and ‘dun indices’ were also calculated (Kassambara, 2017).[1] Alongside participants’ opinions on whether an ‘aggregate’ or ‘dashboard’ approach suffices for a reliable poverty measure, results from these robustness checks helped to guide the decision of ‘where to cut the hierarchy of clusters’; i.e. the optimum number of dimensions to include within a multidimensional poverty measure (Everitt, 2011). Results from this clustering exercise were analysed using R and fed into the questionnaire for Day 2.

3.4.2.2 Delphi Day 2

On Day 2, before the main iterative exercise, participants were asked the same demographic and perceived expertise questions as were administered on Day 1. After completing these questions, each participant was again asked to describe what poverty in Bogor City means to them. Participants were again reminded to answer these questions as representatives of their people. Participants were then encouraged to base all their answers within subsequent iterative rounds on how they defined poverty. This exercise of requiring participants to first define poverty, before choosing its dimensions, indicators and weights, sought to establish what Ringen (1988) calls a ‘logical line of deduction’ between definition and measurement.

Within each of the Delphi’s iterative rounds, a structured questionnaire was utilised, building on the list of indicators available within the Susenas. These indicators were clustered into dimensions according to the results of the clustering exercise conducted on Day 1.

[1] Results for these robustness tests are presented within the Appendix to Chapter 4.

Participants were requested to indicate which dimensions and indicators were of value within the construction of a measure of poverty for their community and to assign weights to these items. These weights aimed to represented degrees of importance and possible levels of trade-off between dimensions and indicators. The list of dimensions and indicators were presented in random order for each participant to minimise potential bias that may occur if the same dimension or indicator appeared first or last for every participant. As noted within sub-section 3.4.1.3, for the weighting exercise, participants were required to assign weights with values ranging from “0” to “1”.

Between each round, graphical summary statistics consisting of measures of central tendency and dispersion of answers were presented to each participant. A link to these results was distributed to participants; thus, they could view them on each of their computer screens. These results were also displayed on big projected screens at the front, back, left and right walls of the Conference Room where the Delphi study took place. In addition to this, before each consecutive round, the researcher allocated 30 minutes to discuss these results with participants. Within these periods, participants, if they were willing, were given the chance to ask questions and discuss these results. The researcher took particular care to present results in an objective way, focusing on the data at hand, without the addition of the researcher’s own subjective views.

Feedback was calculated and analysed by the researcher during 45-minute coffee and lunch breaks between rounds. Within each break, participants were encouraged to join in seminars and video presentations that strove to be unrelated to the Delphi study’s subject matter.[1] To minimise bias, it was crucial that participants remained within the Conference Room and did not discuss answers among themselves or other non-Delphi participants.[2] Although there is no way of guaranteeing that this did not happen, presenting participants with activities between rounds sought to mitigate this challenge. The next sub-section describes the robustness checks conducted to assess the reliability of the Delphi’s results.

3.4.3 Post Delphi Steps

Aware of the rapidity of the Delphi study conducted, robustness checks needed to be done in order to determine whether the consensus reached was not what Dietz (1987) terms as

[1] For the break between Rounds 1 and 2, a speaker was invited to talk about climate change and for the break between Rounds 2 and 3, videos of Oxford city and of the University of Oxford were displayed.

[2] On both days, all participants remained in the Conference Room for the whole length of the Delphi process. The room had eating and rest-room facilities which enabled this.

an ‘irrational consensus’; i.e. ‘one which might only be temporary’ (p. 271). This in mind, steps developed within Scheibe et al. (2002) were followed.

Firstly, a ‘Post-Delphi robustness check questionnaire’ was administered at the end of the Delphi’s third round. Secondly, follow-up individual semi-structured interviews with participants and non-Delphi participants were conducted. The ‘Post-Delphi questionnaire’ was primarily utilised to assess whether the feedback between rounds, had any significant effect on changes in participants’ responses and views (Scheibe et al., 2002, p. 270). Within this questionnaire, participants were required to respond to questions on whether in general, they tended to agree with the feedback presented, whether they learnt from this feedback, alongside questions seeking to measure their degrees of satisfaction with the Delphi process itself.[1]

As a further step to check the robustness of the Delphi’s results, individual semi-structured interviews with fifteen Delphi participants and fifteen non-Delphi participants were conducted. Although the Delphi study was conducted anonymously, participants voluntarily recorded their names and contact details on a consent sheet; this list of names was used as a sampling tool to randomly select the fifteen Delphi-participant-interviewees.[2] The fifteen non-Delphi-participant-interviewees were selected from a list of heads of sectors and their deputies, provided by the city government. Purposeful sampling was employed to select these non-participant interviewees with the aim of covering members of different government sectors. Within the first part of these semi-structured interviews, similar to that asked within the Delphi, participants were first prompted to describe how they defined poverty, they were then encouraged to express which dimensions and indicators they thought were relevant, alongside how they should be weighed. After expressing these views, participants were given a booklet summarising the results of the Delphi study. They were then asked whether they agreed or disagreed with these results. Opinions on the Delphi method itself were also gathered. The next section concludes this chapter by discussing the limitations of the Delphi described and how this thesis proposes to mitigate them.

3.5 Limitations of the Study

The approach offered within these pages exhibits a number of limitations. These limitations arise from three main sources, which are: biases that could arise from (1) the vagueness of

[1] Section 4.7 of Chapter 4 discusses this questionnaire and its results in further detail.
[2] For Delphi-participant post-Delphi interviews, even though participants were selected randomly, care was taken to make sure that these interviewees included participants from different sectors with a mix of demographic characteristics, i.e. ‘gender’ and ‘years of experience’.

existing Delphi guidelines, (2) sample selection within the Delphi and (3) use of pre-existing data for the Delphi questionnaire.

I start with the first limitation. As explored within sub-section 3.2.3 of this chapter, the different epistemological paradigms that the Delphi may be attributed to and the flexibility in its application, although contributing to the method's wide use, also serve as its key constraint (Keeney et al., 2010, p. 21). As Harold A. Linstone and Turoff (1977) state, 'if anything is true about the Delphi today, it is that in its design and use, the method can be considered more of an art than a science' (p. 3). Keeney et al. (2010) argue that lack of agreed guidance on how to use the Delphi may cause dilemmas for researchers in the field, which ultimately questions the method's rigour (p. 21). To mitigate this, as illustrated within this chapter, utmost care was taken to follow existing guidelines for the Delphi's application. Through the detailed steps recorded within this chapter, it is also hoped that transparency with regard to the decisions taken within the construction of the Delphi is also achieved.

The second limitation concerns the 'type' of participants included within this thesis' Delphi. As noted within sub-section 3.3.1 of this chapter, participatory methods in poverty measurement would very much benefit from a broad range of participants, consisting of policy makers, NGOs, and most importantly, communities experiencing poverty themselves (Mukherjee, 2006; Narayan, Patel, Schafft, Rademacher, & Koch-Schulte, 1999; Norton & Stephens, 1995). As the primary focus of this chapter and the next is methodological and due to limitations in time and resources, a focus on one type of participant, i.e. 'policy makers', had to be made. When possible, additional Delphi studies that include different types of participants would thus, be beneficial.

The next possible source of bias stems from the use of pre-existing data provided by the Susenas within the Delphi questionnaire. As Duffield (1993) notes, 'this approach may bias responses and limit available options'. Considering this, to mitigate the potential existence of 'missing indicators', each participant within the Delphi was given the chance to add to this list of pre-specified Susenas indicators. Space was always provided within the questionnaire for participants to note these potential 'missing indicators'.

Despite the above limitations, this chapter contributes to existing literature on poverty measurement in two main ways.

Firstly, given the importance of participatory methods in poverty assessments, this chapter provides an overview of the Delphi as an alternative way of gathering opinions on the myriad dimensions and indicators of poverty, alongside their respective weights. This method aimed to mitigate the biases that may result from possible dominant individuals through ensuring a

degree of anonymity; albeit only quasi-anonymity, of its participants. Furthermore, the Delphi's 'mixed-method' approach, allows it to gather both 'qualitative' opinions and 'quantitative' data, thus enabling statistical assessment on whether a reliable level of consensus can be achieved with regard to the choices one needs to make when constructing multidimensional poverty measures.

Secondly, by documenting the detailed steps made within the application of the Delphi method within this thesis, this chapter aims to ensure the systematic and transparent reporting of the decision-making process behind the method's results. This proves important as following Amartya Sen (2004b) and Robeyns (2005b), 'it is not sufficient to argue that people can agree about the different items listed within a measure of poverty'; processes behind how these lists are drawn-up also matter.

The above in mind, the next Chapter, Chapter 4 explores the results of the Delphi study proposed within this thesis.

Chapter 4: Results of the Bogor City Delphi

We need more evidence-based policy to alleviate poverty, such as that offered by this Delphi exercise. Poverty is a multifaceted issue, which can only be eliminated through collaborative work between our different government sectors. This Delphi exercise is the first time we have all gathered together within one room to work on this issue.

Dr Bima Arya Sugiarto, Mayor of Bogor City, August 2016

4.1 Introduction

The previous chapter, Chapter 3, presented a thorough account of the methodological steps taken within the Delphi study proposed in this thesis. The Delphi described within Chapter 3 sought to shed light on the following research questions; results of which, are the subject of this current chapter.

Firstly, participants were asked the question: *How should poverty be defined?* In answering this question, participants were also encouraged to express their views on whether a 'dashboard' or 'composite' approach would be suitable to measure poverty within their district. Secondly, following the 'poverty definition' question, as discussed within sub-section 3.4.2.1 of Chapter 3, Delphi participants were requested to 'cluster' pre-existing indicators available within the Susenas in order to form possible 'dimensions' of poverty. This chapter aims to describe the results of this clustering exercise by addressing the question: *Using similarity scores assigned by Delphi participants, how should the socio-economic indicators within the Susenas be clustered into dimensions?*

As argued within sub-section 3.3.4 of Chapter 3, before examining the dimensions and indicators chosen by participants and how poverty should be measured, one needs to assess whether 'convergence' of these views was evident between the Delphi's iterative rounds. This 'convergence' of views is measured, in particular, by examining whether responses given by participants on dimensions and indicator weights tended to move toward the group's median. The third research question therefore investigates the question: *Do measures of 'stability' between the Delphi rounds indicate that a reliable level of consensus was achieved with regard to participants' views on dimensions and indicators weights?* Since as will be illustrated within this current chapter, 'stability' of responses across the three Delphi rounds was generally achieved, the fourth research question asks: *Which important dimensions and indicators should be included within a reliable measure of poverty?* Finally, based on Delphi participants' preferences on the degrees of importance of these components and the possible trade-offs between them: *How should these dimensions and indicators be weighed against each other?*

The rest of this chapter is organised as follows. Before delving into results, Section 4.2 provides a brief description of the demographic characteristics and perceived expertise of the participants included in the Delphi study. Section 4.3 presents results of the first research question; i.e. how participants defined poverty. Section 4.4 describes the outcomes of the 'clustering exercise'. Section 4.5 addresses the third research question; i.e. whether a level of consensus is achieved with regard to participants' choice of dimensions, indicators and weights. Section 4.6 analyses results of the fourth and fifth research questions; i.e. which dimensions and indicators are thought of as important to include within a multidimensional poverty measure, alongside how they should be weighed. Section 4.7 discusses results of robustness checks utilised to assess the findings within Section 4.6. Finally, Section 4.8 ends this chapter with conclusions drawn from its findings.

4.2 Description of Delphi Participants

Before describing the characteristics of Delphi participants, a note of clarification is needed. As was discussed within sub-section 3.4.2 of Chapter 3, the Delphi within this thesis was conducted in two days. All policy makers invited, attended the study; however, a small level of attrition was evident. Two participants from Day 1 could not continue to Day 2 and sent their deputies to take their place; the total number of participants on both days, thus, remained the same. Except for the results of the 'clustering' exercise, conducted on Day 1, descriptive analyses within this current chapter will primarily focus on responses given by participants who attended Day 2 of the Delphi. This decision was made as the study's main exercise of 'selecting' and 'weighting' dimensions and indicators through the Delphi's iterative rounds, was conducted on Day 2. In addition, as will be noted within the description below, the replacement of the two participants on Day 2, did not significantly alter the demographics of the study's participants.

4.2.1 Participants' Demographic Characteristics

To assess the fulfilment of 'expert' criteria stated within sub-section 3.4.1.4 of Chapter 3, following Scheibe et al. (2002), a demographic questionnaire was administered before the Delphi's iterative rounds in order to collect data on the age, gender, education and years of experience of participants. This questionnaire was given to participants on Day 1 and Day 2 of the Delphi. Given the decision to place primary focus on Day 2 participants, Table 4.1 presents a summary of demographic characteristics of these participants.

Table 4. 1 Demographic Characteristics of Delphi Participants [(a)]

Demographic Characteristic	Number of Participants (N)	Percentage of Participants (%)	Mean	Median	Min	Max
Gender						
Female	18	50				
Male	18	50				
Age			44	44	21	56
Education Level						
Elementary	0	0				
Secondary	0	0				
High School	5	14				
1 year or 2 year diploma	0	0				
3 year diploma	1	3				
Undergraduate	15	42				
Postgraduate masters level	13	36				
Doctorate level	2	6				
Total number of participants	**36**					
Years of experience (How long have you worked within the Bogor City Government?)			18	18	2	34

Note: [(a)] These characteristics are of participants from Day 2 of the Delphi

Source: Author's computations (Bogor City Delphi, 2016)

As can be seen within Table 4.1, a gender balance was achieved; i.e. out of a total of thirty-six participants included within the Delphi's iterative rounds, eighteen were female and eighteen were male. The mean and median age of the participants was 44 years; only three participants were in their twenties, with one aged twenty-one and two aged twenty-seven. Seventy-eight percent of the participants had an undergraduate or postgraduate (masters level) degree. A further six percent had PhDs. Participants' mean and median years of experience of working within the Bogor City government, was eighteen years.[1]

Although participants were aware that to be involved within the study they needed to possess substantial experience in the area of poverty alleviation, semi-structured interviews with younger participants; i.e. those in their early twenties and thirties, who had less than ten years of experience working as a policy maker, were conducted *after* the Delphi took place.[2,3] In addition to facilitating the robustness checks illustrated within sub-section 3.4.3 of Chapter 3, these interviews aimed to assess their suitability for the study, i.e. whether they had enough experience for their responses to be included within the Delphi's final conclusions. These interviews established that these younger participants had degrees specialising in social policy and development, that they worked closely with their heads of sectors and were responsible for delivering policies directly to the people, which led them to have substantial knowledge of the field. Their responses were thus of great value to the Delphi.[4]

4.2.2 Participants' Perceived Expertise

In addition to the above demographic characteristics, to further establish participants' level of 'expertise' and their 'commitment' to the study, as described within sub-section 3.4.2 of Chapter 3, information on various aspects of participants' personal confidence on the issues of 'defining and measuring poverty', were collected. Table 4.2 presents these results.

[1] Participants from Day 1 also exhibited this same gender balance, however, had a slightly higher average age of 45. Their average 'years of experience' of working within the district government was however, the same as participants included on Day 2, i.e. both averaging eighteen years of experience.

[2] There were five participants aged 20-35, who were involved within both days of the Delphi study. These semi-structured interviews were conducted with all these five participants.

[3] As it was important not to divulge critical information on the Delphi process before the study took place to avoid possible biases in how participants may have viewed the study, the decision to conduct these interviews after the Delphi took place, had to be made.

[4] These post-Delphi semi-structured interviews were also conducted with 'senior' participants, i.e. those with more than ten years of experience. Results from these interviews confirmed these participants' knowledge of poverty issues within their district, besides their passion for reform in measuring poverty and delivering sound policy to support poverty alleviation efforts.

Table 4. 2 Delphi Participants' Perceived Expertise[(a)] [(b)]

Statement 1 I have a definite idea of the needs of the people of my district			Statement 2 If compared to the other participants within this room, my skills on poverty alleviation policy are about here			Statement 3 I think my ideas are in essence, in agreement with the rest of the other participants			Statement 4 I know most of the people in this room very well			Statement 5 I am anticipating that the Delphi is going to be a good method to identify poverty dimensions and indicators for my district		
Likert Category	**N**	**%**	**Likert Category**	**N**	**%**	**Likert Category**	**N**	**%**	**Likert Category**	**N**	**%**	**Likert Category**	**N**	**%**
1	3	8	1	4	11	1	1	3	1	0	0	1	0	0
2	1	3	2	4	11	2	2	6	2	1	2.8	2	0	0
3	2	6	3	7	19	3	5	14	3	4	11.1	3	0	0
4	11	31	4	8	22	4	8	22	4	8	22.2	4	3	8
5	10	28	5	6	17	5	7	19	5	7	19.4	5	6	17
6	3	8	6	3	8	6	9	25	6	11	30.6	6	13	36
7	6	17	7	4	11	7	4	11	7	5	13.9	7	14	39

Note: [(a)] Likert scale categories:
"1": Totally disagree with statement
"7": Totally agree with statement
"4": Presumed neutral category
[(b)] Participants described here are from Day 2.

Source: Author's computations (Bogor City Delphi, 2016)

Following Scheibe et al. (2002), participants were asked to respond to five statements using a 7-category Likert scale to represent their opinions. Category “1” indicated ‘lowest level agreement’; “7” indicated ‘highest level of agreement’ and category “4” indicated a presumed ‘neutral’ level of agreement (Matell & Jacoby, 1971; Sullivan & Artino, 2013; Symonds, 1924).

From Table 4.2, referring to ‘Statement 1: Whether participants had a definite idea of poverty issues within their district’, 31 percent of participants selected the ‘neutral’ category, whilst 53 percent of participants selected categories “5” and above, indicating that most participants felt they had the necessary knowledge with regard to the possible welfare needs of the people within their district.

Examining ‘Statement 2: Participants’ perceived skills in poverty issues, compared to others involved within the study’, 36 percent of participants assigned categories of ‘5’ and above; i.e. they ‘perceived’ that they had a somewhat higher level of knowledge with regard to the issue of poverty, compared to other participants. This may indicate the possibility of ‘dominant’ voices if the exercise were conducted through a ‘discussion’ format. Additionally, this relatively ‘lower’ percentage of participants perceiving that they were ‘more knowledgeable’ may indicate that, in general, participants tended to respect the views of others included in the study. This reasoning is to an extent, supported by the results of the post-Delphi questionnaire examining whether participants valued the ‘feedback’ given between rounds, which will be described in Section 4.7 of this chapter.

To assess the possibility of achieving consensus within this group of participants, Statement 3: ‘whether participant’s ideas would tend to agree with others in the group’, found that 55 percent of participants responded with categories of ‘5 and above’, i.e. they tended to think that their views *would* coincide with other participants included in the Delphi.

As was described within sub-section 3.4.2 of Chapter 3, the Delphi within this thesis was conducted within the setting of one conference room. It was thus important to assess the degree of quasi-anonymity within this study; Statement 4, refers to this and found that 64 percent of participants assigned categories “5” and above, indicating that they knew other participants within the group. This result confirmed the crucial need to ensure that responses were not publicly linked to certain participants. Last but not least, to examine participants’ potential expectations of the Delphi before the study was conducted, ‘Statement 5’, shows that 92 percent of participants had relatively ‘high hope’ that the Delphi process would aid the construction of a reliable multidimensional poverty measure for their district.

In sum, the descriptive analyses above indicates that policy makers included within the Delphi, in general, met the criteria set within sub-section 3.4.1.2 of Chapter 3. Their demographic characteristics seemed to confirm their degree of 'experience' with regard to poverty related issues within their districts. Furthermore, if, following Keeney et al. (2010), 'level of education' was utilised as an 'inclusion criteria', participants within the Delphi were also shown to fulfil this specification as most were highly educated. These findings were further supported by results within the 'perceived expertise' questionnaire. On this, a majority of participants perceived that they had the necessary skills in recognising poverty issues within their district. In addition, a majority tended to acknowledge the possible utility of the Delphi as a way of facilitating consensus with regard to important aspects of multidimensional poverty.[1] These findings act to support the legitimacy of the Delphi results described within the following sections.

4.3 The Definition of Poverty

As Keeney et al. (2010) state, when presenting results of a Delphi process, 'analysis involves both qualitative and quantitative data' (p. 84). This section starts with an evaluation of qualitative responses of the open-ended question of 'how poverty should be defined in the case study area', thus, aiming to address the first research question within this chapter. [2]

The open-ended nature of this first question necessitated a method of analysis commonly utilised within qualitative research, namely 'content analysis'. Cole (1988) defines 'content analysis' as a 'method of analysing written, verbal or visual communication messages'. Keeney et al. (2010) state that the main aim of 'content analysis' is to 'group statements generated by participants into similar areas'; they further suggest that, when analysing Delphi results, 'a simple approach to content analysing usually works well' (p. 72). This in mind, two types of content analysis generally exist; 'deductive' and 'inductive'. For the purposes of this thesis, 'inductive analysis' proved to be beneficial, as, moving from 'specific' statements of poverty definitions, and then 'combining' them into 'larger or more general statements', needed to be done (Chinn & Kramer, 1983; Elo & Kyngäs, 2008).

[1] Results from Day 1 participants also came to the same conclusions.
[2] The question of 'how poverty is defined', was asked of participants at the start of Day 1 and Day 2 of the Delphi process. As described within Section 4.2 of this chapter, a degree of iteration was evident in that two participants from Day 1 were replaced on Day 2. This section presents results from Day 2. These results however, did not differ significantly from Day 1. The percentages of 'definitions' found to match each 'category' were overall found to be very similar. In particular, all participants on Day 1 also noted that poverty is multidimensional.

Within inductive analysis, Elo and Kyngäs (2008) suggest the use of 'coding or categories' in order to organise each response into groups. These 'categories' should 'not simply act as a means to bring together responses which are similar or related'; rather, they are utilised to 'classify answers as belonging to a particular group' (Dey, 2003, p. 102). When formulating categories, Dey (2003) notes that the exercise should be both 'conceptual' and 'empirical''; i.e. 'categories must relate to an appropriate analytical context and be rooted to a wider conceptual theory' (p. 102-103). This in mind, each 'definition of poverty' provided by Delphi participants was examined. Keywords from these answers were then identified in order to differentiate and group these responses into specific 'definition categories'.

The process of examining participants' 'definitions' started with translating each statement from Bahasa Indonesia, the language of instruction for the Delphi questionnaire, to English. Each statement, alongside its translated version, was inputted into an excel spreadsheet to allow for 'cutting and pasting', thus easing the process of identifying 'keywords' (Keeney et al., 2010, p. 85). Keeney et al. (2010) note that wording for these keywords, should be kept as true as possible to the statements provided by participants of the study. On this, Elo and Kyngäs (2008) suggest that 'content-characteristic words' derived from participants' statements may act as keywords. These empirically derived 'keywords' should then be matched with widely applied theoretical definitions of poverty to form more general 'categories'. Table 4.3 lists these 'keywords', the 'categories' they correspond to, alongside definitions of these 'categories'. [1]

As shown within Table 4.3, five 'categories' were identified based on existing theories on poverty definitions and measurement. These five categories are, (1) the 'income' poverty line approach to defining poverty, (2) the 'Basic Needs' approach, (3) Townsend (1979a)'s idea of relative poverty, (4) Sen's Capability Approach and (5) Guio, Gordon, and Marlier (2012b)'s idea of 'material deprivation' as a key aspect of poverty. Chapter 1 offered a general discussion on these theoretical approaches; summaries of what each theory entails are however presented within Table 4.3. After grouping participants' 'definition statements' into one or more of these five 'theory categories', these statements were also grouped into two 'broader' categories. These 'broader' categories were based on whether the statements point toward a 'uni' or 'multidimensional' understanding of poverty. A statement is categorised as referring to a 'multidimensional' definition of poverty, if more than one indicator or type of deprivation is mentioned as a characteristic of experiencing poverty.

[1] A full list of participants' 'definition statements' is available within the Appendix of this chapter.

Table 4. 3 Keywords and Categories Used to Group Delphi Participants' Definitions of Poverty[a]

Keywords[b]	Category[c]	Theoretical Definition	Theorist
Income	Income poverty line	Poverty is having income below a certain threshold; having income below a poverty line.	Lydall (1968);Pigou (1929); D. M. Gordon (1972); Anthony Barnes Atkinson, Rainwater, and Smeeding (1994)
Participation; customary; according to what is valued in the community; standard of living considered adequate; fitting in to society; activities necessary in society; ostracized by society	Townsend's Idea of Relative Poverty	Individuals, families and groups in the population can be said to be in poverty when they lack the resources to obtain the types of diet, participate in the activities and have the living conditions and amenities which are customary, or are at least widely encouraged or approved, in the societies to which they belong. Deprivation is the lack of socially-perceived necessities	Townsend (1979)
Basic needs; daily needs	Basic needs	Poverty is the lack of satisfaction of the elementary needs of a population, especially in education and health. Focusing on income only is not enough.	Streeten (1994); Streeten et al. (1981)
Capability, ability, freedom	Capability approach	Poverty is when a person does not have the freedom to actually do or be the things that he or she may value doing or being .	Amartya Sen (1976a); Sen (2001)
Material; material deprivation	Material deprivation	Material deprivation is defined as the enforced lack of a combination of items depicting material living conditions, such as housing conditions, possession of durables, capacity to afford basic requirements.	Guio, Gordon, and Marlier (2012a)

Note: [a] Results within this table are based on participants responses on Day 2; [b] Keywords were derived from empirical statements made by Delphi participants; [c] Categories were derived from existing theoretical definitions of poverty.

Source: Author's summary (Bogor City, 2016)

Table 4. 4 Summary of Delphi Poverty Definitions by 'Theoretical' and 'Uni or Multidimensional' Categories[(a)]

Category	N	%
Theoretical Category		
Capability approach	21	58
Basic needs	17	47
Relative poverty	11	31
Income poverty	9	25
Material deprivation	3	8
Broader Uni or Multidimensional Category		
Multi-dimensional	36	100
Uni-dimensional	0	0

Note:
[(a)] This table presents results from Day 2. Results from Day 1, however did not differ significantly.
[(b)] The sum total of 'Number of Participants' choosing each 'Theoretical Category' does not equal to thirty-six, as it was possible for a participant's definition to match more than one 'Theoretical Category'.
[(c)] The denominator for these percentage values is the total number of participants (thirty-six). The sum total of these percentages does not equal 100% for the reason within note (b) above.

Source: Author's computations (Bogor City, 2016)

With the above in mind, Table 4.4 presents a summary of the number and percentage of participants whose 'definition of poverty' corresponds to each 'theoretical' and 'uni or multidimensional' categories. Examining the broader 'uni or multidimensional' categories within Table 4.4, it is shown that *all* participants (100 percent) perceived 'poverty' as a 'multidimensional' phenomenon concerning deprivations within multiple aspects of life. Delving deeper into the 'theoretical categories', a majority of participants' definitions mirrored the 'Capability Approach'; i.e. around 58 percent of participants included keywords pointing toward this theoretical understanding of poverty within their responses and acknowledged that poverty should be conceptualized and measured in terms of people's 'freedoms' and 'capabilities to function' (Robeyns, 2005a, p. 95).

Keywords connected to the 'basic needs' approach were found within 47 percent of the statements. These statements were shown to echo Streeten et al. (1981)'s suggestion that measuring poverty using 'income' alone is not enough; one also needs to focus on satisfying the 'elementary needs of society' (p. 25).

Keywords relating to Townsend's idea of 'relative poverty' were found within 31 percent of the statements. These statements described poverty as occurring when a person 'lacks the resources to participate in the activities and maintain living conditions that are customary within the society in which they live' (Townsend, 1979a).

The importance of income as a way of measuring poverty was acknowledged within 25 percent of the participants statements. Last but not least, Guio et al. (2012a)'s use of 'material deprivation' when describing poverty, was only discovered within 8 percent of the statements. The lack of popularity of this approach to defining poverty, may be due to the fact that the idea of 'material deprivation' has been more widely operationalised within the context of developed rather than developing nations (Guio, Fusco, & Marlier, 2009; Guio et al., 2012a; Guio et al., 2016; Whelan & Maître, 2007).

In sum, considering participants' definitions, all statements included multiple aspects of poverty; none considered poverty as a uni-dimensional phenomenon. Measurement should thus, reflect this multifaceted nature of poverty. When examining these definitions in further detail however, differences were found with regard to the specific aspects of poverty that matter most; i.e. whether poverty should be understood as lack of 'freedoms and capabilities', un-fulfilment of 'basic needs', as 'exclusion from societal standards', as a lack of 'income' or as suffering from numerous 'material deprivations'. If, however, a majority was to be defined as 'a category chosen by more than fifty percent' of the participants, the Capability Approach (CA), seemed to be regarded as the most relevant definition. As noted within Chapter 1, the formulation of the CA acted as a 'tipping point', sparking significant debate on the need to recognise and measure poverty through a multidimensional framework. The importance of this approach seems to also resonate within the definitions of poverty stated by Delphi participants. Even though this is the case, unanimity on the CA's importance was not achieved. This existence of 'disagreements' within participants' statements, strengthens the need for a mechanism of achieving at least some level of consensus with regard to the dimensions and indicators of poverty, which may best describe these definitions.

4.3.1 Dashboard or Composite Measure

As summarised above, given the importance of employing a multidimensional framework, participants were asked to choose whether an 'aggregate' or 'dashboard' measure would be the most effective tool to monitor poverty in Bogor City.[1,2] This question also provided an open-ended text box, where participants could state reasons for their choice.

Results show that participants unanimously opted for the use of a 'dashboard' measure. The importance of a 'composite' measure was however, also acknowledged as a 'supplement' to

[1] This question was asked of participants on Day 1 and Day 2 of the Delphi. Both Days indicate the same result; i.e. a dashboard measure was unanimously found to be favourable. As this is the case, this section describes results from Day 2 for reasons stated within Section 4.2 of this chapter.

[2] Section 5.2 of Chapter 5 discusses the key arguments within this composite vs. dashboard debate.

the dashboard. This reasoning was evident within the open-ended responses provided by participants explaining their opinions on this issue. Nevertheless, out of thirty-six, only six participants provided an explanation for their choices. These were:

> *'Poverty comprises multiple areas of interest; as a policy maker, I need to monitor progress for each of the priority areas set by the government. An aggregate index is also needed but as a supplement'*
>
> *'I understand the merits of one index to monitor progress, but how do we determine this one index. Setting the cutoff is complicated. Furthermore, the government could benefit from more than one cutoff, indicating not only 'poverty', but also, those who are 'vulnerable to poverty' and those who are in 'extreme poverty'. Using these different cutoffs, we can implement aggregate measures to monitor poverty across our sub-districts'*
>
> *Look at all these people in the room. They are representatives of different sector areas in the government. We need to transparently know how each of our sectors is progressing, in order to identify key areas where more work is needed. A composite measure is then needed when we aim to monitor progress across all dimensions'*
>
> *'We need an index that comprises dimensions that the people of Bogor City value. These dimensions are multiple. Monitoring them with a dashboard index would be beneficial'*
>
> *'Dashboards provide important information for us. But, I understand that examining whether households experience multiple deprivations, is also important'*
>
> *'We need to know how we are doing in each important dimension; health, education, for example. Dashboards present a clear way of doing this. Knowing whether households experience more than one short-coming at the same time, is however, also needed'*

The above statements confirm the importance of the 'dashboard' approach, whilst noting key uses of 'composite' measures. From these statements, 'dashboards' are recognised as providing the ability to 'monitor progress within important dimensions and indicators of poverty separately', whilst 'composites' enable the examination of 'progress across all these dimensions and indicators', alongside providing an analysis of 'households experiencing multiple short-comings'.[1]

The above results in mind, the next step within this part of the questionnaire required those participants who chose the 'dashboard approach' to respond to the additional question of: *how many dimensions of poverty (groups of indicators) should be included within a dashboard measure relevant for your district.* Table 4.5 illustrates the results of this question.

[1] Results from the post-Delphi robustness interviews explored within Section 4.7 of this chapter confirm these findings.

Table 4. 5 Number of Dimensions within Dashboard[(a)]

Number of Dimensions	
Min	3
Max	15
Mean	6.6
Median	5
Mode	5
SD	3.9

Note: [(a)] This table presents results from Day 2 of the Delphi
Source: Author's computations (Bogor City, 2016)

As shown, participants suggested a minimum of three dimensions and a maximum of fifteen.[1] Thus, as this substantial spread between the 'minimum' and 'maximum' number of dimensions suggested, was evident, the 'median' was taken to represent the 'optimum' number of dimensions to be included within the proposed dashboard. [2]

As a final note, since all participants chose the 'dashboard' approach, none provided responses with regard to the 'poverty cutoff' question; i.e. how many indicators a household should be deprived in, in order to be categorised as 'poor'. Post-Delphi robustness check interviews, discussed within Section 4.7 of this chapter, confirm this 'unwillingness' of defining a 'poverty cutoff'.

These interviews found that with regard to 'poverty cutoffs', being able to set multiple 'cutoffs' measuring the amount of those who are 'vulnerable to poverty' and those in 'extreme poverty', would be more beneficial than setting an absolute threshold. In addition, being able to observe the total distribution of multiple deprivation across households was found to be crucial as it would provide an indication as to the level of inequality within their district (Bogor City Delphi Non-Participants, 2016; Bogor City Delphi Participants, 2016).[3] These findings in mind, the next section presents results of the Delphi's 'clustering indicators into dimensions' exercise.

[1] Results within Table 4.5 are from Day 2. As a comparison, participants on Day 1 also chose a median of 5 dimensions. A difference was, however observed as to the maximum number of dimensions to consider; on Day 1 a maximum value of 14 dimensions was found.

[2] Merits of the use of the 'median' within Delphi studies is discussed within sub-section 3.3.4 of Chapter 3. In addition, results from the post-Delphi robustness interviews, examined in Section 4.7 of this chapter, confirm agreement as to the decision to group indicators into five dimensions.

[3] This issue is discussed further within Chapters 5 and 6 of this thesis.

4.4 Clustering Indicators into Dimensions

This section addresses the second research question within this chapter, i.e. how pre-existing indicators of poverty available within the Susenas were clustered into dimensions.

As illustrated within sub-section 3.4.2.1 of Chapter 3, the 'clustering of indicators into dimensions' exercise was conducted on Day 1 of the Delphi; before the iterative rounds were conducted on Day 2. Participants were required to 'cluster' twenty-eight pre-existing indicators into 'groups' of dimensions by assigning 'similarity scores' ranging from "0" indicating complete similarity and "4" indicating complete un-relation.[1]

To analyse results from this 'clustering exercise', following the method developed within Norman Crolee Dalkey et al. (1969), Norman C Dalkey and Rourke (1971) and Norman C Dalkey (2002), the 'mean' of the 'absolute value of the similarity ratings across all participants' was calculated. These 'mean absolute ratings' were then transformed into a group 'dissimilarity matrix' and were processed with an agglomerative hierarchical clustering routine. Within this routine, each indicator is 'initially considered as a cluster of its own; the most similar clusters are then successively merged until there is just one single big cluster' (Kassambara, 2017, p. 42). Analysis of this 'dissimilarity' matrix was undertaken using the hclust package within R.[2]

Results of the above hierarchical clustering routine are commonly presented within a tree-based diagram, known as a 'dendrogram'. Within this diagram, a multilevel hierarchy of clusters is presented, where 'clusters at one level are joined together to form larger clusters at the next level' (Kassambara, 2017, p. 42). Langfelder et al. (2008) state that 'clusters' are 'defined by cutting branches off the dendrogram'. A decision thus needs to be made with regard to the 'level' at which to 'cut the tree' in order to generate the suitable number of clusters. Norman Crolee Dalkey et al. (1969) note that, 'the number of clusters should be compatible with the data and or any theoretical or empirical considerations within the procedure' (p. 24). This in mind, within this chapter, this decision was based on the results within sub-section 4.3.1 above; i.e. that the 'optimum' number of clusters should be equal to the 'median' number of dimensions stated by Delphi participants. As the median was found

[1] To note, participants were reminded of the definitions of these indicators within the Susenas before conducting this 'similarity' exercise.

[2] In particular, guidelines within Langfelder, Zhang, and Horvath (2008) and Galili (2015) were followed.

to be five, the dendrogram within this chapter was 'cut' at the point where the hierarchical clustering solution offered *five* dimensional clusters of indicators.[1]

Following the above decision, in hierarchical clustering, an additional consideration needs to be made as to how to define 'distance or similarity levels' between indicators 'within' and 'between' dimensional clusters. Everitt (2011) and Hamilton (2012) provide a detailed summary of possible methods aiming to address this issue. They particularly explore the implications of applying 'single linkage', 'complete', 'average' and 'ward linkage' methods, on the structure of the clusters.[2] Within this chapter, the use of all these alternative methods was tested and the resulting clusters were examined. It was found that all four methods yielded almost homogenous clusters; i.e. the 'single', 'complete', 'average' and 'ward' linkage methods tended to cluster similar indicators into the same five dimensions. Nevertheless, subtle differences were still found.[3,4]

A choice thus needed to be made with regard to which method to utilise. In order to arrive at this decision, the validity and reliability of the five clusters formed using the four methods, were tested through a series of 'cluster validation' techniques suggested within Hamilton (2012) and Kassambara (2017).[5] These techniques concluded that, compared to the other methods, the 'complete' clustering method, formed the most 'well' clustered dendrogram.

---

[1] Langfelder et al. (2008) note that within R's hclust package, one may choose to 'cut' the dendrogram into 'groups' by either specifying the desired 'number' of groups or indicating the 'cut height' at which to 'cut' branches of the diagram. Within this chapter, the 'desired number of groups'; i.e. five, was specified.

[2] To briefly summarise, within 'single linkage' clustering, the similarity between two clusters is defined by the similarity of their most similar members. The 'single linkage' merge criterion is thus 'local'. Schütze, Manning, and Raghavan (2008) note that in 'single linkage' clustering, 'attention is only paid to the area where two clusters are joined together; other more distant parts of the clusters and the clusters' overall structure are not taken into account'. The opposite is true for 'complete linkage'; where 'inter-group distance is defined as the largest distance between any two individuals, one from each group' (Everitt, 2011). Everitt (2011) notes that an intermediate option between the use of 'single' and 'complete' clustering is the use of 'group average clustering', where the 'inter-group dissimilarity is defined as the average dissimilarity between individuals from both groups' (p. 61). As for the 'ward method', it relies on 'minimum variance' to determine distance between clusters; i.e. the distance between clusters is the error sum of squares (SSE) between two clusters over all of the variables (Ward, 1963). When utilising the 'ward' method, 'at each stage of the clustering procedure, the within SSE is minimised over all partitions obtained by combining two clusters from the previous stage' ("Forming Clusters," 2020).

[3] Referring to the dendrogram within Figure 4.1, which utilises the 'complete' method, the 'average' method grouped one 'Health' indicator into the 'Living Standards' dimension, rather than including it into the 'Health' dimension. Another example was that, compared to 'complete' clustering, the 'ward' method grouped 'child mortality' with the 'Employment and Asset' indicators; rather than including it within the 'Birth attendant and contraception' indicators.

[4] As an alternative, Cronbach's alphas were calculated to assess the statistical reliability of the five dimensions. Relatively low alphas were found with an average of 0.2. Even though this is the case, to stay true to the decisions and consensus achieved through the Delphi, results from this hierarchical clustering exercise are still utilised to define the five dimensions.

[5] Results of these tests are presented within the Appendix of this chapter.

This in mind, the 'complete' method was utilised to form the 'final' set of dimensions resulting from the Delphi's 'clustering exercise'. Figure 4.1 presents the dendrogram alongside the five-dimensional clusters which were formed.

Within Figure 4.1, X1-X28 represent the 28 pre-existing socio-economic characteristics available within the Susenas, which were included as possible 'indicators' of poverty in the Delphi's iterative questionnaire.[1] The red boxes were generated by the hclust package, to represent the five dimensional clusters resulting from the 'complete' method. Following Keeney et al. (2010), these five clusters were assigned labels, which were developed in the hopes of being 'true to the wording of the indicators within those dimensions' (p. 85). Considering the 'indicators' within them, the five dimensional labels were thus: (1) Employment and Asset Ownership, (2) Health and Safety, (3) Education, (4) Child Health and Contraception and (5) Living Standards.[2] These five dimensions then were inputted into the iterative questionnaire implemented on Day 2.

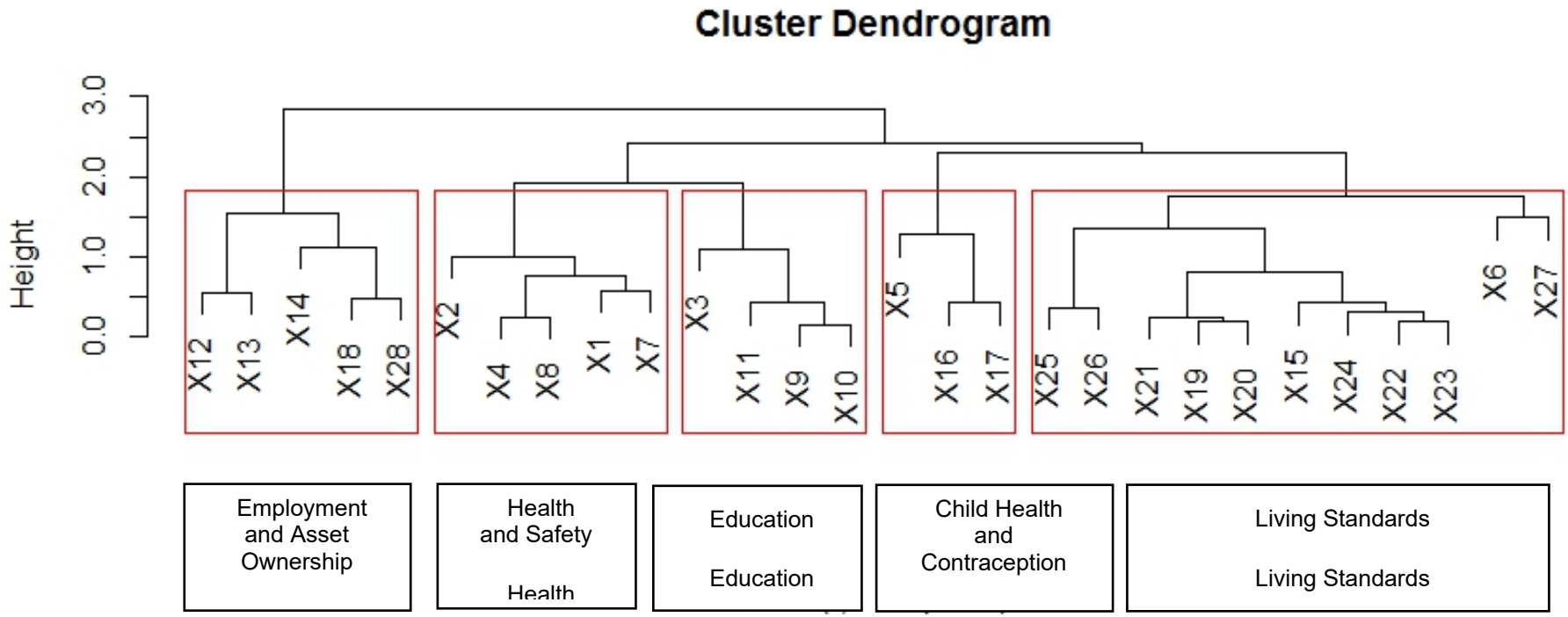
Following the steps discussed within sub-section 3.4.2.2 of Chapter 3, on Day 2, participants were requested to firstly 'select', out of these five possible dimensions, which were thought to be of particular importance, to include within a measure of multidimensional poverty. Participants could select any number of dimensions. They were then asked to select relevant *indicators,* which may best measure deprivation within these 'selected' dimensions. Finally, participants' views on how these 'dimensions' and 'indicators' should be *weighed*, were solicited. Participants gave these opinions and had the chance to review them through the Delphi's three iterative rounds.

Before examining participants' choice of dimensions, indicators and weights, the next section first assesses whether 'stability' was evident with regard to these views.

[1] For the purposes of this chapter, wording and labels for the indicators were translated from Bahasa Indonesia into English and simplified into three to five words.

[2] As a note of clarification, within the fifth cluster, i.e. 'Living Standards', referring to Figure 4.1, at a glance, it seems that the indicators X6 and X27 form a separate sixth cluster. The hclust package, however, joined this 'sixth' cluster with the cluster directly beside it; i.e. that containing the indicators: X25, X26, X19, X20, X15, X24, X22 and X23, as they are both connected by a 'higher' branch above them. These two 'lower' clusters form the fifth 'Living Standards' cluster, as the dendrogram was 'cut' at the 'higher' branch above these two clusters. The height at which this 'higher' branch was cut, corresponds to the heights where the four other dimensional clusters are cut.

Figure 4. 1 Bogor City Delphi Cluster Dendrogram[(a)(b)]



Note:

[(a)] X1-X28: Indicators of poverty:

X1: Safe Living Environment	X7: Immunization and Vaccinations	X13: Employment	X19: Roof	X25: Electricity
X2: Ability to Travel	X8: Children to be Breast Fed	X14: Credit	X20: Wall	X26: Cooking fuel
X3: Ownership of Birth Certificate	X9: Literacy	X15: Household Consumption	X21: Floor	X27: Gov. help
X4: Healthy life	X10: Children's education	X16: Child mortality	X22: Drinking water	X28:AssetOwnership
X5: Modern Health Facilities (for birth)	X11: Adult education	X17: Contraceptives	X23 Washing and Cooking Water	
X6: Traditional Health Facilities (for birth)	X12: Access to the internet	X18: House or Land Ownership	X24: Sanitation	

[(b)] The red boxes were generated by R to indicate the five dimensional clusters.

[(c)] These results are from Day 1 of the Delphi.

Source: Author's computations (Bogor City, 2016)

4.5 Stability and Consensus

This section addresses the third research question within this chapter: can consensus be achieved with regard to which dimensions, indicators and weights, to include within an empirical measure of multidimensional poverty?

As a brief review, after selecting dimensions and indicators, participants were requested to assign weights to these chosen components. Values for these weights ranged from “0” to “1”; with “0” indicating that a specific dimension or indicator is irrelevant, thus, should be excluded from measurement, and “1” specifying that the said component is of the utmost importance. The sum of dimensional-weights had to equal “1”; likewise, the sum of indicator-weights within each dimension also had to equal “1”.

Dimensions and indicators, which were not selected by participants were thus, given a weight of “0”. If participants only selected one dimension or one indicator as important, this component was given a weight of “1” and all other available dimensions and indicators were given a weight of “0”. The weights chosen by participants, thus, ultimately determined whether a dimension or indicator, should be included within the final empirical measure. Considering this, the measurement of ‘consensus’ with regard to the choice of dimensions and indicators, within this chapter, is based on measures of the statistical dispersion of the *weights* chosen by participants across the three Delphi rounds. Table 4.6 illustrates these results.

Following the steps set within sub-section 3.4.1.4 of Chapter 3, to analyse whether consensus is likely, the ‘group stability’ of responses is analysed. Firstly, the standard deviations (SDs) of opinions on weights across the three rounds are examined. On dimensional-weights, Table 4.6 shows that, for all but one dimension, i.e. ‘Child Health and Contraception’, standard deviations *decreased* from Round 1 to Round 2. This *may* indicate that after receiving statistical summaries of the group’s response from Round 1 within their feedback sheets, views on dimensional weights within Round 2, started to converge toward the ‘group mean’. From Round 2 to Round 3, however, except for views on dimensional weights for ‘Employment and Asset Ownership’ and ‘Child Health and Contraception’, standard deviations are shown to *increase*. This illustrates a degree of ‘divergence’ from the ‘group mean’, after feedback from Round 2 was presented to participants. Even though this was the case, when examining responses from the first and final round, standard deviations are still shown to have decreased; except for the ‘Child Health and Contraception’ dimension, in which the values for standard deviations did not change from Round 1 to Round 3.

Table 4. 6 Measures of Dispersion of Weights Given to Dimensions and Indicators Across Delphi Rounds

Dimensions and Indicators	Round 1		Round 2			Round 3					
	SD (a)	MAD (b)	SD (a)	MAD (b)	CV (c)	SD(a)	MAD (b)	CV(c)	$\lvert CV_{R2} - CV_{R3} \rvert$ (d)	SD Trend (e)	MAD Trend(f)
1. Education	**0.31**	**0.13**	**0.15**	**0.07**	**0.44**	**0.19**	**0.05**	**0.67**	**0.23***		
Children's Education	0.33	0.25	0.18	0.15	0.45	0.22	0.10	0.62	0.17*		
Literacy	0.17	0.20	0.14	0.1	0.5	0.20	0.05	0.75	0.25*		
Adult Education	0.24	0.08	0.19	0.1	1.14	0.19	0.15	1.25	0.11*		
Ownership of Birth Certificate	0.12	0.00	0.13	0.1	0.82	0.13	0.10	1.08	0.26*		
2. Employment and Asset Ownership	**0.27**	**0.18**	**0.14**	**0.1**	**0.69**	**0.13**	**0.05**	**0.71**	**0.02***		
Stable Employment	0.37	0.10	0.25	0.20	0.72	0.22	0.18	0.69	0.02*		
House/Land Ownership	0.20	0	0.13	0.15	0.92	0.16	0.15	1.07	0.15*		
Asset Ownership	0.19	0	0.15	0.1	0.92	0.16	0.16	1.11	0.19*		
Access to the Internet	0.07	0	0.12	0.1	1.14	0.10	0	1.41	0.27*		
Access to Credit	0.11	0	0.11	0	1.47	0.09	0	1.34	0.13*		
3. Living Standards	**0.21**	**0**	**0.13**	**0.1**	**0.83**	**0.18**	**0**	**1.13**	**0.29***		
Household Consumption	0.12	0	0.19	0.1	1.37	0.25	0.10	1.34	0.03*		
Government Help and Subsidies	0.18	0	0.16	0.08	1.46	0.15	0.09	1.3	0.16*		
Sanitation	0.14	0	0.08	0.05	1.21	0.07	0.05	1.12	0.09*		
Drinking Water	0.08	0	0.09	0.06	1.27	0.07	0.03	1.22	0.05*		
Cooking and Washing Water	0.07	0	0.09	0.1	1.05	0.09	0.03	1.39	0.33*		
Access to Traditional Health Facilities	0.09	0	0.04	0	1.6	0.05	0	2.26	0.66		
Electricity	0.07	0	0.08	0.05	1.22	0.07	0	1.45	0.23*		
Cooking Fuel	0.06	0	0.07	0	1.34	0.07	0	1.43	0.09*		
Floor	0.06	0	0.07	0.05	1.2	0.05	0	1.27	0.08*		
Roof	0.07	0	0.07	0	1.34	0.05	0	1.38	0.04*		
Wall	0.06	0	0.06	0.05	1.11	0.05	0	1.24	0.14*		

Table 4.6 (continued)

Dimensions and Indicators	Round 1		Round 2			Round 3					
	SD[(a)]	MAD[(b)]	SD[(a)]	MAD[(b)]	CV[(c)]	SD (a)	MAD[(b)]	CV (c)	$\lvert CV_{R2} - CV_{R3} \rvert$[(d)]	SD Trend[(e)]	MAD Trend[(f)]
4. Health and Safety	**0.26**	**0.20**	**0.15**	**0.10**	**0.72**	**0.23**	**0.10**	**0.84**	**0.12***		
Healthy life	0.25	0.15	0.19	0.05	0.72	0.22	0.10	0.72	0.00*		
Safe Living Environment	0.27	0.23	0.19	0.10	0.75	0.21	0.10	0.74	0.01*		
Access to Immunization	0.11	0	0.11	0.10	1.03	0.10	0.05	1.16	0.13*		
Ability to Travel (for health)	0.14	0	0.09	0	1.4	0.12	0	1.45	0.04*		
Children to be Breast Fed	0.13	0	0.14	0	1.26	0.12	0	1.21	0.05*		
5. Child Health and Contraception	**0.11**	**0**	**0.12**	**0.10**	**1.13**	**0.11**	**0.10**	**0.97**	**0.16***		
Birth Attendant (Modern Health Facilities for Child Birth)	0.31	0	0.25	0.25	0.99	0.29	0.30	0.97	0.02*		
Contraceptives	0.24	0	0.17	0.15	1.07	0.17	0.20	1.04	0.03*		
Child Mortality	0.14	0	0.17	0.20	1.11	0.16	0.10	1.08	0.04*		

Note:
These results are from the iterative questionnaires given to participants on Day 2.

[(a)] SD: Standard Deviation from the mean.
[(b)] MAD: Median Absolute Deviation from the Median
[(c)] CV: Coefficient of Variance
[(d)] Absolute Value of the difference between CV from Round-2 and the CV from Round-3.
[(e)] Standard Deviation trend from Round 1, to 2, to 3.
[(f)] MAD trend from Round 1, to 2, to 3.
* $\lvert CV_{R2} - CV_{R3} \rvert < 0.5$

Source: Author's computations (Bogor City, 2016)

These results seem to suggest that between the start and end of the Delphi study, views on dimensional weights, which aim to represent the 'degrees of importance and possible trade-off levels' placed by participants on these dimensions, started to converge toward the mean, however, within the last round disagreement was still evident as indicated by the slight increases in the standard deviations. A similar finding is evident for indicator weights, where standard deviations for the most part, are shown to decrease from Round 1 to Round 3, however, slight increases emerged from Round 2 to Round 3.

The next statistic recorded within Table 4.6 is the Median Average Deviation from the Median (MAD), which provides an indication as to whether participants' views on weights tended to converge toward the median values of the group. As argued for within Chapter 3, this measure may act as a more reliable criterion for 'convergence' as participants were reminded of the importance of the 'median' as a measure of 'group response', within their feedback sheets. As can be seen within the last column of Table 4.6, the MAD for all dimensional weights, except for the 'Child Health and Contraception' dimension, is shown to decline between Rounds 1, 2 and 3, decline between Rounds 1 and 3 or decline between Rounds 2 and 3. This indicates a level of convergence toward the median with regard to these dimensional weights.[1] Following von der Gracht (2012), this decline in statistical variance points toward a level of stability of Delphi responses; which, from these results, can somewhat be confirmed for at least four, out of the five dimensions.

When examining MADs for indicator-weights, for most indicators, a similar conclusion to that found when examining dimensional weights is evident; i.e. MADs tended to decline from Round 1 to 3 or Round 2 to 3. This also points toward a level of convergence of views with regard to 'most' of these indicator weights. For some indicators, however, 'stability' of opinion was less likely. As examples are the views on weights for: the 'Ownership of Birth Certificate', 'Employment', 'Government Help and Subsidies', 'Sanitation', 'Birth Attendant (modern health facilities for childbirth)' and 'Contraceptives' indicators, in which their MADs are shown to increase from Round 1 to Round 2 to Round 3.

In addition to the above, for some indicators, conclusions made when standard deviations are examined differ from those when the MADs are analysed. To illustrate, for 'Adult Education', standard deviations tended to decline from Round 1 to 2 to 3; however, its MAD is shown to increase from Round 1 to 2 to 3. This in mind, as noted within Chapter 3, the examination of

[1] A similar conclusion is found If MAD-M (Mean Absolute Deviation from the Median) is used.

trends in ‘deviation’, is, however, only the first step to analysing the ‘stability’ of views within a Delphi. Table 3.1 within Chapter 3, lists additional measures, which hope to further assess the ‘stability’ between the three Delphi rounds. I start with the Coefficient of Variation (CV).

Following the guidelines proposed within Table 3.1 of Chapter 3, to examine ‘stability’, possible changes in the CV between two successive rounds are studied (Dajani et al., 1979; Shah & Kalaian, 2009; Yang, 2003). An ‘absolute difference of 0.5 and below’, between the CV of Round 2 and Round 3, indicates a ‘good degree of stability’, which in turn points toward ‘achievement of a level of consensus’ and ‘no need for additional rounds’ (Dajani et al., 1979; English & Kernan, 1976). As can be seen within Table 4.6, the absolute value of the difference between the CV of Round 2 and the CV of Round 3, for all dimensional weights is shown to have a value below 0.5. This is also true for all, but one indicator weight; i.e. the ‘Access to Traditional Health Facilities’ indicator. For all dimensional weights and most indicator weights, using this ‘CV criterion’ as an indicator of ‘stability’, thus demonstrates movement of participants’ views toward a level of consensus.

In sum, through the use of the above measures, participants’ views on the degrees of importance of particular dimensions and indicators, as represented by their weights, for the most part, tended to converge toward the mean and median values of the group. This is shown by declines in standard deviations from Round 1 to Round 3, alongside declines in MADs from Round 1 to Round 3. This finding is strengthened when examining the CV between the Delphi rounds, which for almost all dimensions and indicators, confirmed the unnecessary need for extending the Delphi’s iterative process to beyond three rounds. Given this observed level of ‘stability’, as argued for within sub-section 3.3.4 of Chapter 3, the results of the ‘final’ Delphi round may be considered as representative of overall ‘group judgment’. Even though this is the case, looking more closely, particularly at the standard deviations of participants’ views in Round 3, a level of disagreement on the ‘exact value’ of these weights was still evident.

4.6 Dimensions, Indicators and Weights

This section aims to tackle research questions two and three; i.e. which types of dimensions and indicators were selected by Delphi participants and how should these components be weighed against one another?

As previously noted, weights were utilised to represent the degrees of importance of each dimension and indicator, alongside possible trade-offs between them. This in mind, considering the degree of stability of responses found within Section 4.5, the median weights from the final

Delphi round were utilised to determine whether to include or exclude a particular dimension or indicator from the 'new' Delphi-based multidimensional measures which will be described within Chapters 5 and 6. A 'non-zero' median specifies the inclusion of a dimension or indicator, whereas a median weight of 'zero', designates the exclusion of that particular dimension or indicator. Table 4.7 illustrates these results and suggests three main points.

Firstly, given their 'non-zero' median weights, participants considered all dimensions as important to include within the 'new' measures.[1] In addition to this, as described within Table 4.7, a hierarchy of importance can be observed when examining the 'median' weights for these dimensions. This hierarchy is to a degree, preserved, if the 'mean' of the weights and the number of participants who voted for the dimension (% vote), are used. To illustrate, 'Education' was assigned the highest degree of importance, compared to the other dimensions, regardless of whether the 'median weight', 'mean weight' or 'percent votes' are examined. Referring to this 'hierarchy of importance' as measured by median weights, the degree of importance of 'Education' is followed by 'Health and Safety', 'Employment and Asset Ownership' and 'Living Standards' which were both assigned the same median weight of 0.20; and of least importance compared to the other four dimensions, 'Child Health and Contraception'.

Secondly, contra to the results for dimensions, not all indicators were considered important. Referring to indicator weights within Table 4.7, ten out of the twenty-eight indicators were assigned median weights of zero. To proxy deprivation within the 'Education' dimension, all four indicators were chosen as relevant. For 'Employment and Asset Ownership', three out of five indicators were chosen. Within 'Living Standards', five out of eleven indicators were chosen. Three out of five indicators were chosen within the 'Health and Safety' dimension. Finally, all three indicators within the 'Child Health and Contraception' dimension were considered as important (this is even though the median weight for this dimension was considered as the lowest out of the five dimensions).

[1] As stated within sub-section 3.4.1.2 of Chapter 3, to mitigate possible biases resulting from using a pre-specified list of indicators, participants were given the chance to note 'missing dimensions or indicators'. No participant reported any missing dimensions. On indicators, two 'missing' indicators were mentioned by five participants; i.e. (1) ability to practice religion freely and (2) access to recreation facilities. As the Susenas does not include these indicators, this constraint can only be mitigated through the collection of additional data. Given this challenge and noting that a relatively small percentage of participants mentioned these indicators as missing, focus within this chapter and the following chapters is placed on what Robeyns (2005b) dubs as a 'feasible measure' as opposed to an 'ideal measure' of multidimensional poverty, which would incorporate all possible relevant indicators.

Table 4. 7 Descriptive Statistics of Weights per Dimension and Indicator[(a) (b) (c)]

Dimension and Indicators	Number of Votes	% Vote	Mean Weight	Median Weight
1. Education	**32**	**89%**	**0.28**	**0.25**
Children's Education	30	83%	0.35	0.30
Literacy	29	81%	0.27	0.25
Adult Education	21	58%	0.15	0.15
Ownership of Birth Certificate	19	53%	0.12	0.10
Total Education Indicator Weights			0.89	0.80
2. Employment and Asset Ownership	**27**	**75%**	**0.18**	**0.20**
Stable Employment	27	75%	0.32	0.33
House/Land Ownership	19	53%	0.15	0.15
Asset Ownership	22	61%	0.14	0.15
Access to the Internet	14	39%	0.07	0.00
Access to Credit	14	39%	0.07	0.00
Total Employment and Assets Indicator Weights			0.75	0.63
3. Living Standards	**26**	**72%**	**0.16**	**0.15**
Household Consumption	23	64%	0.19	0.10
Government Help and Subsidies	20	56%	0.11	0.09
Sanitation	18	50%	0.06	0.05
Drinking Water	18	50%	0.06	0.03
Washing and Cooking Water	19	53%	0.07	0.03
Electricity	15	42%	0.05	0.00
Cooking Fuel	14	39%	0.05	0.00
Floor	16	44%	0.04	0.00
Roof	14	39%	0.04	0.00
Wall	15	42%	0.04	0.00
Access to Traditional Health Facilities (for child birth)	7	19%	0.02	0.00
Total Living Standards Indicator Weights			0.72	0.29

Table 4.7 (continued)[a] [b] [c]

Dimension and Indicators	Number of Votes	% Vote	Mean Weight	Median Weight
4. Health and Safety	**31**	**86%**	**0.27**	**0.20**
Healthy life	30	83%	0.31	0.30
Safe Living Environment	29	81%	0.28	0.30
Access to Immunization and Vaccinations	18	50%	0.09	0.05
Ability to Travel (for health)	16	44%	0.08	0.00
Children to be Breast Fed	17	47%	0.10	0.00
Total Health and Safety Indicator Weights			0.86	0.65
5. Child Health and Contraception	**22**	**61%**	**0.11**	**0.10**
Modern Health Facilities (for birth attendant)	22	61%	0.30	0.33
Contraceptives	19	53%	0.16	0.20
Child Mortality	19	53%	0.15	0.10
Total Child Health and Contraception Indicator Weights			0.61	0.63
Total dimensional weights			1.00	0.90

Note: [a] Results within this table were calculated from the final (third) round of the Delphi conducted on Day 2.
[b] Median weights are arranged from largest (indicating the most important dimension or indicator) to smallest.
[c] Zero median weights indicate the exclusion of a particular dimension or indicator within the final multidimensional measure examined in Chapters 5 and 6 of this thesis.

Source: Author's computations (Bogor City, 2016)

Considering the above, excluding ten out of the twenty-eight possible indicators, led to a somewhat more balanced number of indicators per dimension. Although participants were given space to 'explain' their choices within the Delphi questionnaire, particularly if their views were divergent from the 'feedback' given between rounds, none offered detailed descriptions as to why they thought specific indicators were important within each dimension. The semi-structured robustness check interviews explored within Section 4.7, however, confirm the choice of excluding the ten indicators assigned a median weight of zero within Table 4.7. In addition, these interviews provided some clarity as to why these indicators were not particularly relevant to include in a measure aiming to reflect poverty within the case study area.

To describe the third results from Table 4.7, the 'sums' of the weights are examined. As noted at the start of Section 4.5, participants were requested to assign weights so that the sum of dimensional weights they chose equalled "1" and the sum of indicator weights within each dimension, also equalled "1". Looking at dimensional-weights, Table 4.7 shows that the sum of the *means* for these weights equalled "1".[1] This is not, however, the case when examining *median* values for these dimensional-weights. As will be explained further within Chapter 6 of this thesis, given the importance of 'medians' compared to means when analysing Delphi results, this poses a challenge when incorporating these results into the 'new' empirical measure. This challenge is also present when examining the medians for indicator-weights within each dimension.[2]

With the above key results in mind, Table 4.8 excludes indicators assigned median weights of zero and presents the 'final' list of dimensions and indicators chosen by participants.

[1] This is due to the fact that participants were required to assign weights to all dimensions, even if these weights were 'zero'. By taking these 'zeroes' into account within the calculation of means, the 'sum' of these mean-values, equalled "1".

[2] Within the Delphi, participants had to assign weights for all the five dimensions, even though these weights were zero. For indicators, participants only had to assign weights for those within dimensions that they assigned non-zero weights. When participants did not choose a specific dimension, all indicators within the said dimension will be given weights of "zero". As an example, 9 out of 36 participants did not select the 'Employment and Asset Ownership' dimension; thus, the indicators within this dimension were all assigned zero weights by these 9 participants. Since mean weights for these indicators within Table 4.7, are calculated by dividing the sum of all possible weights assigned by participants by the total number of participants (36), the sum of 'Employment and Asset Ownership' indicator means did not equal "1". If, however, means for 'Employment and Asset Ownership' indicators, were calculated by dividing the sum of weights chosen only by the 27 participants who assigned 'non-zero' weights to this dimension and dividing this sum by 27, the sum of these means would equal "1".

Table 4. 8 Final List of Dimensions, Indicators and Weights

Dimension/Indicator	Median	Mean	MAD	SD	Min	Max
Dimension 1: EDUCATION	**0.25**	**0.28**	**0.05**	**0.19**	**0.10**	**0.60**
Children's Education	0.30	0.35	0.10	0.22	0.10	0.30
Literacy	0.25	0.27	0.05	0.20	0.20	0.60
Adult's Education	0.15	0.15	0.15	0.19	0.20	0.80
Ownership of Birth Certificate	0.10	0.12	0.10	0.13	0.10	0.40
Total Education indicator weights	0.80	0.89				
Dimension 2: EMPLOYMENT AND ASSETS	**0.20**	**0.18**	**0.05**	**0.13**	**0.10**	**0.50**
Employment	0.33	0.32	0.18	0.22	0.20	0.70
House and/or Land Ownership	0.15	0.15	0.15	0.16	0.10	0.50
Assets Ownership	0.15	0.14	0.15	0.16	0.10	0.70
Total Employment and Assets indicator weights	0.63	0.61				
Dimension 3: LIVING STANDARDS	**0.15**	**0.16**	**0.05**	**0.18**	**0.10**	**0.40**
Household Consumption	0.10	0.19	0.10	0.25	0.08	0.75
Government Help and Subsidies	0.09	0.11	0.09	0.15	0.05	0.55
Sanitation	0.05	0.06	0.05	0.07	0.10	0.20
Drinking Water	0.03	0.06	0.03	0.07	0.05	0.25
Cooking and Washing Water	0.03	0.07	0.03	0.09	0.05	0.40
Total Living Standards indicator weights	0.29	0.49				
Dimension 4: HEALTH AND SAFETY	**0.20**	**0.27**	**0.10**	**0.23**	**0.10**	**0.60**
Healthy Life	0.30	0.31	0.10	0.22	0.15	0.70
Safe Living Environment	0.30	0.28	0.10	0.21	0.10	0.60
Access to Immunizations	0.05	0.09	0.05	0.10	0.10	0.30
Total Health and Safety indicator weights	0.65	0.68				
Dimension 5: CHILD HEALTH AND CONTRACEPTION	**0.10**	**0.11**	**0.10**	**0.11**	**0.05**	**0.40**
Birth attendant	0.33	0.30	0.30	0.29	0.25	0.70
Contraception	0.20	0.16	0.20	0.17	0.20	0.60
Child mortality	0.10	0.15	0.10	0.16	0.10	0.50
Total Child Health and Contraception indicator weights	0.63	0.61				
Total dimension weight	0.90	1.00				

Source: Author's computations (Bogor City, 2016)

Referring back to the results within Section 4.5, although convergence was generally observed, a degree of disagreement was still present with regard to the *exact value* of weights for the five dimensions and eighteen indicators included in Table 4.8. This in mind, this table also notes the standard deviations and MADs of these weights, alongside the 'minimum' and 'maximum' weights assigned by participants to these dimensions and indicators. The issue of the 'sum of median weights not equalling one' is again present within Table 4.8. Chapter 6 will discuss how results within Table 4.8 will be incorporated within the construction of a 'new' composite measure of multidimensional poverty based on these Delphi results.

4.7 Robustness Checks

As described within sub-section 3.4.3 of Chapter 3, the robustness steps taken within this section follow those suggested within Scheibe et al. (2002) and primarily aimed to establish whether 'consensus was forced' (p. 271). Two steps were taken; the first through a 'post-Delphi' questionnaire and the second through post-Delphi semi-structured interviews. The post-Delphi questionnaire was administered directly after the three Delphi iterative rounds (on Day 2 of the Delphi implementation process); whereas, the semi-structured interviews were conducted in the weeks following the Delphi. This section starts with results from the first step.

Within the post-Delphi questionnaire participants were presented with six statements, which aimed to primarily assess their opinions on the Delphi process and whether there was a danger that participants were 'pushed' towards consensus unwillingly; i.e. whether it was likely that the 'consensus' achieved through the Delphi would not hold once the process was completed (Scheibe et al., 2002). Participants were asked to provide ratings dependent on whether participants agreed or disagreed with a particular statement. A Likert scale was employed to indicate levels of agreement, with "1" indicating complete disagreement with the statement, "4" acting as a relatively 'neutral' category and "7" indicating complete agreement. Results from this questionnaire are illustrated within Table 4.9.

Examining the first statement on 'views on the Delphi process', 94 percent of participants gave a rating of "5" or above, indicating that in general they felt relatively 'satisfied' with the Delphi that was conducted. Responses to 'Statement 2' seem to strengthen this finding, as 97 percent of participants gave ratings of "5" and above, signalling that they tended to feel that they 'received valuable information' from the Delphi.

For 'Statement 3', 94 percent of participants gave ratings of "5" and above, indicating that participants tended to 'agree with the feedback they were given between rounds'. This somewhat validates the results of the 'stability and consensus' tests described within Section 4.5, which found that participants' views tended to move toward the group means and medians reported within the feedback sheets presented. Even though this was the case, views on 'Statement 4' show that a substantial number of participants (58 percent), tended to agree that 'additional discussion on the conclusions of the Delphi' were needed. This is again in line with that found within Section 4.5; i.e. even though 'convergence' was evident, a level of 'disagreement', particularly with regard to the exact value of dimension and indicator weights, was still present.

Table 4. 9 Post-Delphi Questionnaire Results

Statement 1 I feel satisfied with the conclusions of this Delphi exercise			Statement 2 I feel that I have received valuable input from the feedback between Delphi rounds			Statement 3 In general I agreed with the ideas in the feedback			Statement 4 I feel as if I really wanted to talk to people about my opinions			Statement 5 I think the Delphi could be operational in achieving consensus on important issue, more generally			Statement 6 I think the Delphi exercise went too fast		
Likert Category	**N**	**%**	**Likert Category**	**N**	**%**	**Likert Category**	**N**	**%**	**Likert Category**	**N**	**%**	**Likert Category**	**N**	**%**	**Likert Category**	**N**	**%**
1	0	0	1	0	0	1	0	0	1	2	5.56	1	0	0	1	2	5.56
2	0	0	2	0	0	2	0	0	2	2	5.56	2	0	0	2	4	11.11
3	0	0	3	0	0	3	0	0	3	3	8.33	3	0	0	3	6	16.67
4	2	5.56	4	1	2.78	4	2	5.56	4	8	22.22	4	0	0	4	10	27.78
5	6	16.67	5	8	22.22	5	7	19.44	5	6	16.67	5	7	19.44	5	4	11.11
6	16	44.44	6	14	38.89	6	17	47.22	6	11	30.56	6	12	33.33	6	8	22.22
7	12	33.33	7	13	36.11	7	10	27.78	7	4	11.11	7	17	47.22	7	2	5.56

Note: [(a)]Likert scale categories:
"1": Totally disagree with statement
"7": Totally agree with statement
"4": Presumed neutral category
[(b)] Total number of participants: 36

Source: Author's computations (Bogor City Delphi, 2016)

Furthering this 'need for additional discussion', 'Statement 6' results shows that a number of participants (38 percent) were inclined to think that the Delphi implementation process was 'too fast'. Despite this, views on 'Statement 5', however, indicate that a significant number of participants tended to agree that Delphi may be efficiently applied as a method of achieving a level of consensus on crucial issues, such as poverty.

In sum, results of the post-Delphi questionnaire indicate that participants were inclined to be satisfied with the Delphi process and that they tended to agree with the main conclusions of the study, thus, clarifying that the level of 'consensus' achieved through the Delphi was less likely a result of participants being 'forced' toward convergence. Despite this, however, a number of participants opined that the study would benefit from additional discussion. The second step of the robustness checks, aimed to a degree, to facilitate this need for further 'discussion', whilst preserving the anonymity of those involved.

Following the guidelines set within sub-section 3.4.3 of Chapter 3, the second robustness-check step involved conducting semi-structured interviews with fifteen participants and fifteen non-participants of the Delphi.[1,2,3] Delphi-participant interviews primarily aimed to assess whether they agreed with the conclusions of the Delphi study and whether they felt that they were 'forced' into consensus. The Delphi-non-participant interviews, then sought to ascertain, to a degree, the reliability of the Delphi's findings. Results of these interviews are described below.

All interviewees, whether Delphi or non-Delphi participants, agreed that the Delphi process presented 'an innovative way of gathering opinions from different government sectors with regard to how best to measure poverty' (Bogor City Delphi Non-Participants, 2016; Bogor City Delphi Participants, 2016). As reasons behind this level of 'satisfaction', all interviewees stated the fact that 'their opinions counted' toward the building of a poverty measure for their district

[1] How these interviewees were selected is outlined within sub-section 3.4.3 of Chapter 3. As with the Delphi, care was taken to ensure the anonymity of these interviewees.

[2] The list of Government Sectors and Divisions where these interviewees were based, is available within the Appendix to Chapter 3.

[3] All participants and non-participants contacted for these post-Delphi semi-structured interviews agreed to be interviewed. The interviews were conducted within each of the policy makers' Departmental Office. The duration for the semi-structured interviews ranged from 60 to 90 minutes. Following Bleich and Pekkanen (2013)'s call for reporting 'saturation' rates of interviews, 'saturation was reached within and across the interviewee categories'; i.e. all thirty interviewees delivered similar responses with regard to their views of the Delphi study and the conclusions arising from the study. Interviews were recorded using a voice recording device. Consent was achieved before the recording of interviews was done. Interview recordings and results were stored in a safe and password protected location. Anonymity was strictly ensured, whenever requested.

made them 'feel valued' and prompted them to 'do their best to represent their people' within the study. Furthermore, all interviewees brought up that the study's conclusions; i.e. 'discovering a way of monitoring deprivation and poverty that is relevant to their district', complied with the district government motto of 'from the people, for the people'. This then confirms the findings within Chapter 2; i.e., that within Indonesia, district specific measures are important to assess poverty and promote action. Three (Delphi participant) interviewees, in particular, further discussed the feasibility of implementing the Delphi at a larger scale, particularly to include those experiencing poverty themselves.[1] This level of satisfaction of the Delphi process is in line with results found within the post-Delphi questionnaire described at the start of this section.

In addition to the above, the interviews provided the researcher with the opportunity to explore the possible reasons behind the decision to *exclude* the ten indicators listed within Section 4.6 of this chapter. All interviewees agreed on the importance of the five dimensions and the order of their importance; with 'Education' given the highest weight, followed by 'Health', 'Employment and Assets', 'Living Standards' and 'Child Health and Contraception'. Assigning a roughly equal number of indicators to each dimension was also agreed upon.

Looking at the twenty-eight possible indicators, 'access to the internet' was seen as an aspect of well-being that should not be included within the government's 'crucial' list of priorities, particularly as Bogor City, being an urban area, had existing internet coverage, which was already relatively extensive.[2] For Bogor City, an urban area with a substantial number of small government lending banks and private microfinance facilities, 'Access to Credit', was also considered less important compared to the other indicators assigned non-zero weights. Access to 'Electricity' and the use of hard materials for 'Cooking Fuel' were also found to be less of a challenge within Bogor City. Even though Bogor City still continues to face significant challenges with regard to slum areas, issues of inadequate dwellings, defined within the Susenas as those houses made of wood or mud, were also considered as less important, compared to challenges

[1] One interviewee offered the concrete suggestion that: 'it would be feasible to include representatives of the people at the sub-sub-district level; if it is too challenging to determine who to include, specifically, at least heads of these sub-sub-districts (all 68 of them) could be included. These heads are not civil servants; they are chosen by their communities. Some may also come from difficult and poor backgrounds' (Bogor City Delphi Participants, 2016). As a follow up to this particular discussion, I was invited to present about the Delphi process to forty sub-sub-district heads. Although beyond the scope of this thesis, the presentation generated substantial interest, pointing toward the possibility of applying the Delphi within future studies.

[2] Interviewees' opinions on internet coverage in urban areas in Indonesia is supported by analysis within the Central Bureau of Statistics (CBS) (2014) report on 'Internet Usage in Indonesia'.

the city was facing with regard to 'Sanitation' and 'Clean Water'.[1] Traditional medical facilities (for childbirth), defined within the Susenas as 'shaman medicine', was also considered less important. As for 'Breast Feeding', the general opinion found was that the role of the government was to *promote*, rather than *impose* the need for new-born babies to be breast fed; the decision to 'Breast Feed' thus depends on individual circumstances experienced by both mother and child.

Considering the above, interviewees agreed with the final choice of indicators as listed within Table 4.8. Despite this, less consensus was found on how dimensions and indicators should be weighed. As noted above, the hierarchy of importance with regard to dimensions, as indicated by their median weights from the third Delphi round, was agreed to. Interviewees, however, did not agree on the *exact* values for these weights. This divergence of opinion was even more evident when discussing indicator weights. Although this was the case, the weights suggested by interviewees all fell within the minimum and maximum range for these weights listed within Table 4.8; all interviewees thus, unanimously agreed upon the relevance of this range of weights.

In sum, results from the two-step robustness-checks described above, suggested three main points. Firstly, they confirmed the relevance of the Delphi as a consensus building method, in particular as a way of facilitating decision making with regard to poverty and well-being issues within the case study area examined. Secondly, supporting the results within Sections 4.5 and 4.6 of this chapter, they established that agreement was generally achieved with regard to the *types* of dimensions and indicators which were thought to be crucial to include within a measure of multidimensional poverty. Thirdly, disagreement was however, still evident with regard to the exact value of weights to assign to these dimensions and indicators; this points to the importance of the range of weights listed within Table 4.8.

4.8 Conclusion

The main aim of this chapter was to present a descriptive analysis of results from the Delphi study conducted within this thesis. From these results four main findings can be deduced.

Firstly, unanimous agreement was achieved with regard to the multidimensionality of poverty, thus indicating the need to supplement consumption and income-based measures with

[1] This issue of slum areas and lack of adequate sanitation can also be found within a report on Bogor City published by the Central Bureau of Statistics (CBS) (Central Bureau of Statistics Bogor City, 2018).

measures that aim to capture the multifaceted nature of poverty. This finding is further supported by a relatively large number of participants whose definitions of poverty coincided with Sen's Capability Approach. When specifically asked which 'form' of measurement would best depict the multidimensional nature of poverty, participants unanimously selected the 'dashboard' form. Furthermore, participants viewed that a five-dimensional dashboard would best support the monitoring of multidimensional poverty within their district. Despite this, participants noted that 'aggregate' measures should also be constructed, however, as a supplement to this dashboard.

Secondly, a degree of consensus was achieved with regard to the importance of including all five dimensions within multidimensional measurement. These five dimensions were: 'Education', 'Employment and Asset Ownership', 'Living Standards', 'Health and Safety' and 'Child Health and Contraception'. A level of agreement was also achieved with regard to the inclusion of eighteen indicators to proxy deprivation within these five dimensions. Thirdly, despite the conformity of views on the types of dimensions and indicators which were thought to be important to include within the measurement of multidimensional poverty, disagreements were evident on the exact values of weights to assign to these aspects.

Finally, particularly confirmed by results of the post-Delphi robustness checks conducted, the Delphi was perceived as an effective and innovative tool to address the issue of selecting components to include within a multidimensional measure of poverty. Participants generally agreed as to the merits of this method as a way of enabling policy makers from different parts of the government to conduct quasi-anonymous group discussions on important and crucial issues, such as poverty.

To summarise, the findings of this chapter strengthen the theoretical arguments discussed within Chapter 1 and empirical findings described in Chapter 2 of this thesis, on the need to supplement uni-dimensional measures of poverty, with measures, which adopt a more multidimensional approach. This chapter and the previous, Chapter 3, demonstrate the merits of the Delphi as a method by which to mitigate the challenge of 'selecting' relevant dimensions, indicators and weights, to include within such multidimensional measures. More importantly, as this chapter has proceeded to demonstrate, the Delphi conducted became a key platform, which enabled participants from all parts of the district government involved, to express opinions on how to measure and monitor poverty, ultimately supporting evidence-based policy and promoting increased collaborative-action towards the alleviation of poverty within the case study area included in this study.

Chapter 5: A Dashboard Multidimensional Poverty Measure based on Delphi Results

Poverty comprises multiple areas of concern; policy makers need to monitor progress for each of the priority areas set by the government within a dashboard-type measurement tool. However, an overall composite index, which aggregates deprivation within these priority areas, is also needed as a supplement

(Bogor City Delphi Non-Participants, 2016; Bogor City Delphi Participants, 2016; Natih, 2016b)

5.1 Introduction

As described within Chapter 1 of this thesis, Amartya Sen (1976a) states that 'in the measurement of poverty, two distinct problems must be faced, viz., (1) identifying the poor among the total population, and (2) constructing an index of poverty using the available information on the poor' (p. 219). Chapters 3 and 4 of this thesis aimed to offer solutions with regard to solving Amartya Sen (1976a)'s first challenge through the use of a Delphi survey in selecting the dimensions and indicators of poverty, alongside their respective weights, which would be ideal to utilise in *identifying* the poor within the case study area studied. Following Chapters 3 and 4, the upcoming chapters within this thesis, Chapters 5 and 6, aim to propose ways in which to address Amartya Sen (1976a)'s second challenge; i.e. how best to *aggregate* the multiple indicators and dimensions chosen within the Delphi exercise, as described within Chapter 4, into a meaningful and relevant empirical measure of poverty.

When seeking to aggregate multidimensional poverty, Chiappero-Martinetti and von Jacobi (2012) state, 'there are two directions that are broadly followed'; 'on one hand, a dashboard is considered an appropriate manner for monitoring the trends of development'; 'on the other, composite indices of poverty are often formulated with the aim to facilitate comparisons over time or across countries and to simplify interpretation' (p. 69-70). This thesis argues for the importance of both approaches, with this chapter, Chapter 5, aiming to illustrate how Delphi results may be applied into forming a *dashboard* measure of poverty, whilst the next chapter, Chapter 6, seeks to tackle the challenges of creating a *composite* index of poverty based on outcomes of the Delphi exercise.

On dashboards, Stiglitz et al. (2010b) argue that they act as an 'initial step' in the analysis of poverty, a phenomenon, which is, 'by nature highly complex'; furthermore, focus on 'multiple

indicators' also 'encourages national and international statistical offices to improve on the measurement of these indicators' (Stiglitz et al., 2010b, p. 63). Martin Ravallion (2011b) however, states that 'some degree of aggregation' may still be needed even within a dashboard measure, in the sense that monitoring poverty may benefit from indicators of poverty being aggregated into dimensions. Following this argument, within this chapter, indicators are aggregated into dimensions based on the 'clustering exercise' results described within Section 4.4 of Chapter 4.

The above in mind, this chapter aims to address two main research questions. First, it aims to explore the extent of deprivation within each indicator: *what are the deprivation rates within each of the indicators chosen by Delphi participants?* Once a household's deprivation status within each indicator is established, research question two investigates the question*: what is the extent of deprivation within each dimension of poverty chosen by participants?* Both these research questions aim to describe deprivation within each aspect of multidimensional poverty considered as important by Delphi participants, before presenting these components in the form of a composite measure within Chapter 6 of this thesis.

To answer the above research questions, following Chapter 2 of this thesis, data from the Susenas 2013, is employed. The 'household' is again utilised as the unit of analysis. As indicators and dimensions chosen through the Delphi survey were selected specifically to describe deprivation and poverty in Bogor City, differing from Chapter 2, which also provided analysis for West Java and Indonesia overall, analysis within this chapter will focus on Bogor City.[1] Even though data and results for Bogor City are presented, the purpose of this chapter, as is the overall aim of this thesis, is not to be on poverty in Bogor City, per se. It is however to illustrate how outcomes from a Delphi study may be incorporated within a dashboard measure of poverty, alongside demonstrating the uses of dashboard forms of measurement within poverty analysis.

Considering the above, the rest of this chapter is organised as follows. Section 5.2 presents a brief literature review and summary of empirical results from the Delphi survey aiming to further strengthen the need to consider *both* dashboard and composite measures when assessing poverty. This section not only aims to base the decision behind the creation of the dashboard within this chapter, however, also seeks to pave the way for the analysis of the composite measure proposed within Chapter 6. Before addressing this chapter's research questions,

[1] For comparative reasons, results for West Java and Indonesia are presented within the Appendix to this Chapter and the Appendix to Chapter 7.

Section 5.3 presents an overview of local laws and regulations, alongside international criteria, which offer guidance on recognised standards that are often utilised to define levels of deprivation within indicators of poverty. Section 5.4 describes the results of the two research questions noted above. Section 5.5 concludes.

5.2 Dashboard or Composite?

Arguments that poverty concerns deprivations within multiple aspects of life were strengthened by the results of the Delphi. Given this multidimensional nature, the decision on whether to 'collapse all information into one number, or to keep separate the different dimensions' of poverty, becomes crucial (Brandolini, 2008, p. 8). This section aims to first outline the theoretical arguments found within literature underlining the pros and cons of both dashboard and composite measures, before revisiting key results derived from the Delphi on this issue. Through this discussion, this section seeks to clarify the need to focus on both measurement forms, dashboard and composite, when measuring multidimensional poverty.[1]

5.2.1 Literature Review

The 'dashboard vs composite' debate was most prominently brought forward by Martin Ravallion. Martin Ravallion (2011b) argues for a 'dashboard' measure by stating that the aim of poverty measurement should be to create 'a credible set of multiple indices rather than a single multidimensional index' (p. 235). Further emphasising this point, Martin Ravallion (2011a) questions the merits of a composite measure as a tool to base policy formation, for 'surely, policy makers will naturally want to look at a country's attainment in various dimensions, rather than focusing on its performance with respect to a single composite index'; the ability to

[1] As a point of reference, the Oxford Poverty and Human Development Initiative (OPHI) (2016b), notes the importance of differentiating between 'composite' and 'aggregate' indices. Within poverty measurement, they note that 'composite' measures refer to those, which 'first aggregate across unit data and then build an overall measure'. Following this definition, 'composites' thus, aggregate data from different data sources. Considering this, they define the Global MPI as an 'aggregate' index, rather than a 'composite'. Although acknowledging that the measure suggested within Chapter 6, would resemble OPHI's 'aggregate' definition, as it aims to compile information from one data source, within this thesis, Nardo, Saisana, Saltelli, and Tarantola (2005)'s definition of 'composite' measure is utilised; i.e. 'a composite measure is simply the mathematical combination of individual dimensions that represent different aspects of a concept whose description is the objective of the analysis' (p. 7). This definition thus, encompasses both OPHI's interpretations of what 'aggregate' and 'composite' measures entail. It is often the practice within literature to refer to multidimensional indices as 'composites'. As an example, when discussing the Global MPI and HDI, Martin Ravallion (2011b) refers to both as 'composite indices' (p. 4). The aforementioned decision to follow this practice within this thesis, was thus made.

scrutinise deprivations separately is indeed called for, particularly for 'the task of prioritising policies' (p. 237).

As touched upon within Chapter 2 of this thesis, critiques against the use of composite measures are strongly anchored upon the 'arbitrary judgment' basing the decisions behind the selection of components within such measures, particularly on the choice of weights used to aggregate multiple indicators into a single measure (Laurens Cherchye, Moesen, Rogge, & Van Puyenbroeck, 2007; Martin Ravallion, 2011b). On this, Ravallion (2011a) states that existing composite measures of multidimensional poverty, possess an 'unusually large number of moving parts, which the producer is essentially free to set', more often than not without the guidance of theory and sound methodological assumptions (p. 2). He dubs these indices 'mashup indices' and defines them as 'composite measures for which the producer is only constrained by the availability of data in choosing what variables to include and their weights' (Martin Ravallion, 2011a, p. 2).[1]

The choice of 'weights' are of particular concern within 'mashup indices', as often, the practice is to utilise 'equal weights' to reflect the argument that 'all components within such measures are of equal importance' (Decancq & Lugo, 2008). This reasoning is, however, questionable as 'weights of zero are assumed for all indicators not included within the measure'; thus, 'similar to any other weighting scheme, the equal weighting scheme implies trade-offs between indicators that may or may not be reasonable' (Decancq & Lugo, 2008, p. 15).

What further puts the use of 'equal weights' into question, however, is the fact that value judgments made behind their use are rarely made explicit (Martin Ravallion, 1997). This is crucial, as arbitrary judgment opens the possibility of 'misinterpretation and manipulation of results', which could lead to important policy implications (Chiappero-Martinetti & von Jacobi, 2012, p. 70; Martin Ravallion, 2011a, pp. 2-3). Considering this, Martin Ravallion (2011b) himself, however, does not rule out the importance of composites as they enable assessment of the 'extent to which different deprivations are jointly experienced by the same individuals or households' (Martin Ravallion, 2011b, p. 236). Nevertheless, he argues that composites should ideally only be constructed when substantial data is available and information on weights and trade-offs between indicators can be gathered. This information should, in turn, be based on the

[1] Ravallion (2011a) categorises the Global MPI, examined within Chapter 2 of this thesis, as one such index. He also includes the Human Development Index (HDI) within this category.

opinions of those with ‘a stake in the outcomes’, whilst taking into account ‘the specific country and policy context’ where measurement is needed (Ibid, p. 246).[1]

In sum, critiques of composites seem for the most part, to be focused on how current practices of aggregating multidimensional poverty mainly base their indices on arbitrary assumptions, particularly with regard to the types of indicators they include and the use of weights, which are rarely grounded within theory or based on empirical evidence. Less criticism, however, seems to be aimed at the *need* for composite measures.

F. H. G. Ferreira and M. A. Lugo (2013) state that multidimensional poverty analysis is ‘interesting’ because ‘the dependency structure’ of deprivations across different dimensions ‘may affect how poverty is assessed’ – ‘a dashboard that reports deprivation for each dimension separately would miss this aspect of joint distribution’ (p. 4). These authors, however, do not explicitly argue for composite measures, particularly, as discussed above, among other considerations, a composite measure’s need for weights makes it cumbersome. Instead, to complement a dashboard, they advocate for a ‘direct representation of the dependency structure’ between deprivations across various dimensions; i.e. an exploration of the proportion of households which are deprived in one dimension, two dimensions, continuing to a description of those deprived in the maximum number of dimensions considered (F. H. G. Ferreira & M. A. Lugo, 2013, p. 225). Unlike composite indices however, this exercise cannot generate complete orderings of households or regions. Furthermore, it is possible that ‘a set of univariate analyses done independently for each dimension to conclude that poverty in A is lower than poverty in B while a multivariate analysis concludes the opposite’ (Duclos, Sahn, & Younger, 2006, p. 945). This may result from interaction of the various dimensions and their correlation in the sampled population; ‘a one-at-a-time comparison of poverty across separate dimensions, cannot capture these interdependencies’ (Ibid).

In addition to the above, Stiglitz et al. (2010b) state that ‘as a communication instrument, a dashboard measure, lacks what has made GDP a success: the powerful attraction of a single headline figure allowing simple comparisons of socioeconomic performance’. Without composite measures, the simple questions of ‘Who is poor overall? How many poor people are there? How poor are they?’ cannot be addressed (Sabina Alkire & James Foster, 2011, p. 504). Indeed,

[1]Martin Ravallion (2011b) further notes that the decision basing which indicators to include and how they should be weighed against one another will ‘often be political’; it is thus crucial, to ensure that this choice is a ‘well informed one’ (p. 246).

'composite indicators are much easier to interpret than trying to find a common trend in many separate dimensions' (Nardo et al., 2005, p. 6). The parsimonious way composite measures present information may also help attract public interest and advocacy (Michaela Saisana & Tarantola, 2002; Saltelli, 2007, pp. 65-67).

Given the above, it is clear that *both* dashboards and composites exhibit valuable properties. Dashboards provide a detailed description of deprivation rates within each dimension, enabling an assessment of areas of well-being needing urgent action. Composites complement this, by allowing the analysis of whether multiple deprivations are experienced simultaneously, alongside enabling the construction of a 'headline figure', which may be compared across time and space. With these theoretical considerations in mind, the sub-section below presents a summary of empirical results from the Delphi study conducted within this thesis, concerning this debate.

5.2.2 Empirical Considerations

This sub-section revisits the results described within sub-section 4.3.1 of Chapter 4. As found within Chapter 4, although results of the Delphi cannot claim to be generalizable, a reliable level of consensus with regard to the final results was achieved within the panel of policy makers included in the study.

With regard to the 'dashboard vs. composite' argument, even though Delphi participants valued dashboards as the 'main form' of measurement crucial to monitor deprivation within their district, merits of composites were also recognised. Participants found it crucial for a measure of poverty to be able to shed light on deprivation rates within the dimensions they chose to be relevant. A composite measure aggregating these dimensions was then considered as an important supplement to enable poverty comparisons (Bogor City Delphi Non-Participants, 2016; Bogor City Delphi Participants, 2016; Natih, 2016b).

Semi-structured interviews, described within Section 4.7 of Chapter 4, conducted to test the robustness of these results, further confirmed this conclusion. In particular, interviewees (whether Delphi-participants or non-participants) made arguments, which echoed that made by Chiappero-Martinetti and von Jacobi (2012) who state that a 'dashboard', where dimensions are analysed separately, provides an 'appropriate way of monitoring trends of development within each key dimension of poverty'; whilst a composite index 'facilitates comparisons over time or across countries and simplifies interpretation' (p. 69-70).

With the above argument for forming both dashboard and composite measures in mind, as stated within Section 5.1 above, this chapter will focus on building a dashboard based on Delphi results. Before discussing the results of this dashboard, Section 5.3 discusses the choice of indicator cutoffs utilised to define conditions of deprivation within each of the indicators considered.

5.3 Indicator Cutoffs

As described within sub-section 3.4.1.3 of Chapter 3, indicators within this chapter were taken from an existing dataset, i.e. the Indonesian National Socioeconomic Survey (Susenas) 2013. Participants of the Delphi were requested to group these pre-existing indicators into dimensions through a 'clustering' exercise conducted at the start of the study.[1] From these newly formed dimensions, participants were asked to select which were most relevant to include within a measure of poverty for their district. Following this, participants' views on important indicators, which may be utilised to proxy deprivation within the said dimensions, were solicited. As described within Section 4.6 of Chapter 4, five dimensions were chosen, i.e. (1) Education, (2) Employment and Asset ownership, (3) Living Standards, (4) Health and Safety and (5) Child Health and Contraception. Eighteen indicators were then selected to proxy these dimensions.[2]

Within dashboard measures, two types of deprivation cutoffs need to be considered.[3] The first are *indicator* cutoffs, which define the level of deprivation a household needs to experience in order to be considered as deprived within a particular indicator. Once these indicator cutoffs are determined, this thesis follows the method employed within S Alkire and James Foster (2011), which assigns a score of "1" to a household which is *deprived* in a specific indicator and a score of "0" when the household is *non-deprived*.[4] The second type of cutoff, *dimensional* cutoffs, define the total number of indicators a household needs to be deprived in to be considered as deprived within a specific dimension. This current section will focus on indicator cutoffs, whilst dimensional cutoffs will be the subject of Section 5.4.2 of this chapter.

[1] Section 3.4.2.1 of Chapter 3 discusses the methodological steps within this 'clustering' exercise. Section 4.4 of Chapter 4 describes its results.

[2] As a brief reminder, within this thesis, 'dimensions' thus, refer to the broader categories encompassing multiple indicators. 'Indicators' then serve to proxy conditions of deprivation within the said dimensions.

[3] Following Bourguignon and Chakravarty (2003) and S. Alkire, Foster, and Santos (2011), the term 'cutoff' within this thesis is used to refer to 'the level of deprivation considered as the basis for determining who is deprived in which dimension or indicator'.

[4] Sub-section 6.3.2 of Chapter 6, further discusses the arguments behind the choice of this method.

Within this chapter, indicator cutoffs are based on existing welfare standards found within international and nationally accountable literature and legal documents.[1] In the case where international and national standards differ, national standards are chosen in the hopes of 'embodying the prevailing norms of what it means to be poor' in the case study area selected (S. Alkire et al., 2011, p. 310). In addition, as Indonesia is a vast country, divided into hundreds of decentralised districts, which are allowed to set their own local standards, whenever it is available, localised district level criteria is employed to determine these cutoffs (Brodjonegoro, 2003). These cutoffs aim to reflect the minimum accepted standards with regard to these indicators. This decision of setting indicator cutoffs could arguably be given to the Delphi participants themselves. As noted within Chapter 3, Harold A. Linstone and Turoff (1977), Dietz (1987) and Keeney et al. (2010), among others, however, warn of the significant danger of 'fatigue' occurring within Delphi surveys. An informed choice was made to focus participants' time and attention on the choice of dimensions, indicators and their respective weights. When time and resources permit however, an additional Delphi focused on indicator cutoffs may be beneficial, not only to support the construction of the poverty measure, but also to highlight possible areas where current international and national standards need revision.

With the above in mind, Table 5.1 summarises these cutoffs.

5.3.1 Education Indicators

The importance of 'Education' as a dimension of poverty is not only supported by the views of Delphi participants, however also recognised by both internationally accepted norms and national level laws within Indonesia (Republic of Indonesia National Constitution, 1945; UN General Assembly, 1948; United Nations, 2016; United Nations Department of Public Information, 2009). Delphi participants chose four indicators to proxy deprivation within 'Education', i.e. (1) Children's education (for individuals who are of compulsory school age, aged seven to fifteen years), (2) Adult education (for individuals aged sixteen years and over), (3) Ownership of a birth certificate (for individuals aged zero to seven years old) and (4) Literacy (for individuals aged fifteen years and above). The rest of this sub-section examines cutoffs for each of these indicators.

[1] This method is also employed to set indicator cutoffs within the Global and Adjusted MPIs.

Table 5. 1 Summary of Indicator Cutoffs

Dimension	Indicator	Indicator cutoff A household is deprived if …
EDUCATION	Children's Education	The household has at least one member who is of pre-school age (3-6 years) that does not have access to pre-school OR has at least one member who is of school age (7-15 years) and is not attending school in school-years one to nine.
	Adult's Education	The household has no adults (16 years and over) who have at least a secondary school level certificate.
	Literacy	There is one or more eligible individual (aged 15 and above) within that household that is illiterate.
	Ownership of Birth Certificate	There is one or more eligible individual within that household that does not own a birth certificate.
EMPLOYMENT AND ASSETS	Employment	The household is workless.
	Ownership of Assets	Household does not own assets such as motorcycles, farm animals, motor boats, or other assets that can be sold for more than 500,000 IDR (around 25 GBP).
	Ownership of House and/or Land	The household does not own any of the two assets in this category.
LIVING STANDARDS	Household Consumption Level	Household consumption is below the government poverty line.
	Sanitation	Household's toilet is shared; if toilet is not shared, but household does not have a flush toilet or latrine, or ventilated improved pit or composting toilet; if waste is disposed in an open space (river, pond, beach, other open areas).
	Drinking Water	A household is deprived if its source of drinking water is not piped water, public tap, borehole or pump, protected well, protected spring or rainwater. A household is also deprived if the distance between the source of drinking water and the household's waste disposal area is less than 10 meters.
	Cooking and Washing Water	Household's source of washing and cooking water is not piped water, public tap, borehole or pump, protected well, protected spring or rainwater.
	Government Help	Household does not have any access to government subsidies.

Table 5.1 (continued)

Dimension	Indicator	Indicator cutoff A household is deprived if …
HEALTH AND SAFETY	Healthy life (not experiencing health problems that hinder daily activities).	Household has at least one individual who experiences illness that leads to the inability to conduct daily activities for more than 3 days (the median number of days lost to illness).
	Access to immunizations.	If there is one or more eligible children within that household who has not received the appropriate immunizations.
	A safe environment (safety from crime).	A household is deprived if there is one member who has experienced crime within the last year.
CHILD HEALTH AND CONTRACEPTION	Access to modern health facilities (for birth attendant).	Household has at least one eligible woman who does not have access to modern health facilities for childbirth.
	Child mortality.	At least one woman within the household who has ever experienced child mortality.
	Contraception	Household has at least one woman (who does not want to have more children, who has no medical and cultural reasons against the use of contraceptives) who does not have access to contraceptives.

Source: Author's summary

5.3.1.1 Children's Education

The UNICEF (2007) report on a 'Human Rights Approach to Education for All' states that 'the right to education is a means to reduce disparity and poverty' (p. 15). For children specifically, attending school provides a chance to 'acquire capacities, fulfil aspirations, increase emotional development and forge friendships outside the family' (UNICEF, 2007, p. 21). Education not only acts as a 'route through which economically and socially marginalized children can escape poverty and participate fully in their communities', it also, 'safeguards children from exploitative and hazardous labour and sexual exploitation' (Ibid). In Indonesia, the crucial need for children to be educated is recognised within the National Constitution, i.e. clause 31, which states that 'every Indonesian citizen has the right to be educated' (Republic of Indonesia National Constitution, 1945). Specific regulations on educational standards for children are then found within the nation's Ministry of Education laws (Republic of Indonesia Ministry of Education and Culture, 2003, 2014a, 2014b). Standards set within these laws are summarised below.

The Government of Indonesia's Law Number 47 Year 2008 on Compulsory Education, states that 'every child aged seven to fifteen has to attend compulsory education (years one to six in Primary School, and seven to nine in Secondary School)' (Republic of Indonesia Ministry of Education and Culture, 2008, 2015, 2018). Alongside the importance of 9 years of compulsory education, the role of early childhood learning is also increasingly being recognised in Indonesia; 'early childhood learning is crucial to maintain, as it determines the readiness of a child to enter his/her compulsory education years. Access to pre-school education for children aged three to six, thus needs to be ensured by the government as it plays an important roles in shaping a child's future success' (Republic of Indonesia Ministry of Education and Culture, 2014a, 2014b, 2018).

Considering the above, following the standards set by the Indonesian Government, the cutoff for 'Children's education', is set as: *a household is deprived if it has at least one child of pre-primary age (aged three to six) who does not have access to early childhood education or if the household has a child of compulsory-school-age (aged seven to fifteen), who is not attending school in school-years one to nine.* [1]

5.3.1.2 Adult Education

The importance of 'compulsory education' explored above also bases the setting of the cutoff for 'Adult Education'. Following the Government of Indonesia's law, which stipulates

[1] Following the practice within the Global MPI, households with no eligible individuals applicable to a particular indicator are consider non-deprived within the said indicator.

that 'each individual citizen should achieve at least nine years of compulsory education', *a household is considered deprived if there are no adult members (aged sixteen and above) within the household who have a middle school certificate or higher* (Republic of Indonesia Ministry of Education and Culture, 2008). The more stringent criteria would be to require that *all* adults within the family possess middle school or higher-level certificates.[1] However, given the importance of the household as a unit in Indonesia, the ability of at least one family member to achieve a higher level of education, is assumed to create benefits for all household members. This follows S. Alkire and Santos (2010a)'s reasoning for 'shared positive effects' within the household (p. 15).

5.3.1.3 Ownership of Birth Certificate

UNICEF (2013) states that, 'registering children at birth is the first step in securing their recognition before the law, safeguarding their rights, and ensuring that any violation of these rights does not go unnoticed'. Registration also serves 'a statistical purpose; collecting essential data for planning and implementing development policies' (UNICEF, 2013, p. 6). In Indonesia, without a birth certificate, individuals cannot apply for a passport and citizenship card, cannot enter schooling and are not recorded as potential recipients of Government subsidies, which include educational subsidies (*Republic of Indonesia National Law Number 23 Year 2006 on Citizenship*, 2006). [2] With these considerations in mind, the cutoff for this indicator is thus, *all eligible members of the household should have a birth certificate for the household to be considered non-deprived in the 'birth certificate' indicator.* Within Susenas 2013, data on 'ownership of birth certificate' is available for household members who are aged zero to seventeen years old. Thus, for this indicator 'eligible members' refers to those aged zero to seventeen within the household. [3]

5.3.1.4 Literacy

For the final 'Education' indicator, 'Literacy', the cutoff is that, *all household members over the age of fifteen should be able to read and write for the household to be considered non-deprived in literacy.* The data for 'Literacy' available within the Susenas is in the form of 'self-reported abilities to read and write'. The above cutoff was chosen based on the criteria set within the Government of Indonesia's national development plan on the role of 'literacy for

[1] As a comparison, if this more stringent cutoff were utilised, an extremely high level of deprivation (over 50 percent of households) was found to be evident.

[2] As an example, Bessell (2009)'s work on child poverty in Indonesia found that 'a majority of urban street children were excluded from accessing public services (such as school and healthcare), usually because they did not hold a residency card' (p. 531). A birth certificate is a pre-requisite document, which citizens have to own in order to obtain a residency card.

[3] For this indicator, physical evidence of ownership of a birth certificate is requested for all eligible members of the household (Central Bureau of Statistics, 2013b).

welfare', which states that 'every eligible member of the household (aged fifteen and above) has to be able to read and write' (Republic of Indonesia Ministry of National Development Planning, 2019). This goal is supported by international standards, in particular the Sustainable Development Goals' (SDGs), fourth goal, which stipulates that 'by 2030, nations should ensure that all youth and a substantial proportion of adults, both men and women, achieve literacy and numeracy' (United Nations, 2016).

5.3.2 Employment and Assets Indicators

5.3.2.1 Employment

Before describing the cutoff for this indicator, a short discussion on how 'employment and work' is measured within the Susenas is warranted. The Central Bureau of Statistics (2018a) defines being 'unemployed' as 'those who are not working, however are looking for work; those who are not working, however are setting up their own business; those who are not working and who are discouraged to find work; and those who have a promised job, however, are currently not working'.

Information on unemployment in Indonesia is gathered through the Indonesian Labour Force Survey (Sakernas). The Sakernas cannot be linked to the Susenas, as different households are sampled within these two surveys. Acknowledging this, the Susenas, however, *doe*s collect some information on 'work'. Even so, within this survey, it is somewhat unclear how to discern between a person who is 'unemployed' according to the Central Bureau of Statistics (2018a)'s definition above and a person who the survey records as 'economically inactive'. The Susenas defines 'economically inactive' as those who are 'conducting activities other than work during the last week, including those who are not seeking to be employed, however want a job and those who do not want a job' (Central Bureau of Statistics, 2013b; The Australia Indonesia Partnership for Economic Governance (AIPEG), 2017). The Susenas then defines 'work' as 'an activity that is done for the purpose of gaining an income or profit and is done for at least one hour per week in a continuous and routine way, during the last week' (Central Bureau of Statistics, 2013b, p. 80).[1]

Considering the above, within this chapter, focus can only be given to those who 'work' vs. those who are dubbed as 'economically inactive'. From the data available within the Susenas, we can thus, identify what Leaker (2009) dubs as 'workless' households, i.e. households where 'none of the adults of working age are in employment' (p. 46). Leaker

[1] Within the Susenas, individuals aged ten and above are asked the question of whether they have 'worked during the last week'. The official legal working age for Indonesia is, however, 'fifteen-years and older, whilst no upper age limit is imposed' (Central Bureau of Statistics, 2018a; Suryadarma, Suryahadi, & Sumarto, 2007a, p. 544).

(2009), Platt (2010) and Gregg and Wadsworth (2011), among others, point to increasing interest on the effects of a reduction in economic activity on welfare and highlight the detrimental effects of living within such 'workless' households. These authors further advocate the need for policy to 'reach out' to these households.[1] In particular, Gregg and Wadsworth (2001) show that 'workless households are highly likely to be living in poverty and to have attendant social problems' (p. 762-764).[2] Furthermore, workless households with children are shown to be even more vulnerable, as 'long term impacts of worklessness may lead to intergenerational transmissions of deprivation' (Platt, 2010, p. 73).

The importance of 'stable work' is also recognised within the Republic of Indonesia National Constitution (1945), Article 27 Paragraph 2, which states that 'every citizen has the right to work and to receive an adequate income to sustain a level of welfare that is deemed appropriate'. Although the link between 'household worklessness' and poverty is recognised, the mechanisms by which it may impact welfare, are relatively under-researched in Indonesia.[3] Considering the above, the cutoff chosen for this indicator follows that defined within international literature, i.e. the definition utilised within Leaker (2009), among others, which is: *a household is deprived if no adults work within it (if it is workless)*.[4]

[1] Leaker (2009) uses the 'international standard' definition to define the 'economically inactive', i.e. 'those who are not in employment or unemployed, which includes those who want a job but have not been seeking work in the last four weeks, those who want a job and are seeking work but are not available to start work, and those who do not want a job' (p. 42). This definition is similar to that utilised within the Susenas.

[2] Further evidence on this for the UK can be found within Gregg, Harkness, and Machin (1999), for the EU within Micklewright and Stewart (1999) and de Graaf-Zijl and Nolan (2011), for Asia within Hwang (2011), among others.

[3] The impact of unemployment on poverty is however, more thoroughly analysed. Alisjahbana and Manning (2006), Suryadarma et al. (2007a) and Suryadarma, Suryahadi, and Sumarto (2007b), among others, provide further discussions on this.

[4] Given the lack of concrete guidance on how deprivation within this indicator is defined in the Indonesian context, the use of more stringent cutoffs was tested. Using data from Bogor City, it was found that applying these stricter cutoffs significantly increased the number of deprived households. For example, requiring at least one adult to work within the household, leads to a significant increase in the households deprived within this indicator (from five when the 'workless' cutoff is used to seventeen percent). Requiring at least two adults to work, increases the deprivation rate to 55 percent. In addition to this, if, for example, the 'at least two adults work' cutoff were to be applied, one-person households may automatically be considered as deprived, even if said household member works. Without adjustment models for single person households, utilising this more severe cutoff may thus, be problematic. As these adjustments were beyond the scope of this thesis, the decision to utilise the less stringent 'workless' cutoff was made. In utilising this cutoff, however, acknowledgement needs to be given, underlining that the deprivation rate resulting from this cutoff, is likely, to be the minimum level of deprivation, which may be evident within this indicator. This follows the aim of the cutoffs, which as stated at the start of Section 5.3 above, seeks to reflect minimum levels of suffering within the indicators considered.

5.3.2.2 Ownership of Assets

Alongside 'stable work', the 'ownership of assets' is seen to 'contribute directly to the household's income generation process' or 'as means to ensure future income', particularly when households experience adverse shocks (Berloffa & Modena, 2013; Central Bureau of Statistics, 2005). Furthermore, Lerman and McKernan (2008) state that 'economic security throughout the life course is intrinsically linked to not only income, however, also asset ownership' (p. ii). Indonesia's Ministry of Development Planning highlights the role of 'asset-ownership' as a form of 'coping mechanism' by stating that 'the loss or the decrease of the ownership of assets is an important risk factor that often becomes the cause of poverty, particularly when negative shocks occur' (Badan Perencanaan Pembangunan Nasional (BAPPENAS), 2018).

To reflect these considerations, the Indonesian Central Bureau of Statistics (2005) defines a minimum criterion for 'asset ownership', which aims to enable households to maintain their livelihoods and weather possible sudden economic shocks. The cutoff for this indicator follows this definition, i.e. *a household is deprived if it does not own assets such as motorcycles, farm animals, motorboats, or other assets than can be sold for more than 500,000IDR (around 25 GBP).*[1]

5.3.2.3 Ownership of House/Land

The importance of property rights is acknowledged within the Government of Indonesia's official short and median term strategies for development; in particular, the Indonesian Ministry of National Development Planning (BAPPENAS) recognises that for both the rural and urban poor, property rights over housing provide shelter, dignity and a means for wealth accumulation. Property rights also act as 'a buffer to smooth consumption in times of shocks, providing more stability, when compared to those who are landless' (Badan Perencanaan Pembangunan Nasional (BAPPENAS), 2018; Meinzen-Dick, 2009). With this in mind, the cutoff for 'House or Land Ownership', is that *a household is deprived if it does not own the house or land that it lives in'.*

5.3.3 Living Standards Indicators

5.3.3.1 Household Consumption

Martin Ravallion (1994) states that most analysts using household data for developing countries utilise 'current consumption', as opposed to income, as an indicator of living

[1] The Susenas 2013 documents the ownership of fourteen assets; i.e. the ownership of a bicycle, motorbike, car, boat, motor-boat, cable TV, AC, refrigerator, water heater, LPG gas canister, landline-phone, mobile phone, PC and laptop (Central Bureau of Statistics, 2013b).

standards; this is due to the fact that, 'incomes of the poor often vary over time in fairly predictable ways (and sometimes in quite un-predictable ways)' (p. 13). As described within sub-section 2.2.1 of Chapter 2, 'consumption expenditure' acts as the main measure utilised to assess poverty in Indonesia. This in mind, the cutoff for this indicator follows the Central Bureau of Statistics (2013a)'s official definition of the poverty line, in that *a household is deprived if its consumption level is below this line.*[1]

5.3.3.2 Sanitation

Access to adequate sanitation and clean water is closely linked to a household's standard of living. As stated by the World Health Organization (WHO) (2004), 'significant progress in standards of living has been possible with simple sanitation improvements'; indeed, it is noted that, 'safe drinking water and adequate sanitation are crucial for poverty reduction and sustainable development' (United Nations (UN), 2015). The standards set within the "Republic of Indonesia Law Number 3 Year 2014 on Total Coverage of Sanitation for Communities" 2014) coincides with that stated by the World Health Organization (WHO) (2018a). These standards are thus followed to set the cutoff for 'Adequate Sanitation', i.e. *a household is deprived if its toilet is shared; if toilet is not shared, but household does not have a flush toilet or latrine, or ventilated improved pit or composting toilet; if waste is disposed in an open space (river, pond, beach, other open areas).*

5.3.3.3 Drinking Water

The cutoffs for 'clean drinking water' follow specific standards set by the Republic of Indonesia Ministry of Health (2018) and Central Bureau of Statistics (2018b); i.e. a *household is deprived in 'drinking water' if its source of drinking water is not piped water, public tap, borehole or pump, protected well, protected spring or rainwater.* A household is also deprived *if the distance between the source of water and the household's waste disposal area is less than 10 meters.* This criteria is also recognised within the World Health Organization (WHO) and the United Nations Children's Fund (UNICEF) (2017).[2]

[1] Despite limited literature on multidimensional poverty in Indonesia, Hanandita and Tampubolon (2015) provide discussion on the creation of a multidimensional composite measure, using data from the Susenas. The selection of the indicators within this measure aimed to mirror those included within the HDI. This in mind, these authors also propose the inclusion of 'consumption' within their measure; however, they argued for a more stringent criterion for this indicator, i.e. a cutoff, which is above the Indonesia official national poverty line. As there exists no empirical evidence to base the decision behind which poverty line to use and following the aim of the cutoffs within this chapter to reflect minimum levels of deprivation, the decision was made to use the national poverty line.

[2] As noted within Chapter 2, there are missing values evident for this indicator. These missing values make up 8 percent of the sample in Bogor City. Thus, the retained sample for drinking water in Bogor City is 92 percent of the total sample. Table 7.10 within the Appendix to Chapter 7 lists indicators, where missing values are present.

5.3.3.4 Cooking and Washing Water

The cutoffs for 'clean cooking and washing water' utilised the same standard set for 'drinking water' within sub-section 5.3.3.3 above, however, following how adequacy within this indicator is defined within the Susenas, the additional requirement that 'the water source should be at least 10 meters from waste disposal areas', does not apply to this indicator.[1]

5.3.3.5 Government Help

For 'Government help', the Susenas provides data on government subsidies and transfers that aim to help households smooth their current consumption levels and plan for the future. This information covers details of whether households receive government rice subsidies (RASKIN), conditional cash transfers, which aim to keep children in school and/or promote prenatal care for households with pregnant women, government soft loans, school scholarships, pensions and free health care (Central Bureau of Statistics, 2013b).

As stated within the start of Section 5.3 above, the cutoffs selected aim to reflect a *minimum* level of deprivation within indicators. Setting these 'minimum accepted standards' cutoffs hopes to mitigate potential arguments, which may arise, particularly if the final measure identifies significantly higher levels of poverty compared to existing measures (Laurens Cherchye & Kuosmanen, 2004; Laurens Cherchye et al., 2007). As Laurens Cherchye et al. (2007) argue, setting the cutoffs to reflect a minimum level of deprivation, makes it 'difficult for policy makers to argue that the construction of the measure itself has led to higher levels of poverty, compared to that identified by existing measures' (p. 125).

This in mind, the cutoff for this variable aims to reflect the minimum amount of 'government help' a household may be eligible to receive. As 'help' is anchored to household consumption levels, the indicator most frequently utilised as a criterion when selecting recipient households, the cutoff for this indicator is thus, *a household is deprived if it is poor according to the government consumption measure, and if it receives no help or subsidies from the government.*

[1] As will be described within Figure 5.1 of sub-section 5.4.1, this omission of the 'distance from waste disposal area' criteria leads to significant differences in the number of households categorised as deprived within the above two water related indicators.

5.3.4 Health and Safety Indicators

5.3.4.1 Healthy Life

The first indicator, 'Healthy Life', refers to whether illness affects an individual's ability to fulfil daily activities (Central Bureau of Statistics, 2013b).[1] The World Health Organization (WHO) (2009) describes the correlation between health status and poverty in that, 'ill-health can contribute to losses in individual utility or social welfare both directly (because people prefer to be more healthy than less healthy) and indirectly by reducing the enjoyment or utility associated with the consumption of goods and services unrelated to health, or by compromising other economic objectives, such as producing income that allows people to consume market goods' (p. 3).

With regard to how health influences poverty, literature talks of the 'cost of illness framework'. Within this framework, two approaches are commonly used to calculate 'economic cost' incurred by households due to illness; the input-based and output-based methods (World Health Organization (WHO), 2009, pp. 3-6). The 'input-based approach' measures the 'production losses as a consequence of illness through assuming that the duration of an individual's absence from work fully corresponds to the market value of those lost days' (Rice, 1967; Tarricone, 2006, p. 57; World Health Organization (WHO), 2009, p. 6). The 'output-based approach', assesses the 'value of the production of the sick person compared to the counterfactual of what would have happened in the absence of illness' (Attanayake, Fox-Rushby, & Mills, 2001, p. 597; Goldschmidt, 1987; Liu, 2016; World Health Organization (WHO), 2009, p. 74; Yi, Heckman, Zhang, & Conti, 2015).

The World Health Organization (WHO) (2009), points to the weaknesses of using the input-based approach, as it may 'overestimate economic losses due to the approach overlooking the coping strategies by households to mitigate the adverse circumstances imposed by sickness', and suggests the use of the 'output-based' approach, as it measures actual (rather than potential) net losses in income' (p. 6).[2] Despite the clear advantages of the output-based approach, most literature on the estimation of the 'cost of illness', chooses to

[1]The Susenas defines 'an individual experiencing illness' as, a person who during the last one month, experienced 'chronic illness, whether physical or mental, experienced injuries as a result of crime or work or non-work related accidents', or 'experienced either or more than one of the following: fever, cough, cold, asthma, trouble breathing, diarrhoea, continuous headaches, toothache, and other debilitating health complaints' (Central Bureau of Statistics, 2013b, p. 52). The Susenas also provides data on the number of days a household head forfeits his or her daily activities as a result of illness. Daily activities are defined as 'work, schooling, taking care of the household or other activities' (Central Bureau of Statistics, 2013b, pp. 52-53).

[2] Hanandita and Tampubolon (2015) also utilise 'number of days lost' to proxy 'morbidity' within their proposed index of multidimensional poverty. The argument behind their chosen cutoff for this indicator (4 days lost) is, however, unclear.

utilise the input-based approach, notably due to the 'complexity of questions around income and lost production' (World Health Organization (WHO), 2009, p. 73). This chapter follows this trend, albeit with a somewhat 'hybrid' interpretation of the 'input-based approach' due to the data limitations within the Susenas, which are outlined below.

Within this chapter, the indirect costs of illnesses are estimated by the number of days lost due to illness. A monetary value cannot be attached to these 'number of days lost' as the Susenas does not provide an estimate for individual wages. Using 'days lost' rather than monetary estimates of a household's loss of income due to illness follows Goldschmidt (1987)'s method of using 'actual output units', which represents a 'hybrid' approach, coupling aspects from both the input and output-based approaches. Use of this hybrid method, however, leads to an additional consideration, i.e. how many days lost due to illness warrants a household's status to be considered as deprived?

This chapter follows Attanayake et al. (2001), in that the distribution of the number of days lost within the sample dataset is analysed and the *median* days lost is chosen as an arbitrary cutoff for this indicator. In addition to this, the following issues also need to be addressed; i.e. *which* individuals' health and *how many* individuals experiencing illness, influence the household's deprivation status within this indicator? On this issue, S. Alkire and Santos (2010a) argue for the consideration of 'negative externalities' when setting health related indicator cutoffs. In this case, these authors argue that having a sick person within the household influences the welfare of the whole household. The World Health Organization (WHO) (2013) report strengthens this idea as although the impact of illness may refer to the 'monetary costs' a household has to incur, more importantly, however, 'it also involves the stress of having a sick person within the household and changes in the family's daily routine because of this situation' (p. 13).

With the above considerations in mind and given the importance of the household as a unit in Indonesia, the cutoff for this indicator follows the 'negative externalities' idea, i.e. *a household is deprived if it has at least one person who has lost their ability to conduct their daily activities for more than the median amount of days lost as recorded within the Susenas 2013 dataset (3 days).*[1]

[1] Goldschmidt (1987) and Attanayake et al. (2001) offer further suggestions on the issue of 'whose' health may most influence the household's level of welfare. These authors propose the exclusion of 'time loss suffered by economically inactive individuals such as those who go to school, job seekers and people with learning difficulties, within the household' (p. 597). By this criterion, the 'cost of illness' calculations only apply to those who work. Rice (1967), however, warns of the dangers of not taking into account the value of services conducted within the household, which are seldom considered as 'economic contributions'. She points to the distortions due to the omission of 'services of the housewife', which may lead to striking conclusions on the effects of illness to productivity that

5.3.4.2 Immunization

The World Health Organization (WHO) (2018b) defines immunization as 'the process whereby a person is made immune or resistant to an infectious disease'. Immunization is 'a proven tool for controlling and eliminating life-threatening infectious diseases, making it 'one of the most cost-effective health investments that governments can undertake' (Crocker-Buque, Mindra, Duncan, & Mounier-Jack, 2017; Grundy et al., 2019; World Health Organization (WHO), 2018b). It is also 'one of the most powerful tools to end preventable child deaths, saving up to three million children a year' (UNICEF, 2012, p. 2).

Within the Delphi, participants understood 'Access to Immunization' to not only express access for children, but also to represent the ability of that child to grow into a healthy and immune adult; immune children will most likely mean immune adults (Bogor City Delphi Non-Participants, 2016; Bogor City Delphi Participants, 2016; Natih, 2016b). The Susenas, however, only has information on 'immunization records for those aged zero to fifty-nine months'; no information is available for the adults within the household (Central Bureau of Statistics, 2013b, pp. 61-63). Particular focus can thus only be given, within this chapter, to children's access to immunization.

With regard to 'child immunization', the *Republic of Indonesia National Law Number 36 Year 2009 on Health* 2009) states that *every child* needs to have a full set of basic immunizations (emphasis added). With this in mind, the cutoff for this indicator is: a *household is considered deprived if it has at least one eligible child who was not given the required set of immunizations.*[1]

5.3.4.3 Safety

Following the Delphi's results, 'health' also goes hand in hand with the 'feeling of safety'. The United Nations (2014) states that 'crime is both a cause and consequence of poverty'. Within academic literature, Berk, Lenihan, and Rossi (1980), among others, state that 'strong correlations between poverty and official measures of crime are perhaps among the most firmly established of social science empirical generalizations' (p. 766). Nationally, the Government of Indonesia acknowledges 'safety' as a fundamental right, which needs to be

are biased toward one sex (Rice, 1967, p. 428). This in mind, this chapter follows Rice (1967) in considering the importance of the health status of all individuals, as potentially influencing the experience of deprivation within the whole household.

[1] The Republic of Indonesia Ministry of Health (2016) defines the 'basic set of immunizations' as including, 'Hepatitis B (given at 0-7 days), Tuberculosis (given at 1 month), three sets of Diphtheria vaccinations (given at 2, 3 and 4 months), four sets of Polio vaccinations (given at 1, 2, 3 and 4 months) and Measles (given at 9 months)'. Within the Susenas there are missing values for this indicator. These are listed within Table 7.10 of the Appendix to Chapter 7.

upheld in order to maintain well-being within its *Republic of Indonesia National Law Number 2 Year 2002 on Policing and Safety 2002)*. Within the Susenas, 'safety' refers to 'safety from crime'; i.e. whether household members experienced 'murder, burglary, rape and/or fraud during the last year' (Central Bureau of Statistics, 2013b, pp. 38-40). As guidance for a possible cutoff for this indicator, Bourguignon (2009) argues that the 'social cost of crime' should not only be measured as 'the monetary value of lost property from thefts, fraud and embezzlements, among others', but also the 'non-monetary costs of pain, suffering and possible disappearance of human capital, in cases where households experience homicides' (p. 186). In the case of the 'experience of crime', arguments for the importance of possible 'negative externalities' may thus, also hold. To illustrate, a rape experienced by one household member, may likely influence the well-being of the whole household. Considering this, the cutoff for this indicator is: *a household is deprived if at least one individual within it experienced crime*.

5.3.5 Child Health and Contraception Indicators

5.3.5.1 Modern health facilities for childbirth

Within the Government of Indonesia's Long Term Development Plans, healthcare for children and mothers, alongside reproductive care are considered major priority areas, which are crucial to maintaining well-being and reducing poverty (Badan Perencanaan Pembangunan Nasional (BAPPENAS), 2014).[1] On this, the Republic of Indonesia Ministry of Health (2014) states that *all* mothers should have access to modern health facilities for childbirth (emphasis added). This in mind, the cutoff for this indicator is: *a household is deprived if at least one member within it does not have access to modern health facilities for childbirth.* [2]

5.3.5.2 Child mortality

Links between child mortality and poverty in Indonesia are investigated within Lanjouw, Pradhan, Saadah, Sayed, and Sparrow (2001) and Baird, Ma, and Ruger (2011), among others. Baird et al. (2011) specifically illustrate the government's commitment toward reducing child mortality in Indonesia and ensuring maternal health. The government pledges that 'infant mortality' is an issue that has to be addressed and no mother should have to go through the death of a child (Republic of Indonesia Ministry of National Development

[1] Modern health facilities in Indonesia are provided through small government clinics (Puskesmas), small government maternal clinics (Posyandu), government and private hospitals, and government and private midwife practices (B. Triyono et al., 2014, p. 12).

[2] The Appendix to Chapter 7 lists missing values in this indicator. Households with no eligible women are considered as non-deprived in this indicator.

Planning, 2004, 2009). With this in mind, the cutoff for this indicator is: *a household is considered as deprived if at least one member has experienced child mortality.* [1]

5.3.5.3 Contraception

The availability of decent healthcare for pregnant mothers and their children goes hand in hand with the provision of family planning and access to contraception. The Government of Indonesia follows the United Nations Population Fund (UNFPA) (2010) in acknowledging that 'women's and adolescents' right to contraceptive information and services are recognized as key human rights'. In addition, decent healthcare (prenatal, natal and throughout adulthood), coupled with the right to contraception, are not only prerequisites to alleviating poverty, they are 'critical to achieving gender equality and ensuring that women can participate as full members of society' (B. Triyono et al., 2014, pp. 1-2; United Nations Population Fund (UNFPA), 2010). The Republic of Indonesia Ministry of Health (2014) pledges that *all* women should have the choice to be able to access contraceptives. With this in mind, the cutoff for this indicator is: *a household is considered deprived if it does not have access to modern contraceptive devices.* [2, 3]

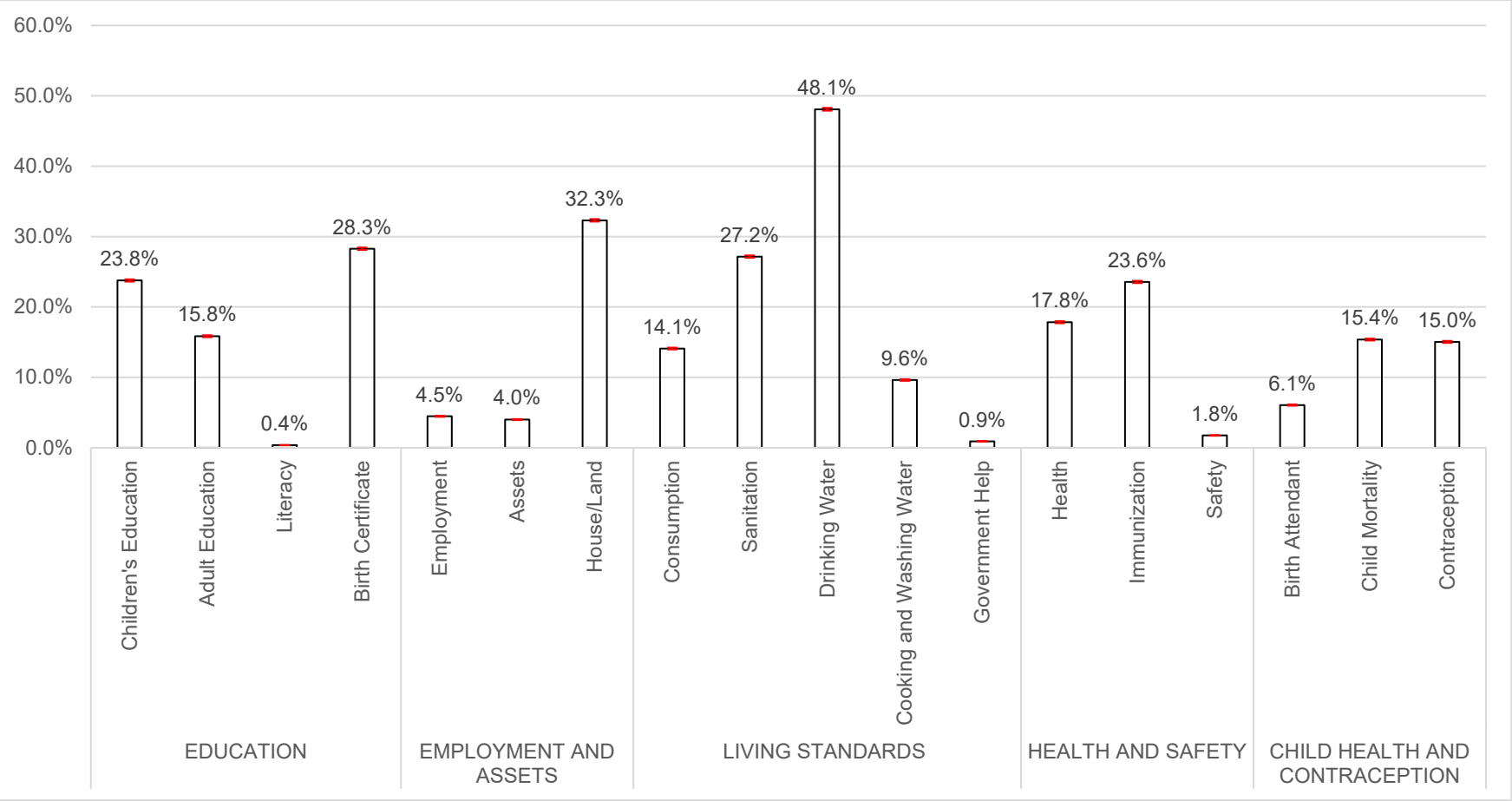
5.4 Results

5.4.1 Indicator Deprivation Rates

Given the indicator cutoffs described above, this section aims to answer Research Question 1: *what are the deprivation rates within each of the indicators chosen by Delphi participants?* Thus, seeking to illustrate the main functions of a dashboard measure, which as was described in the Introduction to this chapter, provides the ability to monitor multiple aspects of poverty individually, alongside uncovering indicators, which could benefit from improvements with regard to how they are measured. Figure 5.1 summarises the results of the above question and suggests two main findings.[4]

[1] Table 7.10 within the Appendix to Chapter 7 lists the number of missing values for this indicator.
[2] The Susenas records not only information on the type of contraceptive device women within the household have access to, it also collects responses on why women chose not to access these devices (Central Bureau of Statistics, 2013b). This in mind, those who do have access to modern contraceptives, however, chose not to access them due to 'wanting more children' or 'religious beliefs', were not considered as deprived.
[3] For Bogor City, there are 7 percent of the sampled households, which have missing values within this indicator.
[4] If terms within literature discussing the Global MPI were followed, the deprivation rates reported within Figure 5.1 would be dubbed as 'raw' headcounts, i.e. the proportion of the total number of households, which are deprived in each indicator (S. Alkire & Santos, 2010a, pp. 10-16).

Figure 5. 1 Percent of Households Deprived in Each Indicator (Bogor City, 2013)[(a)]



[(a)]Note: 95% Confidence Intervals are shown in red. Results are calculated using population weights available within the Susenas.
Source: Author's computations

Firstly, when examining the eighteen indicators, the highest level of deprivation was found within 'Drinking Water'.[1] Referring back to the cutoff for this indicator described within sub-section 5.3.3.3, close to fifty percent of households in Bogor City were shown to not have access to clean sources of drinking water.[2] Drawing upon urban poverty literature, Chaplin (1999), Montgomery (2009), Lüthi, McConville, and Kvarnström (2010), Joshi, Fawcett, and Mannan (2011), McFarlane, Desai, and Graham (2014), among others, find similar evidence describing 'high levels' of lack of access to adequate drinking water and sanitation facilities within cities in the developing world that could lead to health threats, which 'income alone cannot always fend off'. These authors urge for the importance of considering such issues as policy priorities.[3]

The second finding concerns the deprivation rate within the 'Consumption' indicator. Although this indicator is utilised as the main tool to measure and monitor poverty in Indonesia, the level of deprivation within it does not seem to correspond to similar levels of deprivation within the other indicators considered. To further illustrate this fact, tetrachoric correlations were calculated between household deprivation status in 'Consumption' and its deprivation status within the other indicators. Except with the 'Government Help' indicator, which as described within sub-section 5.3.3.5, is anchored to household consumption, modest levels of correlation (an average of 0.38) were found between 'Consumption' and the other indicators examined.[4] This highlights the 'mismatch' between lack of 'consumption expenditure' and deprivations in other social indicators.

Given the main conclusions above, the following discussion briefly describes deprivation rates within each of the indicators considered aiming to further illustrate the function of a dashboard measure and pinpoint possible policy priorities within each of the five dimensions.

[1] The total number of households in Bogor City in 2013 were 248,161; 48% of total households amounts to more than 100,000 households.

[2] Somewhat hand in hand with deprivation within 'Drinking Water', the level of deprivation in 'Sanitation'; i.e. households, which do not have access to adequate toilet facilities, was also found to be relatively high compared to the other seventeen indicators considered. Deprivation within the 'Cooking/Washing Water' was however significantly lower. Looking back at sub-section 5.3.3.4, the cutoff defined for this indicator differs with that set for 'Drinking Water' in that it does not include the additional criteria that the household's 'cooking/washing water source should be at least 10 meters away from waste disposal areas'. The significant difference in deprivation levels between these two indicators, further illustrates the importance of transparency with regard to the decisions made behind the selection of indicator cutoffs.

[3] These issues have become even more crucial within developing nations, particularly in response to the current COVID-19 pandemic. To illustrate, 'policies on social distancing and cleanliness to curb the spread of the virus become less effective in densely populated areas and cities, where access to clean sanitation and water are still unresolved' (Natih, 2020; The Lancet, 2020).

[4] Modest correlations were also found when using data for West Java and Indonesia (an average correlation of 0.23 and 0.3 were found, respectively).

Within the 'Education' dimension, a relatively high number of households (28 percent), were shown to be deprived in the 'Ownership of Birth Certificate' indicator. Cameron (2012) and Bhatia, Ferreira, Barros, and Victora (2017), among others, illustrate the impact of 'weak central registration services' on preventing those who are vulnerable from 'accessing benefits tied to birth registration'. These authors found that 'poor children' were most at risk, as without birth certificates, they may likely be 'excluded from receiving social assistance aiming to support their attendance in school' (Bhatia et al., 2017). In Indonesia, this issue is further exacerbated, as possessing a birth certificate often acts as a pre-requisite before being able to enrol within formal schooling. A relatively high level of deprivation within this indicator may thus, partially contribute toward deprivation within the 'Children's Education' indicator.[1] Looking at Figure 5.1, close to a quarter of households in Bogor City suffer deprivation within 'Children's Education'. This is evident, despite schooling facilities within the city being, for the most part, accessible and adequate.[2]

The next indicator, 'Adult Education', however, exhibits a relatively lower level of deprivation compared to 'Children's Education'. This may point to the likeliness of adults obtaining their 9-year compulsory education qualifications (middle school certificates) later in life through non-formal education centres.[3] M. B. Triyono and Mateeke Moses (2019), among others, highlight the key roles these 'non-formal education centres' play within urban cities in Indonesia, in upgrading skills, providing training and lessening educational inequalities, particularly for those who are vulnerable and who wish to enter the formal job market. The results of the dashboard within Figure 5.1, may thus act to further support the effectiveness of such centres.

Compared with levels within 'Children's and Adult's Education', 'Literacy', exhibits a significantly lower deprivation rate (at only 0.4 percent). This follows the findings within SDG monitoring reports, which highlight Indonesia's achievements in reaching a close to universal literacy rate (at 95.5 percent) across its provinces (Hendarman, Hadiat, & Aruan, 2018). Referring back to sub-section 5.3.1.4, however, how 'Literacy' is measured in the Susenas, i.e. 'self-reported' skills in reading and writing, may not necessarily reflect concrete ability to apply these skills in life. Results of the PISA 2015 and 2019 studies somewhat confirm that

[1] As further reference, a tetrachoric correlation of 0.45 was found between household deprivation status in 'Ownership of Birth Certificate' and 'Children's Education'. This indicates some level of correlation between these two indicators, however, still at a relatively modest level.

[2] The Central Bureau of Statistics Bogor City (2018) report confirms the availability of adequate schooling facilities in Bogor City.

[3] Since the 1970s the Indonesian government has been implementing what is known as the Kejar Paket (catch-up) programmes, which are 'non-formal education equivalency programmes for adults in response to the government's drive for equity in the provision of basic education' (Brennan, 1997; Brock & Symaco, 2011, pp. 58-59; M. B. Triyono & Mateeke Moses, 2019).

this may be the case, as they found that ‘Indonesian school children ranked significantly below their peers in most other Southeast Asian countries’ (OECD, 2018; "PISA 2015 Country Overview of Indonesia," 2015; Tehusijarana, 2019). This is thus an area within the dashboard, which may particularly benefit from measurement improvements.

Within the ‘Employment and Assets’ dimension, ownership of ‘Land/House’ exhibited a relatively high level of deprivation. On the impact of ‘precarious land ownership’, Douglass (1992) highlights the fact that ‘there are few welfare issues not embedded in the questions of land ownership and tenure, particularly within urban settings’ (p. 16). Particularly for urban areas, this is due to land markets being ‘vastly competitive and subject to speculative investments, thus, leading many to resort to crowding into slums and squatter settlements’ (Douglass, 1992, p. 17; Griffin, Khan, & Ickowitz, 2002). Even if ‘rental markets’ emerge to cater to the poor, many ‘occur on the fringe of cities, have precarious tenures and do not have access to public services’ (Ibid). These issues were evident in Bogor City. Slum areas are a recurring challenge; the City has a minimum of seventeen areas, close to three-hundred hectares, which are categorised as shanty towns (Public Relations Office Government of Bogor City, 2017).

Within the two health-related dimensions, ‘Access to Immunization’ is shown to exhibit the highest level of deprivation (at 24 percent deprived households). Referring back to Section 4.7 of Chapter 4, results from the post-Delphi semi-structured interviews, may shed light as to why this may be the case. These interview results found that ‘barriers to the access to immunization’ not only included the need for more ‘health facilities’ and ‘healthcare personnel’, however, more importantly, involved the need to ‘educate the people’ on the importance of vaccines, as ‘some traditional and local beliefs prohibit their use’ (Bogor City Delphi Non-Participants, 2016; Bogor City Delphi Participants, 2016; Gosakti, 2020; Natih, 2016b; Republic of Indonesia Ministry of Health, 2016). This challenge of ‘local beliefs’ is also recognised within literature (Pelčić et al., 2016). Gosakti (2020) in particular, explores the case of Indonesia and highlights the role of a recent *fatwa* (religious sanction), which was placed on vaccinations, on the decrease in the number of children receiving immunizations, particularly for measles.

In addition to the above, the ‘Health’ indicator, measuring the loss of the ability of household members to conduct daily activities, also exhibited a substantial level of deprivation (at close to 20 percent deprived households). As discussed within sub-section 5.3.4, health conditions could act as negative shocks, debilitating a household’s savings and ability to fulfil daily needs. Likewise, is the possible impact of experiencing ‘Child Mortality’ within the household. Interestingly, the relatively high level of deprivation within ‘Child Mortality’ (at 15

percent) is evident even though a lower level of deprivation is found within the 'Birth Attendant' indicator (at 6 percent). On this issue, Montgomery (2009) argues that 'even though urban populations have an advantage in health, as is so often asserted, then it seems that this advantage must be very unequally shared' (p. 399). This may be the case in Bogor City, as even though access to modern health facilities for childbirth may be available, adequate health care for the poor, could still be precarious. In addition to this, the deprivation level within 'Contraception' (at 15 percent deprived households), further underlines the importance of policies, which not only aim to protect mothers through birth, however also, seek to monitor the health of new-born babies, whilst maintaining the well-being of mothers following birth. These 'aftercare' policies should include efforts to increase awareness and access to contraceptive devices, particularly amongst those who are most vulnerable (United Nations Children's Fund (UNICEF), 2020).

In sum, as suggested at the start of this sub-section, the description above acts to demonstrate the key uses of dashboard measures. Firstly, by reporting detailed figures on deprivation levels for *all* the eighteen indicators, the dashboard helps to uncover possible areas needing urgent policy action. Looking back at Figure 5.1, access to adequate drinking water, house/land tenure and the basic citizenship rights to a birth certificate, exhibit the highest levels of deprivation and thus, may act as such priority areas.[1] Secondly, looking back at the way indicators are currently measured within the Susenas, the dashboard reveals areas, which may benefit from advancements in measurement; this was particularly evident for the 'Literacy' indicator. As the Susenas continues to be the main source of information utilised to monitor poverty and welfare in Indonesia, identifying aspects which may benefit from improvement is thus, crucial.

Although as illustrated above, uncovering degrees of deprivation within each indicator enables us to highlight areas where urgent assistance is needed, monitoring progress within a dashboard consisting of *eighteen* indicators, may likely be cumbersome. Following Martin Ravallion (2011b), a set of 'grouped indicators' or 'dimensions', enabling the examination of, as examples, 'Health' or 'Education' poverty, may thus be needed. The following sub-section explores deprivation levels within these 'dimensions'.

[1] A brief note is given on how these results above compare if the geographical unit of analysis is broadened to include the whole of West Java (the province where Bogor City is located) and Indonesia. Table 7.7 of the Appendix to Chapter 7, displays these results. A similar pattern is seen, regardless of geographical unit; i.e. access to clean drinking water, also exhibited the highest level of deprivation, followed by significant deprivation within 'Ownership of Birth Certificate' and 'House/Land Ownership', when West Java and Indonesia are considered. 'Consumption' was also found to be only moderately correlated with the other indicators.

5.4.2 Dimension Deprivation Rates

This sub-section aims to describe results of Research Question 2: *What is the extent of deprivation within each of the five dimensions of poverty chosen by Delphi participants?*

Before this question may be answered, the issue of how one may aggregate across the different indicators within each dimension, needs to be addressed. On this issue, the use of 'data reduction' methods, such as factor analysis and principal component analysis, are often suggested.[1] As described within Chapters 3 and 4, however, within this thesis, the purely statistical exercise offered by such methods, was replaced by the 'clustering' exercise conducted by Delphi participants. Following the results of this exercise, the simplest method to summarise the extent of deprivation in each dimension, is the 'number of deprived indicators within each dimension' (Whelan et al., 2014, p. 184).[2] Anthony B Atkinson (2003) dubs this method the 'counting approach'.[3]

A decision then needs to be made as to a cutoff, which determines the number of deprived indicators within a dimension the household needs to suffer, to constitute deprivation within said dimension. As noted at the start of Section 5.3, I dub these, *dimensional* cutoffs.

When discussing deprivation cutoffs, F. H. G. Ferreira and M. A. Lugo (2013) recall the fundamental choice between the 'union' and 'intersection' approaches (p. 230). Within the 'union' approach, a household is considered as deprived in a dimension if it is deprived in at least one of that dimension's indicators, whilst within the 'intersection approach', the household is deprived in a dimension if it is deprived in all of the dimension's indicators (Anthony B Atkinson, 2003; Bourguignon & Chakravarty, 2003; Duclos et al., 2006, p. 944). Furthermore, Anderson, Crawford, and Leicester (2008), describe the 'union rule' as treating indicators as 'completely non-substitutable', whilst, the 'intersection rule' by contrast, treats them as 'perfectly substitutable' (p. 177). These assumptions, however, may not hold in reality. Thus, a third approach is also suggested within literature; i.e. the 'intermediate' approach, which aims to present an alternative cutoff level that lies between the two extremes of the union and intersection approaches (S Alkire & James Foster, 2011; Duclos et al., 2006).

[1] Nolan and Whelan (1996), among others, illustrate how factor analysis may be applied. Sahn and Stifel (2000) offer an example of its application within the developing world.

[2] Recall at the start of Section 5.3, it is noted that, if a household is deprived in a certain indicator, it is assigned a score of "1" and if it is not deprived, it is given a score of "0".

[3] In addition to the 'ease of interpretation' it offers, the use of the 'counting approach' forgoes the need to assign weights to indicators within each dimension. This significantly reduces the need for 'value judgments' within the process of aggregating indicators into dimensions (Martin Ravallion, 2011b, 2012).

This sub-section applies these three approaches in determining levels of deprivation within the five dimensions chosen by Delphi participants.[1,2] Figure 5.2 summarises the results and suggests two key points.

Firstly, the choice of *dimensional* cutoff plays a significant role in determining the level of deprivation within each dimension. Of further importance is that differing conclusions arise with regard to *which* dimension exhibits the *most* deprivation when the different cutoffs are applied. To illustrate, when the union cutoff is utilised, the highest level of deprivation is evident within the 'Living Standards' dimension, however, when the intersection cutoff is used, 'Employment and Assets' exhibit the highest level of deprivation.

Secondly, although this is the case, for *all* the five dimensions considered, the union approach consistently identifies the *highest* level of deprivation, whilst the intersection approach always records the *lowest*.[3,4] This confirms Whelan et al. (2014)'s argument that even though these two approaches are easy to understand, they can be 'particularly ineffective at separating the poor from the non-poor'; as the union approach tends to identify 'implausibly large numbers' of the poor, while the intersection captures 'small minorities' (p. 18). This then strengthens the need for the 'intermediate' approach.

---

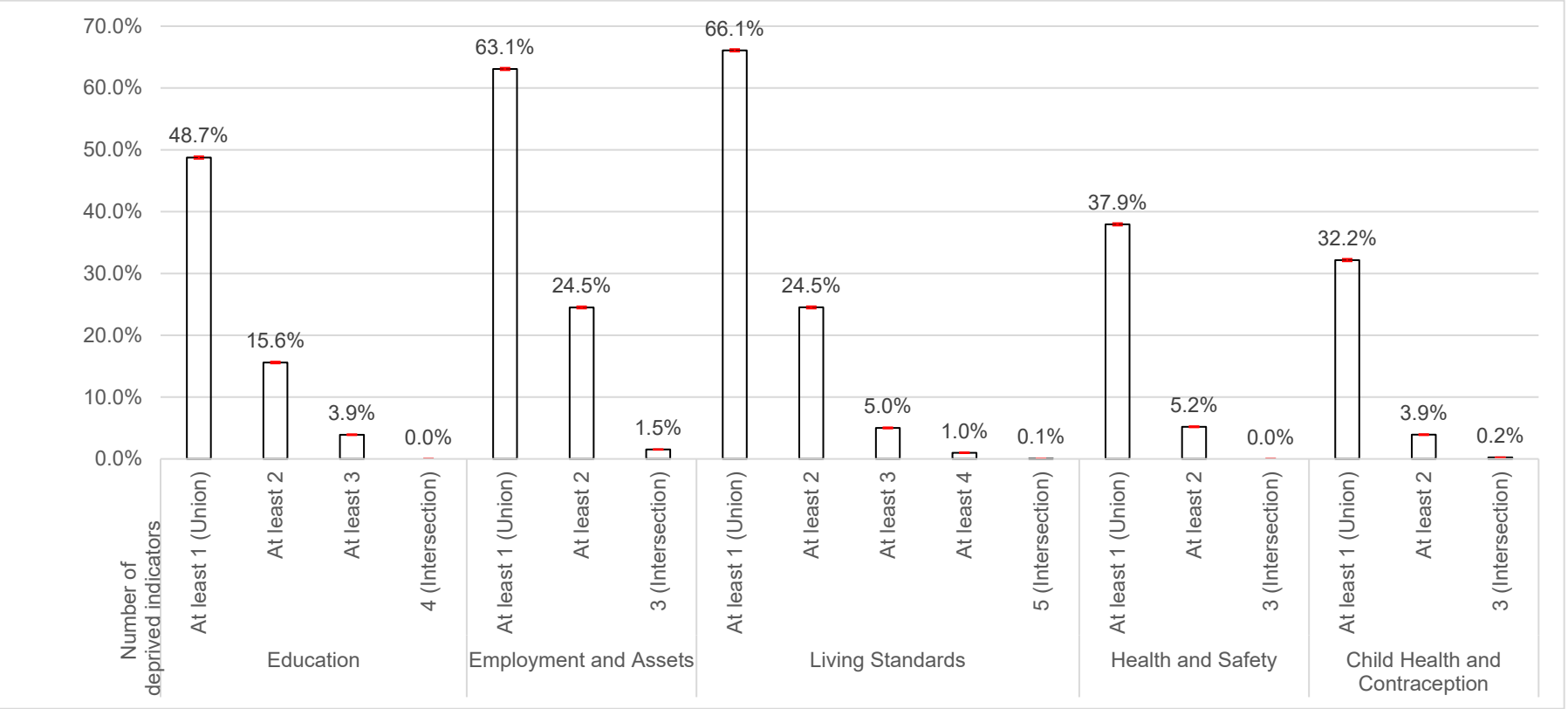
[1] Arguably the choice of dimensional cutoff could have been relegated to Delphi participants; however, as described within Chapter 3, issues of fatigue and the length of the questionnaire were significant factors, which prompted the conscious choice to rely on the 'counting approach' suggested within this chapter. Results of the semi-structured robustness interviews conducted following the Delphi, then supported the decision to test a number of dimensional cutoffs; i.e. union, intermediate and intersection, as, unanimity was not achieved amongst interviewees, on which approach was most adequate. Given this lack of unanimity, when time and resources permit, additional data collection could be conducted to discover whether, at least, agreement on a range of dimensional cutoffs, exists. This range could then be applied to define maximum and minimum values, which in turn may contribute to the creation of what Chiaperro-Martinetti (1994) dub as 'membership functions' and the use of 'fuzzy set' theory in the measurement of poverty (Chiaperro-Martinetti, 1994, p. 371; Qizilbash, 2006; Zadeh, 1965).

[2] Alternatively, Whelan et al. (2014) propose selecting cutoffs for each dimension, which will lead to deprivation rates that are similar to the percent deprived identified by the '60 percent of median income at-risk-of-poverty' threshold. They chose this threshold as a benchmark as it is considered to be widely accepted. Within Indonesia, the 'Consumption' measure may act as such a benchmark, as it is the official measure utilised by the government to identify the poor. As results within Chapters 2, 3 and 4 of this thesis show, however, its acceptance as a sole benchmark for poverty identification, continues to be uncertain (Syebubakar, 2014). This in mind, utilising the method suggested within Whelan et al. (2014) was less possible, as a widely accepted standard was not available.

[3] As comparison, S. Alkire (2011) calculated the rates of poverty if the 'union' and 'intersection' approaches were used to define poverty when the Global MPI is utilised for 104 developing countries. They found an average poverty rate of 58 percent when the 'union' approach was applied and an average rate of zero percent when the 'intersection' approach was utilised. Furthermore, Whelan et al. (2014), calculating the proportion of the poor for EU-SILC countries using these two approaches, discovered similar conclusions; i.e. the 'union' identifies extremely high levels of poverty, whilst the 'intersection' approach identifies close to zero poverty.

[4] This is also the case when the geographical unit is expanded to include the whole of West Java and Indonesia. Table 5.3 of the Appendix to this chapter illustrates this.

Figure 5. 2 Proportion of Deprived Households in Each Dimension According to Number of Deprived Indicators within Dimension (Bogor City, 2013)[(a)(b)]



Note: [(a)] 95% Confidence Intervals are shown in red. Results are calculated using population weights available in the Susenas.
[(b)] Union approach: a household is deprived in a dimension, if it is deprived in at least one indicator used as a proxy of that dimension
Intersection approach: a household is deprived in a dimension, if it is deprived in all indicators used to proxy that dimension
Source: Author's computations

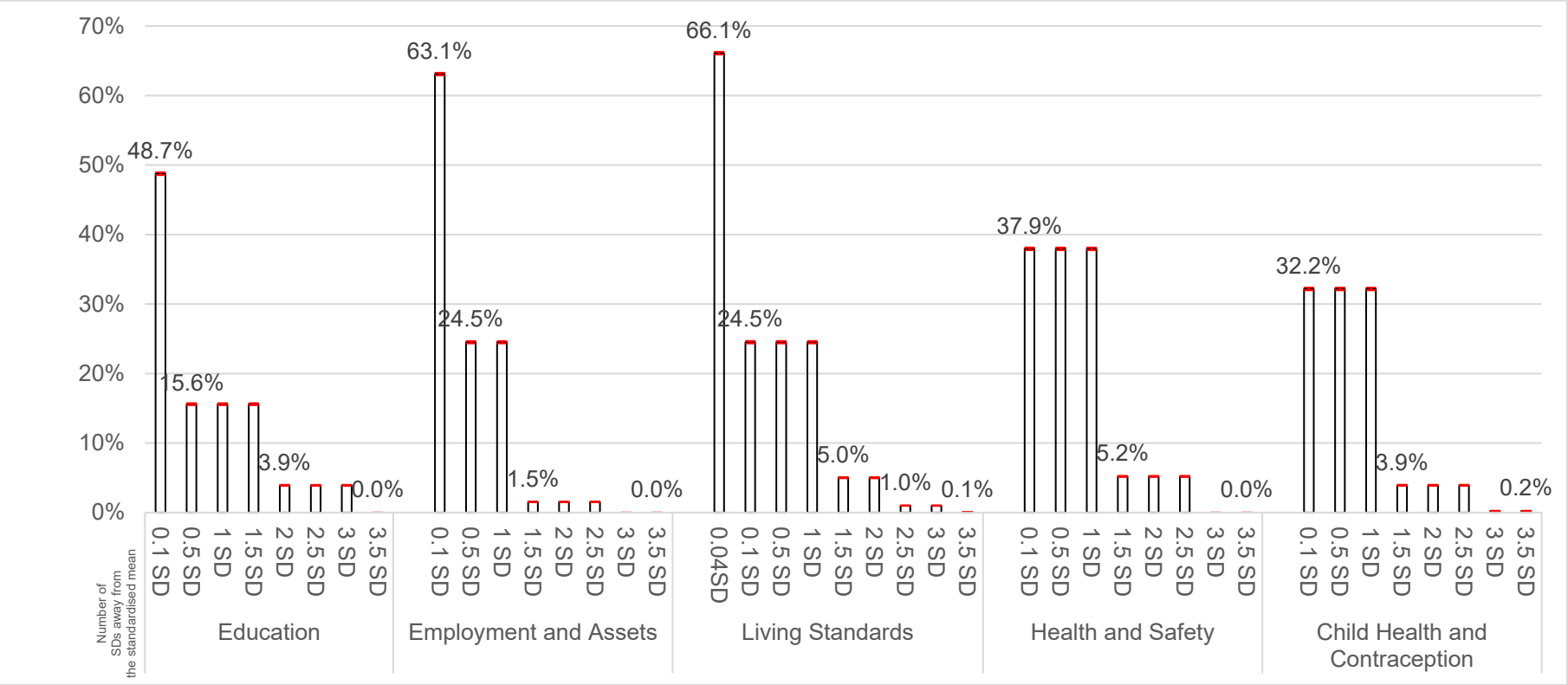
Within Figure 5.2, deprivation levels for the intermediate approach, are exhibited by the cutoffs that lie between the union and intersection cutoffs. A dilemma however arises with regard to the comparability of these 'intermediate' cutoffs across the dimensions considered, as each dimension utilises a different number of indicators to proxy deprivation within them. To demonstrate, the 'Education' dimension has four indicators, however, the 'Living Standards' dimension has five, whilst the 'Employment and Assets', 'Health and Safety' and 'Child Health and Contraception' dimensions have three.

In an attempt to secure comparability across the five dimensions, before setting the intermediate cutoffs, households' total deprivation within each dimension, were first standardised; i.e. the means of the distribution of total deprivation within each dimension were set to zero and their standard deviations were adjusted to equal one. The intermediate dimensional cutoff was then defined as, a household is deprived in a dimension, *if its standardised total dimensional deprivation is equal to or above a certain amount of standard deviations away from the standardised mean*. Figure 5.3 illustrates the results of this exercise.

As can be seen, within Figure 5.3, several intermediate cutoffs are tested, i.e. ranging from -0.04 to 3.5 standard deviations away from the mean.[1] This range encompasses the deprivation levels, which result, if the union and intersection approaches were utilised. To illustrate, for the 'Education', 'Employment and Assets', 'Health and Safety' and 'Child Health and Contraception' dimensions, applying the dimensional cutoff of: *a household is deprived if it has a standardised total indicator deprivation of equal to or larger than 0.1 standard deviations away from the mean*, results in a similar percentage of deprived households, to that, if the union approach was utilised. For the 'Living Standards' dimension, the cutoff of: *equal to or above 0.04 standard deviations away from the mean*, exhibits similar deprivation levels compared to when the union approach is utilised for this dimension. For the 'Education', 'Living Standards', 'Health and Safety' and 'Child Health and Contraception' dimensions, the cutoff of: *equal to or above 3.5 standard deviations away from the mean*, then reports similar levels of deprivation to when the intersection approach is applied. For the 'Employment and Assets' dimension, a similar deprivation rate to when the intersection approach is used, is found when the cutoff is set at *2.5 standard deviations away from the mean.* Considering the above, for the intermediate approach, this chapter examines the cutoffs, which lie between those that result in similar deprivation rates to when the union and intersection approaches are applied.

[1] Following Duclos et al. (2006), 'in theory, the entire area of possible intermediate cutoffs should be tested, however, it is more practical to choose a selection of cutoffs to examine' (p. 959).

Figure 5. 3 Proportion of Deprived Households in Each Dimension According to the Level of (Standardized) Total Household Dimensional Deprivation (Bogor City, 2013)[a][b][c][d]



Note: [a] Total household dimensional deprivation was first standardized to have a mean of 0 and standard deviation of 1, before these percentages were calculated.
[b] SD: standard deviation (standardized to equal 1).
[c] Dimensional cutoff: a household is deprived if its standardized total dimensional deprivation (sum of indicator deprivation within said dimension) is equal or below/above x standard deviations away from the mean.
[d] 95% Confidence Intervals are shown in red. Results were calculated using population weights available in the Susenas.
Source: Author's computations.

Looking at Figure 5.3, it can be seen that different intermediate cutoffs, lead to differences with regard to which dimensions exhibits the most deprivation. Table 5.2 further explores this point. To assess whether comparisons of levels of deprivation are robust to the choice of cutoffs, Duclos et al. (2006) suggest the use of 'multivariate stochastic dominance' analysis. Although conducting detailed dominance analysis is beyond the scope of this chapter, I borrow the logic of this method to study the results within Table 5.2.[1]

Table 5. 2 Dimension Ranks when Different Cutoffs are Applied (Bogor City, 2013)[(a)]

Dimension	0.1 SD	0.5 SD	1 SD	1.5 SD	2 SD	2.5 SD	3 SD	3.5SD
Education	3	5	5	1	4	3	1	3
Employment and Assets	2	4	4	2	5	4	4	4
Living Standards	1	3	3	3	2	5	2	2
Health and Safety	4	1	1	2	1	1	5	5
Child Health and Contraception	5	2	2	4	3	2	3	1

Note: [(a)] Dimensions are assigned ranks 1-5, according to the percentage of deprived households within the said dimension. Rank 1 indicates that a dimension has the highest percentage of deprived households, when that specific cutoff is applied.

Source: Author's computations

In Table 5.2, dimensions are assigned ranks according to the proportion of deprived households found within them when a specific intermediate cutoff is applied. These ranks range from "1" to "5"; with "1" indicating that a dimension has the highest number of deprived households and "5" indicating it has the lowest level of deprivation, when a specific intermediate cutoff is applied. This analysis suggests two key findings.

Firstly, dimensions are ranked differently when different cutoffs were applied. To illustrate, using the '0.5 SD' cutoff, 'Health and Safety' is ranked first, whilst 'Education' is ranked fifth. These ranks hold when the '1 SD' cutoff is applied, however, differ significantly when the '1.5 SD' cutoff is used, where 'Education' is ranked first, rather than last.

Secondly, referring back to the *indicator* deprivation levels summarised within Figure 5.1, even though the highest level of deprivation was found to be evident within 'Drinking Water', an indicator included in the 'Living Standards' dimension, looking at Table 5.2, this

[1] Anthony B. Atkinson (1987), James E Foster and Anthony F Shorrocks (1988) and James Eric Foster and Anthony F Shorrocks (1988), among others, offer further discussions on 'dominance analysis'.

dimension is ranked as having the highest level of deprivation *only* if the '0.1 SD' cutpff is applied.[1] When the other cutoffs were utilised, 'Living Standards', was never ranked as exhibiting the highest level of deprivation. Compared to 'Living Standards', if *dimension* deprivation levels were considered, 'Health and Safety' seems more likely to act as a priority area, as this dimension was ranked first for a majority of the intermediate cutoffs considered. Differences in conclusions, which may arise when examining indicators separately, versus considering the joint distribution of deprivations across various indicators within each dimension, highlights the importance of considering both indicator and dimensional deprivation levels when analysing multidimensional poverty.[2]

In sum, borrowing jargon from multivariate stochastic dominance analysis, results described within Table 5.2, clarify that the level of deprivation within dimensions is not robust to the choice of dimensional cutoff, i.e. different deprivation criteria leads to different conclusions with regard to the dimensions with the most challenging levels of deprivation. This in turn, illustrates the need for care when it comes to the selection of these deprivation thresholds. In addition, it demonstrates that, far from that suggested by its proponents, a dashboard, similar to a composite measure, is not free from significant value judgments. The challenge faced when measuring poverty is thus, not the need for value judgments and the level of uncertainty they may bring, but how best to explicitly present the reasoning behind these choices. The fact that measuring poverty may require judgment, should thus, not be a reason to dismiss the importance of such an exercise.[3] This then paves the way for the need for the composite measure, the subject of the next chapter, Chapter 6, of this thesis.

5.5 Conclusion

This chapter argued for the importance of both dashboard and composite measures when attempting to assess the extent of poverty within a given society. Literature on this issue was drawn upon to highlight the strengths of both approaches. Empirical evidence supporting this argument gathered through the Delphi exercise illustrated within Chapters 3 and 4 of this thesis, was also referred to.

[1] As noted above this cutoff results in similar levels of deprivation in 'Living Standards' to that found when the union approach is applied.

[2] Duclos et al. (2006) argue for the importance of considering joint distributions when analysing poverty as it may be possible that 'deprivation in one indicator may affect how much deprivation declines if there is an improvement in other indicators' (p. 945). This fact also strengthens the need for the composite measure, proposed within Chapter 6 of this thesis, which considers deprivation levels in all the eighteen indicators, within its calculation.

[3] To quote Amartya Sen (1979b), 'there is very little alternative to accepting the element of arbitrariness in the description of poverty and making that element as explicit as possible' (p. 288).

The above in mind, this chapter focused on the construction of the dashboard measure. When creating such a measure, decisions needed to be made with regard to three main aspects; i.e. (1) the types of indicators and dimensions to include within the measure, (2) the indicator cutoffs, the level of deprivation a household needs to experience to be considered as deprived within a specific indicator and (3) the dimensional cutoffs, the number of deprived indicators a household needs to experience to be considered as deprived within a specific dimension. Addressing these aspects enabled analysis of the extent of deprivation evident within each indicator and dimension; the two research questions this chapter aimed to tackle. Before summarising the results of this chapter, a brief overview of the above three aspects is presented below.

With regard to the first aspect, i.e. the selection of indicators and dimensions, this was based solely on the results of the Delphi exercise described within Chapters 3 and 4 of this thesis. For the second aspect, i.e. indicator cutoffs, this chapter argued for the use of existing minimum accepted welfare criteria found within international and national policy standards to guide their selection. For the final aspect, i.e. dimensional cutoffs, three approaches were applied. The widely recognised ‘union’ and ‘intersection’ approaches were utilised, alongside a range of ‘intermediate’ dimensional cutoffs. The effects of the use of these different cutoffs on the level of deprivation within each dimension, was then analysed. With decisions on these three aspects in mind, this chapter uncovered two main conclusions.

Firstly, with regard to the first research question, the dashboard presented a detailed report on the extent of deprivation within all the eighteen indicators considered as important by Delphi participants. When analysing in further detail, this chapter found that the extent of indicator deprivation in the case study area chosen for analysis within this thesis, mirrored that evident within existing literature discussing urban poverty in developing countries. Indicators of adequate ‘Drinking Water’, alongside decent ‘Sanitation’ facilities were shown to exhibit the highest levels of deprivation. Following these indicators, deprivation levels within the ‘Land/House Ownership’ and ‘Birth Certificate Ownership’ also exhibited relatively high levels of deprivation. This points to the possible importance of considering such areas as priority areas, when attempting to address poverty. As described within the Appendix of this chapter this trend was also present when data from West Java and Indonesia were considered. This strengthens arguments made within existing literature, which suggest rapid urbanisation in Indonesia and highlights the increasing importance for policy in Indonesia to place greater focus on welfare issues faced by the urban poor (Lewis, 2014).

Secondly, to address the second research question, i.e. dimensional deprivation rates; confirming that found within existing literature, the ‘union’ approach identified the largest

levels of deprivation within all dimensions, whilst the ‘intersection’ approach, identified the lowest levels for all the five dimensions examined. These extreme values strengthened the need for applying the ‘intermediate’ approach to setting dimensional cutoffs. This in mind, a range of ‘intermediate’ cutoffs were tested and dimensional ranks, based on the number of deprived households within the said dimension when different cutoffs were applied, were determined. This exercise found that levels of deprivation within dimensions were highly sensitive to the dimensional cutoff chosen. Robust conclusions with regard to which dimension exhibited the highest number of deprived households, could thus, not be made.

The above key findings in mind, this chapter concludes that as argued for within Ravallion (2011a), a dashboard measure indeed, offers a simplified way of presenting data on deprivation rates for multiple indicators and dimensions. Furthermore, this form of measurement enables the identification of priority areas, drawing the attention of policy makers to aspects of poverty where significant intervention may be needed. Considering the impact of the choice of cutoffs on deprivation rates, dashboards are, however, not free from value judgments. As with any form of poverty measure, clarity and discussion with regard to the reasons behind the choices made within measurement are crucial.

Although as noted above, this chapter illustrates the merits of using dashboards in exploring levels of deprivation for each indicator and dimension, this form of measurement falls short of explaining which households are poor *across* these various aspects of poverty. Where a dashboard cannot offer the ability to construct such ‘complete orderings’ of those experiencing poverty, a composite measure may act as a supplement. This is the subject of the next empirical chapter of this thesis, Chapter 6.

Chapter 6: A Composite Multidimensional Poverty Index based on Delphi Results

A composite measure is a necessary supplement to a broad dashboard of indicators, enabling the comparisons of poverty rates across our sub-districts

(Head of the Bogor City District Development Body Statistics Division, 2016)

6.1 Introduction

This chapter forms the second part of tackling Amartya Sen (1976a)'s challenge of addressing how best to measure poverty given its multidimensional nature. As argued for within Chapter 5, this thesis proposes the use of two approaches, i.e. the dashboard method, which was described within Chapter 5 and the composite method, which is the subject of this current chapter.

Section 5.2 of Chapter 5 highlighted the ability of composite measures to produce 'complete orderings' of households and countries, allowing comparisons of the extent of poverty across time and space. In addition, composites enable the assessment of the experience of joint deprivations within multiple indicators of poverty. Despite these strengths, the methodological arbitrariness by which current composite measures are constructed, has led to numerous critiques against their use. This chapter aims to contribute to the overall improvement of these existing measures. Particular focus is given to one crucial source of technical weakness within composites, what Saltelli (2007) dubs as the 'key objection' to aggregation, i.e. the arbitrary nature of the *weighting* process by which multiple indicators of poverty are combined (emphasis added, p. 68).

To address the above concern, this chapter proposes a 'new' composite measure, which utilises dimensions, indicators and weights selected by Delphi participants. I dub this 'new' measure, the Delphi Composite Index of Multidimensional Poverty, hereafter the Delphi-CIMP. By constructing this measure, this chapter seeks to address two research questions. Firstly, it aims to suggest a novel and robust method of incorporating opinions from a Delphi exercise within a composite measure of poverty. As noted above, specific attention is placed on how views concerning the *weights* for these multiple aspects of poverty could be utilised. This in mind, the first research question asks: *Based on Delphi results, what are the optimal set of weights to utilise within the Delphi-CIMP?* I label these Delphi-derived weights, 'expert' weights.

As noted within Section 5.2 of Chapter 5, the use of 'equal' weights within current composite measures, has been a source of substantial debate. The 'expert' weights suggested within this

chapter aim to offer an alternative to these commonly applied weights. Considering this, the second research question aims to examine the impact of the use of these two weighting schemes, 'expert' vs. 'equal'. Within this chapter, particular scrutiny is given to the use of 'equal nested' weights, the weighting scheme utilised within the Global and Adjusted MPIs. Research question two thus, investigates: *To what extent does the Delphi-CIMP identify different sizes and composition of the poor if 'expert' weights are applied as opposed to 'equal nested' weights?* Results for this question aim to shed light as to how sensitive the newly developed composite measure is to the choice of weights within its construction.

As with Chapter 5, calculations within this chapter are conducted using data from the case study area chosen, i.e. Bogor City, available within the Susenas 2013.[1] The household is again taken as the unit of analysis. The rest of this chapter is organised as follows: Section 6.2 provides a brief literature review of methodological guidelines that should be considered when constructing a composite index of poverty. Following these guidelines, Section 6.3 describes the methodological choices adopted to build the Delphi-CIMP. Section 6.4 presents the results of this chapter. Section 6.5 then concludes with a brief summary of findings and possible policy implications.

[1] As noted within Chapters 5 and the Appendix to Chapter 7, a number of missing values exist with regard to three indicators included within the Delphi-CIMP, i.e. 'Drinking Water', 'Child Mortality' and 'Contraception. Following Global MPI guidelines, household which have missing values in any indicator are dropped from the final analytical sample. Sabina Alkire et al. (2019) note that for the Global MPI, given its 33.33 percent poverty cutoff, household who are deprived in 33.33 percent of the Global MPI's indicators despite having missing values could be calculated as 'poor'. For the case of the Delphi-CIMP, as will be described within this chapter, in order to determine whether a household is poor, the said household's 'total weighted deprivation' was needed. The poverty cutoff was then set at several standard deviations away from the standardised mean of these weighted deprivations. Sabina Alkire et al. (2019) suggestion above to reduce the number of missing households, thus could not be applied. This in mind and considering the missing values for the three Delphi-CIMP indicators stated above, the total retained sample for Indonesia was thus 81 percent, for West Java 81 percent and for Bogor City 84 percent. Following Global MPI guidelines, the retained sample size for Indonesia is below the 85 percent threshold for preventing possible bias. This in mind, steps for 'bias analysis' within the Global MPI guidelines were followed. Firstly, the major source of sample reduction was identified. For the Delphi-CIMP this was the 'Contraception' and 'Drinking Water' indicators. For both these indicators, the sample was then divided into two groups; i.e. those with complete information on all indicators and those with missing information within a specific indicator. The second step then involved checking the headcount ratios of deprivation within the other indicators, for these two groups. Following Sabina Alkire et al. (2019), if there is a systematic and statistically significant difference between these headcounts within the two groups (at a significance level of 1%), the sample does not fulfil the criteria for the bias analysis test. Carrying out this test for the national, provincial and district levels within this chapter, for all these levels, no significant difference was found. Thus, the sample fulfils the bias analysis criteria.

6.2 Review of Existing Methodological Guidelines

Following Chiappero-Martinetti and von Jacobi (2012), to construct composite measures, a number of essential steps need to be taken; these steps range from 'the choice of indicators to include, their transformation and combination, to the specification of their weights' (p. 71). Decisions made within each step require careful consideration, as they may impact how the final measure quantifies poverty, which in turn could significantly influence policy. The sub-sections below discuss key methodological recommendations within these steps.

6.2.1 Step 1: Theoretical Framework, Dimensions and Indicators

Nolan and Whelan (1996) highlight the importance of 'linking empirical measurement' to 'how poverty is defined'. When constructing composite measures of poverty, a sound theoretical framework should thus be outlined before measurement takes place. This theoretical framework would then guide the selection of the measure's components, i.e. dimensions, indicators and weights. On the selection of these components, the OECD's (2008) "Handbook on Constructing Composite Indices" notes the importance of 'involving experts and stakeholders' in order to take into account their 'multiple viewpoints' and to 'increase the robustness of the set of indicators chosen' (p. 22). As argued for within Chapters 3 and 4, the Delphi aimed to act as a way of soliciting these 'expert' views.

6.2.2 Step 2: Transformation of Indicators

Once a degree of consensus has been achieved on the definition and possible components to include within a composite, before aggregating and computing the measure, 'indicators that are measured in different units must be transformed into the same unit' (Michaela Saisana & Tarantola, 2002, p. 9). Transformation methods enable one to 'trim raw data distributions, remove the scale effect from all indicators simultaneously and avoid having extreme values that overly dominate results' (Nardo et al., 2005, p. 11). Jacobs, Smith, and Goddard (2004) and Nardo et al. (2005), in particular, list a number of transformation methods, which may be utilised. To select the most appropriate method, these authors argue that this choice should be based on the 'objectives' of the composite and 'data properties' of its indicators.

6.2.3 Step 3: Weighting and Aggregation

Following the 'transformation' step outlined above, once indicators are commensurate with each other, 'central' to the construction of a composite index is the need to combine the different

indicators in a meaningful way (Nardo et al., 2005, p. 11). This step 'implies a decision on the weighting model and aggregation procedure' (Ibid). I start with a discussion on the meaning of weights and how this connects to the aggregation function chosen to combine indicators within composites.[1]

6.2.3.1 The Meaning of Weights and the Choice of Aggregation Function

Weights represent 'important value judgments about the (vague) notion of well-being; any choice of weights should thus, be open to questioning, made explicit, comprehensible and open to public scrutiny' (S. Anand & Sen, 1997, p. 6; Decancq & Lugo, 2008, p. 2). To further describe the role of weights within composite measures, S. Anand and Sen (1997) suggest the use of partial derivatives and propose the examination of Marginal Rates of Substitution, hereafter MRSs, between indicators. The MRS is defined as the amount a household is willing to give up in one indicator, to gain an extra unit in another, while maintaining the same level of wellbeing (Decancq & Lugo, 2013, p. 16).[2] Given how indicators are weighed, the MRS enables one to assess how the composite index may react to changes in deprivation levels within its indicators. It also illustrates the rate of trade-off between indicators, which as Martin Ravallion (2011b) points out, is simply the 'marginal weight' of deprivation within one indicator relative to that within another (p. 11).

The importance of recognising that 'setting weights' also carries assumptions on possible trade-offs between indicators is particularly so when 'linear' or 'geometric' aggregation functions are employed to compute the composite. The use of these functional forms to combine indicators implies 'the idea that deficits in one indicator may be offset by surpluses in another' (Nardo et al., 2005, p. 12). The meaning of weights within composites, thus, also, goes hand in hand with the choice of aggregation function. Nardo et al. (2005) suggest that in general, the choice of aggregation function comes within three broad categories, i.e. 'linear aggregation (where indicators are summed up), geometric aggregation, (where indicators are multiplied) and aggregation using non-linear techniques' (p. 12). For its parsimony and ease of interpretation, however, academic literature on multidimensional poverty, to date, has focused mostly on the

[1] Referring back to Chapters 3 and 4, participants of the Delphi were asked to assign both dimension and indicator weights. To simplify the discussion within this current chapter, particular focus is given to indicator weights, to underline the fact that the 'final' composite, aims to aggregate across all the eighteen indicators chosen by participants.

[2] Following Martin Ravallion (2011b), while the term 'MRS' is more commonly used with reference to the trade-off built into an individual's utility function, it can be defined for any suitably analytic function of two or more variables' (p. 11).

use of 'linear' aggregation' as a way of combining deprivations (S. Alkire & Santos, 2010a; Martin Ravallion, 2011b).[1] When linear aggregation functions are utilised, 'the higher the indicator-weight, the better a household will perform if it fares better in that indicator; in other words, indicator weights determine the 'contribution of that indicator to the experience of poverty, alongside, as discussed above, forming part of the trade-offs between indicators' (Decancq & Lugo, 2008, p. 14).[2]

6.2.3.2 Selecting Weights within Composite Measures

Decancq and Lugo (2008) suggest a number of general procedures, which could be utilised to determine weights within composite indices of poverty. These procedures range from the use of (1) equal weights, (2) data driven methods, i.e. frequency-based methods, benefit-of-the-doubt procedures, statistical and regression-based weights and (3) normative weights. In the interests of conciseness, this sub-section focuses on a theoretical review of the weighting schemes most relevant to this chapter, i.e. 'equal', 'normative' and 'benefit-of-the-doubt' weights.[3]

As discussed within Section 5.2 of Chapter 5, equal weights are the most commonly used approach to weighting in multidimensional measures.[4] This approach offers the most straightforward and simple way of setting weights, supporting ease of interpretation and requiring the least amount of time and resources behind its determination (S. Alkire & Santos, 2010a, p. 18).[5] As Decancq and Lugo (2008) note however, 'despite its popularity, equal weighting is far from uncontroversial' (p. 14). Indeed, the questionable assumptions behind the use of 'equal weights' has become a major source of criticism, which may outweigh the simplicity this approach hopes to offer. Considering this, as noted within Chapter 1, Carlucci and Pisani (1995) and Decancq and Lugo (2008), among others, argue that as an alternative to equal weights, 'normative weights' tend to be 'closer related to the meaning of weights as trade-offs and thus, may be expected to lead to more reasonable results' (Carlucci & Pisani, 1995; Decancq & Lugo, 2008). Selecting 'normative weights', however, presents significant challenges

[1] Linear aggregation is utilised within the Human Development Index (HDI), the Global and Adjusted MPIs (S. Alkire & Santos, 2010a; Decancq & Lugo, 2008; Martin Ravallion, 2011b).

[2] Weights can be interpreted 'directly as trade-offs, only when certain additional assumptions are applied within the aggregation function'; i.e. when 'perfect substitutability between indicators' is assumed (Decancq & Lugo, 2008, p. 14).

[3] Decancq and Lugo (2008) offer a review of the other possible methods of deriving weights.

[4] Indeed, a variant of this method; i.e. the 'equal nested weighting' scheme, is used within the Global and Adjusted MPIs explored in Chapter 2.

[5] On this issue, T. Atkinson, Cantillon, Marlier, and Nolan (2002), note that 'the interpretation of the set of indicators is greatly eased where the individual components have degrees of importance that, while not necessarily exactly equal, are not grossly different' (p. 25).

as synthetizing different views is often an arduous task. The 'benefit-of-the-doubt' approach, applied through a method of deriving weights known as Data Envelopment Analysis, hereafter DEA, offers a way of addressing these challenges. In order to describe key aspects of the 'benefit-of-the-doubt' approach, details on the steps taken to compute weights within the DEA are first outlined below.

Weights within the DEA, are determined endogenously, i.e. their values are dependent on the 'observed performances' of the entities analysed (Laurens Cherchye & Kuosmanen, 2004; L. Cherchye et al., 2008; Melyn & Moesen, 1991; Nardo et al., 2005).[1] These weights are estimated using a 'linear optimisation model to the best advantage for each entity so as to maximise its relative performance' (Despotis, 2005, p. 969). Within this chapter, the 'performance of an entity' refers to a 'household's deprivation level' and is measured by the weighted sum of the said household's total deprivation within the indicators included in the composite measure (Mahlberg & Obersteiner, 2001, pp. 5-7).

To solve the aforementioned 'linear optimisation' model, a 'best-practice frontier', is first constructed.[2] This 'frontier' may be determined through two approaches: (1) computation or (2) setting a 'target' frontier level. For the first approach, the 'frontier' is derived within the DEA itself by computing a 'linear combination' of the 'best performing entities' as observed in the dataset analysed. For the purposes of this chapter, this method would thus involve determining a linear combination of households with the best outcomes with regard to their deprivation levels within each indicator. The second approach determines the 'frontier' by setting a 'target' level, which aims to reflect the most 'preferred combination' of indicator deprivation rates. Ideally, this 'target' frontier, should be based on 'expert' preferences or alternatively, policy documents may be consulted (Korhonen, Tainio, & Wallenius, 2001; Nardo et al., 2005, pp. 59-61). Weights within the composite are then calculated by minimising a household's 'distance', i.e. its actual achievement with regard to its total weighted deprivation, to these previously determined 'frontiers'.[3] Due to this minimisation process, 'higher' weights are assigned to indicators where a

[1] Data Envelopment Analysis (DEA) originates from 'efficiency analysis'. This method offers a non-parametric way of measuring the 'efficiency level of a decision-making unit as it converts the multiple resources it has (its inputs) into the many goods and services it produces (its outputs)' (Banker et al., 1994; Charnes, Cooper, & Rhodes, 1978).

[2] Referring to literature on industrial efficiency, the 'best-practice frontier', a notion first introduced within Farrell (1957), 'delineates the technological limits of what a decision-making unit can achieve through utilising a given level of resources' (Mahlberg & Obersteiner, 2001, p. 5).

[3] Units or entities are classified into 'well-performing' if they are 'at the frontier' or 'worse performing' if they are 'below the frontier' (Mahlberg & Obersteiner, 2001, pp. 5-6).

household has ‘superior’ performance and ‘lower’ weights are assigned to areas in which its performance is ‘relatively worse’ (Mahlberg & Obersteiner, 2001, pp. 5-7). This is where the ‘benefit-of-the-doubt’ approach ‘logic’ comes into play. The core idea behind ‘benefit-of-the-doubt’ weighting is that, indicators, which a country performs relatively well in should be assigned a higher weight’, as good performance within particular indicators reflect the higher importance placed on policies concerning the said indicators (Laurens Cherchye et al., 2007; Lovell et al., 1994). Considering this, weights, which result from the ‘benefit-of-the-doubt’ approach are often dubbed as ‘most favourable’ weights. Compared to the use of ‘equal’ weights, ‘most favourable’ weights, are allowed to vary across the different indicators. This variation is assumed to reflect policy priorities and reveal possible trade-offs faced when different policies are implemented (Nardo et al., 2005, p. 63). Laurens Cherchye et al. (2007) argue for the importance of this approach as ‘composite indices are, for the most part, utilised to monitor policy’ (p. 141).

Furthermore, for the purposes of this thesis, L. Cherchye et al. (2008) highlight the desirability of combining results from a Delphi study with the ‘benefit-of-the-doubt’ approach. This is particularly true as the DEA based optimisation model, utilised within empirical applications of the ‘benefit-of-the-doubt’ approach, are ‘particularly apt’ when dealing with the ‘challenging reality’ of differing opinions on weights within composite indices (Ibid). In addition to this, setting ‘most favourable weights’, through applying the ‘benefit-of-the-doubt’ approach, aims to mitigate possible claims that ‘unfair weighting’ was employed as ‘any other scheme would only worsen the position of the specific area of interest’ (Laurens Cherchye, Moesen, & Van Puyenbroeck, 2004, p. 930).

6.2.4 Step 4: Sensitivity and Uncertainty Analysis

Following the OECD’s (2008) “Handbook on Composite Measures”, uncertainty analysis focuses on how robust the composite’s results are to the choices made within its construction, i.e. the types and number of indicators, transformation method and aggregation function utilised. Sensitivity analysis then measures the level of uncertainty that could be reduced if a particular source of uncertainty is removed from the composite. These checks provide information on whether the results of the composite represents meaningful constructs and reduces the possibility that the measure may send misleading policy messages’ (Nardo et al., 2005, p. 13).[1]

[1] Conscious of space, I refer the reader to M. Saisana, Saltelli, and Tarantola (2005) for a comprehensive guide to conducting these robustness checks.

6.3 Steps in the Construction of the Delphi-CIMP

With the above guidelines in mind, this section focuses on the methodological choices adopted within the Delphi-CIMP.

6.3.1 Step 1: Defining Poverty and Selecting Components of the Delphi-CIMP

As discussed in Chapters 3 and 4, within this thesis, information on how poverty is defined, alongside the dimensions, indicators and weights relevant to proxy its experience, were collected through the Delphi. This sub-section aims to highlight the Delphi's key findings, which are relevant to the construction of the Delphi-CMIP.

Firstly, on how poverty should be defined, Delphi participants unanimously agreed that poverty is multidimensional. In particular, a significant number of participants expressed definitions, which were similar to within Sen's 'Capability Approach', i.e. when measuring poverty, income is not enough, rather focus should be given to the multiple capabilities one needs, to be free from poverty (Bogor City Delphi Participants, 2016; Natih, 2016a, 2016b). Following this, participants then selected dimensions, indicators and weights, which best represented these definitions. Table 6.1 provides a summary of these components.

Given this chapter's particular focus on the role of weights, Table 6.1 also highlights two challenges that need to be addressed before incorporating these results within the final composite measure, i.e. (1) the challenge of mitigating the existence of 'disagreement' with regard to the 'exact value' of the weights to utilise and (2) the challenge of the sum of indicator and sum of dimension weights chosen by Delphi participants, not equalling one.

To describe the first challenge, Table 6.1 includes the standard deviations and range of values for the chosen dimensions and indicator weights. As can be seen, these standard deviations and ranges were sizeable, thus illustrating the existence of a level of disagreement with regard to the exact values of these weights. Although, as discussed within Section 4.7 of Chapter 4, agreement on the importance of the 'lower' and 'upper' bound values for these weights listed within the last two columns of Table 6.1 was achieved, deriving complete orderings of households with regard to their total weighted deprivation levels, necessitates the determination of a *particular* set of weights, which ideally, should be able to take into account the diversity of views noted within Table 6.1.

Table 6. 1 Summary Statistics of Expert Weights (Bogor City, 2016)[(a)]

Dimension/Indicator	Median	Mean	MAD	SD	Min	Max
Dimension 1: EDUCATION	**0.25**	**0.28**	**0.05**	**0.19**	**0.10**	**0.60**
Children's Education	0.30	0.35	0.10	0.22	0.10	0.30
Literacy	0.25	0.27	0.05	0.20	0.20	0.60
Adult's Education	0.15	0.15	0.15	0.19	0.20	0.80
Ownership of Birth Certificate	0.10	0.12	0.10	0.13	0.10	0.40
Total Education indicator weights	0.80	0.89				
Dimension 2: EMPLOYMENT AND ASSETS	**0.20**	**0.18**	**0.05**	**0.13**	**0.10**	**0.50**
Employment	0.33	0.32	0.18	0.22	0.20	0.70
House and/or Land Ownership	0.15	0.15	0.15	0.16	0.10	0.50
Assets Ownership	0.15	0.14	0.15	0.16	0.10	0.70
Total Employment and Assets indicator weights	0.63	0.61				
Dimension 3: LIVING STANDARDS	**0.15**	**0.16**	**0.05**	**0.18**	**0.10**	**0.40**
Household Consumption	0.10	0.19	0.10	0.25	0.08	0.75
Government Help and Subsidies	0.09	0.11	0.09	0.15	0.05	0.55
Sanitation	0.05	0.06	0.05	0.07	0.10	0.20
Drinking Water	0.03	0.06	0.03	0.07	0.05	0.25
Cooking and Washing Water	0.03	0.07	0.03	0.09	0.05	0.40
Total Living Standards indicator weights	0.29	0.49				
Dimension 4: HEALTH AND SAFETY	**0.20**	**0.27**	**0.10**	**0.23**	**0.10**	**0.60**
Healthy Life	0.30	0.31	0.10	0.22	0.15	0.70
Safe Living Environment	0.30	0.28	0.10	0.21	0.10	0.60
Access to Immunizations	0.05	0.09	0.05	0.10	0.10	0.30
Total Health and Safety indicator weights	0.65	0.68				
Dimension 5: CHILD HEALTH AND CONTRACEPTION	**0.10**	**0.11**	**0.10**	**0.11**	**0.05**	**0.40**
Birth attendant	0.33	0.30	0.30	0.29	0.25	0.70
Contraception	0.20	0.16	0.20	0.17	0.20	0.60
Child mortality	0.10	0.15	0.10	0.16	0.10	0.50
Total Child Health and Contraception indicator weights	0.63	0.61				
Total dimension weight	0.90	1.00				

Note: [(a)] This table is a copy of Table 4.8 presented within Chapter 4 of this thesis. As described within Section 4.6 of Chapter 4, this table summarises results from the final Delphi round.

Source: Author's computations (Bogor City Delphi 2016)

To illustrate the second challenge, I briefly summarise the part of the Delphi questionnaire, which prompted the selection of weights. Within the weighting exercise, participants were requested to assign weights (from 0 to 1) to each dimension and indicator they had chosen.[1] They were asked to first assign weights for dimensions; this exercise was then followed by selecting weights for indicators within the dimensions they had previously selected. The sum of each participant's set of dimension weights had to total 1. Likewise, each participant's set of

[1] Weights were thus restricted to be non-negative.

indicator weights within dimensions, also had to total 1. As discussed within Section 4.6 of Chapter 4, particular importance was placed on the 'median' weights from the third Delphi round as indicative of the Delphi's final results. Referring to Table 6.1, it is however, shown that even though the sum of dimension and indicator weights selected by *each* participant had to equal 1, the sum of the median of all weights selected by *all* participants, did not equal 1.[1] This is due to the fact that each participant was given the freedom to not only choose different types of dimensions and indicators, but they were also free to decide the number of dimensions and indicators to include. Before delving into ways to mitigate the above challenges, a decision on how to transform the indicators listed within Table 6.1 into commensurate measurement units needs to be made.

6.3.2 Step 2: Transforming Delphi Indicators

The need to transform indicators is evident within the Delphi-CIMP. Looking back at Table 6.1, the Delphi's chosen indicators are measured by different scales, which renders them non-comparable before transformation. To briefly illustrate, raw data for the 'Children's education' indicator is measured by an ordinal scale variable, 'Ownership of house/land' is in the form of a nominal scale variable, whereas 'Household consumption' is an interval scale variable.

As described within sub-section 6.2.2, the selection of an appropriate transformation method, should take into account the objectives of the composite measure and the data properties of the indicators included. Firstly, on the purpose of the Delphi-CIMP, the main aim of the Delphi study was to address the arbitrariness in the way current measures of multidimensional poverty selected dimensions, indicators and weights included within them. The Delphi study enabled the eliciting of information with regard to these components directly from those involved within poverty alleviation.

The final goal of the exercise, is then to compare the effects of utilising such information with that used within the existing indices examined in Chapter 2, i.e. the Global and Adjusted MPIs.[2] Considering this, in order to focus on the effects of these choices, the transformation method utilised within the Delphi-CMIP follows that employed within the Global and Adjusted MPIs. Using a similar transformation method seeks to highlight the impact of differences in a measure's components, on how these measures identify the size and composition of the poor.

[1] As discussed within Section 4.6 of Chapter 4 the sum of the means and modes of these weights, particularly those of indicator weights, also did not equal one.
[2] This comparative exercise is presented within Chapter 7 of this thesis.

Secondly, as noted above, the data properties of the indicators included within the Delphi-CIMP, which are measured in different units, support the use of this transformation method. This in mind, I briefly summarise the said transformation method, below.

To identify deprived households, the transformation method utilised by the Global and Adjusted MPIs, follows the steps developed within Bourguignon and Chakravarty (2003) in which, a household's achievement within each indicator is transformed into simple "0" and "1" categories, through the use of indicator-specific cutoffs.[1,2,3] If a household's achievement in an indicator is equal to or above the cutoff, the household is assigned a score of "0", indicating non-deprivation. If its achievement is below the cutoff, the household is given a score of "1", indicating deprivation within the specific indicator. Based on this transformation method, each household will then have a set of 18 binary ("0" or "1") outcomes with regard to their deprivation status within each of the eighteen indicators listed in Table 6.1.

6.3.3 Step 4: Aggregation and Weighting within the Delphi-CIMP

6.3.3.1 Choice of Aggregation Function

As touched upon in sub-section 6.2.3 of this chapter, literature has mostly focused on the use of 'linear aggregation' within multidimensional poverty measurement. Indeed, it is also the approach taken by the Global and Adjusted MPIs (S. Alkire & Santos, 2010a; Martin Ravallion, 2011b). Following the logic behind the selection of the 'transformation' function outlined within the previous sub-section, adopting a similar aggregation method to that applied within these existing measures, seeks to enable focus to be placed on how the choice of measure components may impact results. Within this chapter, I thus, limit analysis to this most widely-applied 'linear aggregation' approach; a method which Martin Ravallion (2011b) dubs the 'deprivation aggregation' approach (p. 11).

The 'deprivation aggregation' approach measures deprivation in each indicator separately and then aggregates these indicator-specific deprivations into a composite index (Sabina Alkire &

[1] The specific indicator cutoffs utilised within the Delphi-CIMP are listed in Table 5.1 of Chapter 5.

[2] This transformation method was also utilised within Chapter 5 of this thesis.

[3] Despite its simplicity, this transformation function, however, does not take into account the severity of deprivation within indicators, i.e. how far below the indicator deprivation cutoff, a household's achievements are. Although beyond the scope of this chapter, the ability to take this into account is arguably an important aspect of measuring poverty, which allows one to take into account crucial issues, such as inequality (J. Foster, Greer, & Thorbecke, 1984b, 2010). I refer to S Alkire and James Foster (2011) for methods in which to incorporate the severity of poverty within measurement.

James Foster, 2011, p. 309; Martin Ravallion, 2011b, p. 9).[1] This method offers ease of interpretation, as 'aggregation' is simply conducted through calculating the weighted sum of a household's deprivation across the eighteen indicators included within the Delphi-CMIP. Martin Ravallion (2011b) and S Alkire and James Foster (2011) however, note that within this approach, weights are assumed to *all* possess non-negative values and the sum of these weights is restricted to be equal to one.[2] This assumption bases the need to address the 'second challenge' of incorporating Delphi opinions within the Delphi-CIMP, i.e. the challenge of arriving at a set of non-negative weights of which the sum totals one. The next sub-section describes how this chapter proposes to address this challenge, alongside the first challenge, i.e. the challenge of taking into account the differences in opinion found with regard to the exact value of these weights.

6.3.3.2 Incorporating Expert Weights in the Delphi-CIMP

This section proposes the use of the 'benefit-of-the-doubt' approach to weighting, in order to address the above-mentioned challenges. Key aspects of this approach were described within sub-section 6.2.3 of this chapter. The remainder of this sub-section aims to add to this discussion and delve deeper into the application of this approach.

As was noted within sub-section 6.2.3, the application of the 'benefit-of-the-doubt' approach involves applying a Data Envelopment Analysis (DEA) linear programming model. This involves three main steps, i.e. (1) choosing whether to apply an index maximising or an index minimising DEA, (2) determining a 'best-practice frontier' and (3) calculating 'optimal weights' given steps (1) and (2). Figure 6.2 summarises these steps.

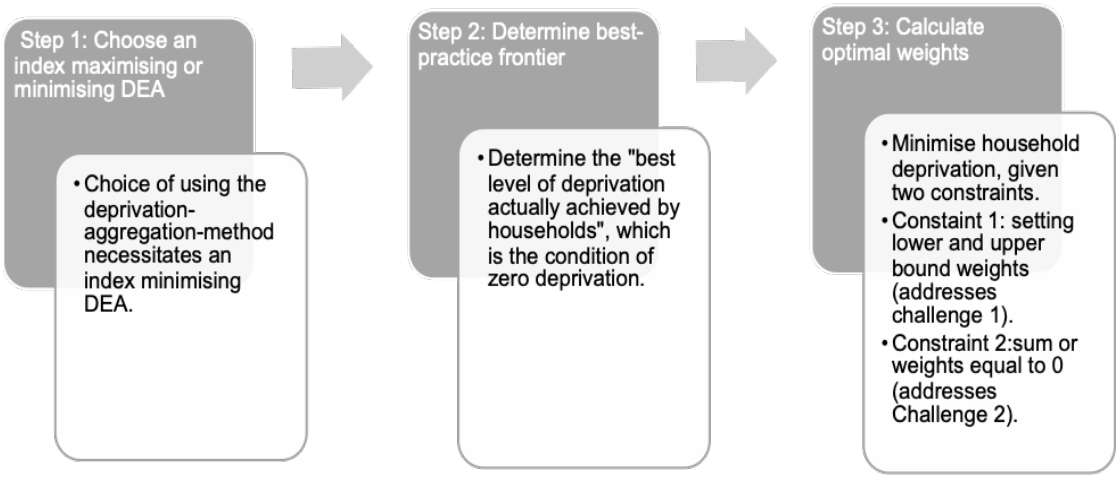
Within the first step, Mahlberg and Obersteiner (2001) note that there exists either index maximising or index minimising DEA models; where an index maximising DEA is applied if the scores of all indicators are preferred to be as high as possible (e.g., GDP, literacy rate, etc.) and

[1] Formal treatment of this approach can be found within Tsui (2002), Bourguignon and Chakravarty (2003), Duclos et al. (2006) and S Alkire and James Foster (2011).

[2] In addition to this, within a linear aggregation function, the elasticity of substitution between each indicator is assumed to be 'infinite', thus, indicators are assumed to be 'perfect substitutes' (Decancq & Lugo, 2008, p. 4). Linear aggregation also assumes a fixed rate at which transformed achievements can be exchanged, which is a 'constant' for all possible attributes (Ibid). Bourguignon and Chakravarty (2003) note that these may not be very satisfactory assumptions, as indicators may not be perfect substitutes in real life (p. 40-41). Despite these assumptions, as noted above, the conscious decision of applying this commonly used aggregation function, was taken. S. Anand and Sen (1997) and Chakravarty and Majumder (2005), however, offer a discussion on the possible implications of considering varying degrees of elasticity of substitution between these multiple indicators.

an index minimising approach is applied if scores are preferred to be as low as possible (e.g. deprivation rates) (p. 6).[1] Within this chapter, given the decision to apply the 'deprivation aggregation' approach, the index minimising DEA is applied. The aim of the minimisation exercise within the DEA is thus to *minimise* a household's level of deprivation within the indicators included in the Delphi-CIMP.

Figure 6. 1 Steps to Determine Optimal Expert Weights within the DEA's Minimisation Model



Source: Author's summary

For the second step, district government development goals were followed in order to determine the 'best-practice frontier', to which the 'performance' of households with regard to their indicator deprivation rates, may be compared. This in mind, the target of achieving 'zero deprivation' within all of the eighteen indicators, was utilised as a benchmark. This follows the press release delivered by the City government following the Delphi exercise, which expressed the 'district government's commitment to ensuring that all basic needs within the indicators chosen through the Delphi, are met' (Bogor City Government, 2016).[2]

[1] Coelli, Rao, O'Donnell, and Battese (2005) provide further discussion on the properties of these models.

[2] In addition to this, the use of 'zero deprivation in all indicators' as the best-practice frontier, somewhat follows theoretical arguments commonly utilised to underpin the selection of frontiers within DEAs. Following Farrell (1957) to solve the optimisation problem in DEAs, 'it is far better to compare performances with the best actually achieved'. Given the distribution of household deprivations in Bogor

In the third step, dimension and indicator weights within the DEA model, are calculated in order to *minimise* the distance between a household's deprivation level and the best-practice frontier, i.e. the condition of 'zero deprivations in all indicators'. It is within this step, that opinions on weights derived from the Delphi exercise, play an integral part. This step also offers solutions to the two challenges faced when incorporating these views, noted at the start of this sub-section.

To address the first challenge 'non-negativity and non-zero constraints' are set, before solving the above minimisation problem. Referring back to Table 6.1, the minimum and maximum weight values stated by participants are used to define this first constraint. By fulfilling this constraint, the solution of the minimisation problem, i.e. the *exact value* of dimension and indicator weights should thus be within the range chosen by Delphi participants. This constraint allows us to take into account the plurality of judgment evident with regard to these weights. In addition to this, applying this constraint mitigates the possibility of 'extreme scenarios', i.e. the minimisation problem assigning a weight of "1" to indicators where households tend to perform best and assigning a weight of 'zero' to indicators where performance is worst (L. Cherchye et al., 2008, p. 241).[1,2]

To tackle the second challenge, a second constraint is imposed. This constraint stipulates that the sum total of dimension weights should equal one; likewise, the sum total of indicator weights within dimensions, should also equal one. This constraint, often dubbed within DEA literature as the 'bounding constraint', fulfils the assumptions behind the use of the deprivation-aggregation approach as expressed within sub-section 6.3.3.1.

The solution of the minimisation problem, whilst taking into account the above two constraints, resulted in *re-weighted* versions of the weights chosen by Delphi participants. I dub these new weights, 'optimal expert' weights.[3] In order to arrive at these 'optimal expert' weights, the above

City, which will be explored further within Section 6.4, this 'best level achieved' was indeed found to be the condition of 'zero deprivations'.

[1] Within DEA literature the possibility of extreme scenarios is known as the 'boundary problem' (Mahlberg & Obersteiner, 2001).

[2] As further reference, Despotis (2005), Thanassoulis (2001), Cooper, Seiford, Thanassoulis, and Zanakis (2004) and Portela, Thanassoulis, and Simpson (2004), among others, offer additional guidance on the issue of 'setting boundaries' within DEAs.

[3] As the Delphi included thirty-six participants, for each dimension and indicator, this minimisation exercise generated thirty-six 'new' weights; I call these weights 'DEA generated weights'. These 'DEA generated weights' were utilised to re-weight the answers given by Delphi participants. The final 'optimal expert' weight for each dimension and indicator was determined by the 'sum' of these re-weighted weights. To note, the 'DEA generated weights' were relatively similar for each participant. The opinions of all participants were thus, considered as relatively equal in value within the calculation of the 'optimal

minimisation problem was solved using Python 3.7's 'gradient descent' minimisation package and the PyCharm CE text editor.[1] Before discussing the results of this exercise, the next sub-section describes the robustness checks conducted within this chapter.

6.3.4 Step 4: Auxiliary Robustness Checks

Uncertainty analysis within this chapter primarily focuses on the effects of two different weighting schemes, i.e. 'equal' vs. 'optimal expert' on how the final Delphi-CIMP measure identifies the different sizes or compositions of the poor, alongside whether the weighting schemes lead to differences in how households are ranked according to their deprivation levels. In addition to this, the possible effects of the use of different *poverty cutoffs*, i.e. the number of deprivations a household needs to be deprived in, in order for that household to be considered as poor, is also examined (S Alkire & James Foster, 2011). This is of particular importance as, following the analysis presented within Section 4.3 of Chapter 4, Delphi participants were unwilling to set such a cutoff.[2]

6.4 Results

6.4.1 Research Question 1: What are the Optimal Expert Weights?

This sub-section describes the outcomes of the linear minimisation problem outlined within sub-section 6.3.3 and aims to answer the first research question within this chapter; i.e. what are the values of the 'optimal expert' weights for each dimension and indicator included in the Delphi-CIMP. Table 6.2 presents these results.

---

expert' weights, which resulted from this exercise. The mathematical formulation for this minimisation problem is presented within the Appendix to this chapter.

[1] The 'gradient descent' method was chosen as it is the most parsimonious and frequently used algorithm utilised within optimisation studies; it's use is, however, particularly common within 'deep learning models' in machine learning (Boyd & Vandenberghe, 2004; Ketkar, 2017, p. 105; Raschka, 2015). To solve minimisation problems, this method first generates an initial random value for each unknown parameter. The maximum number of iterations are then set, alongside a measure stipulating the 'tolerance' of the model. This 'tolerance' level determines the maximum distance between results from the last iteration and outcomes from the previous iteration; in other words, it is used to specify a 'rule' by which to prompt the iterations to 'stop' when a specified 'optimal solution' has been achieved (Ibid, p. 110). Technically, an infinitely small tolerance level could be set. This would require a significantly larger amount of iterations; thus, a longer time period would be needed in order to solve the minimisation problem. This in mind, in order to balance precision and parsimony, for the minimisation problem within this chapter, the 'tolerance' level was set at the level when the difference between the 'optimal expert' weight resulting from the final iteration, with the weight resulting from the iteration preceding this final result, equalled 1E10-10.

[2] This chapter leaves the analysis of possible other sources of uncertainty, which may result from the choice of other parameters, such as the type of transformation method or aggregation function, as topics for further research

Table 6. 2 Optimal Expert Weight vs Equal Nested Weights[(a)]

Dimension/Indicator	Delphi Lower Bound Weight	Optimal Expert Weight	Delphi Upper Bound Weight	Equal Nested Weights
Dimension 1: EDUCATION	**0.1**	**0.29**	**0.6**	**0.2**
Children's Education	0.1	0.1	0.3	0.05
Adult's Education	0.2	0.03	0.8	0.05
Literacy	0.2	0.15	0.6	0.05
Ownership of Birth Certificate	0.1	0.02	0.4	0.05
Dimension 2: EMPLOYMENT AND ASSETS	**0.1**	**0.18**	**0.5**	**0.2**
Employment	0.2	0.08	0.7	0.07
Ownership of Assets	0.1	0.06	0.7	0.07
Ownership of House and/or Land	0.1	0.04	0.5	0.07
Dimension 3: LIVING STANDARDS	**0.1**	**0.14**	**0.4**	**0.2**
Household Consumption Level	0.08	0.05	0.75	0.04
Sanitation	0.1	0.02	0.2	0.04
Drinking Water	0.05	0.02	0.25	0.04
Cooking and Washing Water	0.05	0.02	0.4	0.04
Government Help	0.05	0.03	0.55	0.04
Dimension 4: HEALTH AND SAFETY	**0.1**	**0.26**	**0.6**	**0.2**
Health	0.15	0.09	0.7	0.07
Immunisation	0.1	0.03	0.3	0.07
Safety	0.1	0.14	0.6	0.07
Dimension 5: CHILD HEALTH AND CONTRACEPTION	**0.05**	**0.12**	**0.4**	**0.2**
Birth attendant	0.25	0.07	0.7	0.07
Child mortality.	0.1	0.03	0.5	0.07
Contraception	0.2	0.03	0.6	0.07
Sum of dimension weight		1		1
Sum of indicator weight		1		1

Note: [(a)] The minimisation exercise proposed within sub-section 6.3.3 generated 'optimal expert weights' for each dimension and indicator within dimensions. Following that conducted in the Global and Adjusted MPIs, the *final* 'optimal expert weight' for an indicator was found through multiplying the 'optimal expert weight' found for the dimension, which encompasses said indicator, with the 'optimal weight' found for the indicator itself. In this way, as shown within this table, the sum of all eighteen 'optimal expert' indicator weights equalled one

Source: Author's computations (Bogor City, 2016)

Within Table 6.2, the minimum (lower bound) and maximum (upper bound) values for these weights, as was noted before within Table 6.1, are also listed. In addition, the values of weights if the 'equal nested' scheme were utilised are also listed. The exercise of comparing 'optimal expert' vs. 'equal nested' weights will be the subject of sub-section 6.4.2; for now, focus is given to the sole examination of the 'optimal expert' weights and what their values may entail. Considering this, from Table 6.2, four main findings may be deduced.

Firstly, Table 6.2 shows that 'optimal expert' weights for each dimension and indicator, fulfil the first constraint set within the minimisation problem, i.e. that their values should lie within the lower and upper bound weights chosen by Delphi participants. Referring back to sub-section 6.3.3.2, this constraint was set to address the challenge of the existence of disagreements with regard to the exact values for these weights. The 'optimal expert' weights thus, aimed to act as an objective solution to this diversity of views.

Secondly, 'optimal expert' weights are shown to vary across each dimension and indicator. Referring back to the discussion within sub-section 6.3.3, this variation aims to reflect preferences (particularly policy priorities) and possible trade-off levels between the dimensions and indicators included. To illustrate, looking first at dimensional weights, a hierarchy of importance of dimensions may thus be observed, with deprivation within 'Education' given the highest concern, followed by that within 'Health and Safety', 'Employment and Assets', 'Living Standards' and 'Child Health and Contraception'.[1] As discussed within Chapter 4, results of the post-Delphi semi-structured interviews confirmed the importance of this hierarchy. It is thus encouraging that the 'optimal weights' presented within Table 6.2 preserved this ordering of dimensions.

The third finding relates to whether the 'optimal expert' weights within Table 6.2 may be considered as 'most favourable' weights. To demonstrate this point, 'optimal expert' weights for indicators are examined. Given the importance of the 'Education' dimension as reflected by its relatively high optimal expert weight, I look to indicators within this dimension as examples. Within this dimension, 'Literacy' is assigned the highest weight. Referring back to the indicator deprivation rates presented within Figure 5.1 of Chapter 5, deprivation rates within 'Literacy' were found to be relatively low (at 0.4 percent deprived households), compared to the other indicators. This shows that 'optimal expert' weights are indeed 'most favourable' weights, as

[1] As noted within Section 4.6 of Chapter 4, a similar ordering of dimensions was observed when the 'median' weights of the third Delphi round were examined.

higher weights are assigned to aspects that households perform relatively well in. It is important to note, however, that this 'higher' weight is still within the bounds set by Delphi participants. These 'bounds' thus act to ensure that the values of 'optimal expert' weights are also dependent on empirical data reflecting expert opinions on their importance (Laurens Cherchye & Kuosmanen, 2004; L. Cherchye et al., 2008). Referring to whether these weights reflect existing policy, ensuring that every citizen fulfils the right to be able to read and write has always featured as a major priority within development plans set by the national and district governments in Indonesia (Republic of Indonesia Ministry of National Development Planning, 2019). The importance of this indicator is also reflected by the stringency of the cutoff placed to define deprivation within it; as summarised in Table 5.1 of Chapter 5, following guidelines set by Indonesia's Central Bureau of Statistics (CBS), the *whole* household is considered as deprived if at least one individual within it cannot read or write.

The final finding within Table 6.2, illustrates how the minimisation exercise within the benefit-of-the-doubt approach mitigates the second challenge of incorporating Delphi opinions, i.e. that the sum total of dimension and indicator weights has to equal one. As shown in the last two rows of Table 6.2., utilising 'optimal expert' weights address this challenge.

In sum, the findings above demonstrate the relevance of the use of the 'optimal expert' weights in tackling the two challenges faced when incorporating Delphi opinions. In addition to this, the above analysis acts to confirm the plausibility of these weights as those aiming to reflect the preferences, possible policy priorities and degrees of trade-offs between dimensions and indicators, as chosen by those involved in the Delphi study. Considering this, the next section presents results of the main purpose of this chapter, i.e. comparing the effects of using these 'optimal expert' weights vs. the commonly applied 'equal nested' weights.

6.4.2 Research Question 2: Optimal Expert vs. Equal Nested Weights

This section aims to answer Research Question (2), i.e. whether the Delphi-CIMP will identify different sizes and compositions of the poor if two different weighting schemes, 'optimal expert' vs. 'equal nested' are used.

6.4.2.1 Values of the Weights

Before examining possible differences in the values of the weights within the two weighting schemes, a brief recap of how 'equal nested' weights are determined is presented below.

As described within Chapter 2, within the 'equal nested' weight approach, equal weights are first assigned to each of the dimensions (S. Alkire & Santos, 2010a; Cohen & Saisana, 2014). All of the five dimensions within Table 6.2 are thus, given a weight of 1/5 = 0.2. Equal weights are then assigned to indicators *within* dimensions. As an example, within the 'Education' dimension, each of its four indicators are given an equal weight of 1/5x1/4=1/20=0.05. Indicator weights thus may vary, as dimensions may include a different number of indicators.

Looking closely at the values of the weights listed within Table 6.2, the use of 'equal nested' weights therefore implies 'equal importance' of each dimension, however, different degrees of importance are assigned to each indicator. Compared to the 'optimal expert' weighting scheme, however, it is less clear whether these differences in *indicator* weights within the 'equal nested' scheme reflect priorities and preferences, or if they are merely a product of how indicators are clustered into different dimensions. Furthermore, for dimensional weights, the hierarchy of importance evident in the 'optimal expert' weights, is not present when 'equal nested' weights are used. As noted within the previous sub-section, preserving this hierarchy within dimensions, was considered as essential by participants within the Delphi (Bogor City Delphi Non-Participants, 2016; Bogor City Delphi Participants, 2016). The lower 'equal nested' weights assigned to the 'Education' and 'Health and Safety' dimensions, may thus, underestimate the importance of both these dimensions, compared to if 'optimal weights' are assigned. On the other hand, the higher 'equal nested' weights given to the 'Living Standards', 'Employment and Assets' and 'Child Health and Contraception' dimensions, compared to that if 'optimal expert' weights were used, could overvalue the importance of these dimensions.

Before ending this sub-section, I note that although the above differences exist, the values for the 'equal nested' weights assigned to each dimension and indicator lie within the lower and upper bound weights suggested by Delphi participants. Even though this is the case, referring back to the Delphi results described within Chapter 4, only four out of the thirty six participants (eleven percent of all participants), proposed assigning equal weights to each dimension.[1] The use of 'equal nested' weights within the Delphi-CIMP, thus, remains questionable. This issue is of importance, as differences in weights may affect how poor households are identified and thus, how policy is determined. The sub-sections below explore this further.

[1] Referring to the Delphi results in Chapter 4, except for weights for 'Drinking Water', 'Sanitation' and 'Cooking/Washing Water', which as suggested by the 'optimal expert' weights listed within Table 6.2, should be assigned equal importance, almost all participants chose different weights for each indicator.

6.4.2.2 Household Ranks

To determine household rank, a household's total weighted deprivation when the two weighting schemes are utilised is first calculated. We recall that through the transformation method suggested within sub-section 6.3.2, households are assigned "0" or "1" scores, indicating 'non-deprivation' or 'deprivation', within each of the eighteen indicators. Referring back to sub-section 6.3.3.1 a household's total weighted deprivation is calculated by determining the weighted average of the household's deprivation score within the eighteen indicators in Table 6.2. Households are then assigned ranks according to these weighted sums.[1] Finally, the spearman rank correlation coefficient is computed as a way of assessing whether a household's rank changes when the different weighting schemes are applied. Table 6.3 presents the results of this exercise.

High correlation is shown to be evident between household rank when the two weighting schemes are applied.[2,3] If the main government policy was to target the poorest households according to their total weighted deprivation, the use of both weighting schemes may thus, lead to similar conclusions, i.e. it is likely that the government would tend to deliver social assistance to the same households with the highest levels of total weighted deprivation, as they would be ranked relatively similarly, regardless of the weighting scheme applied.

Table 6. 3 Rank Correlation coefficients between household rank using optimal expert vs. equal nested weights

	Optimal Expert	Equal Nested
Optimal Expert	1	
Equal Nested	0.92***	1

Note: *** Denote significance at 0.1% level
Source: Author's computations (Bogor City, 2013)

On this issue, previous research, mostly aimed at evaluating the weights employed within the Human Development Index (HDI), arrived at similar conclusions, i.e. high levels of correlation

[1] This is similar to how household ranks for the Global and Adjusted MPIs were determined in Chapter 2.
[2] Similar to the above results examining household rank, relatively high correlation (0.93 in Bogor City) was also found between households' total weighted deprivation scores, when the use of the 'optimal expert' and 'equal nested' weighting schemes, were compared.
[3] Significant correlation between household ranks was also evident if the analysis was expanded to include West Java (0.93) and Indonesia (0.93).

were found between country ranks when the use of DEA 'benefit-of-the-doubt' generated weights, was compared with the HDI's equal weights (Laurens Cherchye, 2001; Chowdhury & Squire, 2006; Despotis, 2005; Mahlberg & Obersteiner, 2001). Furthermore, of particular relevance to this chapter, Chowdhury and Squire (2006), proposed the inclusion of opinions of researchers on weights, to re-weight the HDI. These authors found that despite widespread criticism on the use of equal weights, their resulting 'opinion based' weights did not lead to substantial variation in results with regard to how countries are ranked. They concluded that the use of equal weights was not only 'simple' and 'convenient', but importantly also reflected well the opinions of the experts included within their study. Nevertheless, these authors continue to caution against the use of equal weights, as their findings were only based on one sample of experts.

To summarise, based on the results above, if importance is only placed on how households are ranked, the use of the relatively less resource and time consuming 'equal nested' weights, *may* suffice. This is an important finding, as more often than not, policy makers are faced with limited funds and resources. Following Chowdhury and Squire (2006) cautionary note, however, in addition to the possible differing conclusions, which could be found if the 'pool of experts' is widened, 'ranks' comprise only one side of the story. Additional analysis needs to be conducted to test whether the use of the two weighting schemes leads to disparities with regard to the size and composition of the poor. The sub-sections below explore this further.

6.4.2.3 Size of the Poor

Before analysing possible differences in the size of the poor, a 'poverty cutoff', i.e. the amount of deprivation a household needs to experience to be considered poor, first needs to be determined. Sub-section 5.4.2 of Chapter 5 introduced the idea of applying the 'union', 'intersection' or 'intermediate' approaches to determine deprivation cutoffs. It was however found, that the 'union' and 'intersection' approaches tended to identify extreme values with regard to the number of poor.[1] Within this current sub-section, the 'intermediate' approach is thus applied. [2] As described within Chapter 4, Delphi participants were unwilling to define a rigid poverty cutoff, thus, a range of these 'intermediate' cutoffs, is considered.

[1] Furthermore, S. Alkire (2011) states that 'if either the union or the intersection identification approaches are used, no particular weights on indicators are required in order to identify who is poor' (p. 15).

[2] Since within this chapter, *weighted* indicator deprivations are considered, applying the 'union' approach, i.e. a household is considered as poor if it is deprived in at least one of the eighteen indicators, becomes challenging. Employing the 'intersection' approach proves to be less demanding, as determining the

Before setting these cutoffs, to maintain comparability between the distributions of weighted deprivations when the two weighting schemes are applied, households' total weighted deprivation scores, were first standardised to have means of zero and standard deviations of one.[1] A household was then categorised as poor if *its standardised total weighted deprivation is equal to or larger than a certain number of standard deviations away from the standardised mean*.[2]

Considering the above, Figure 6.3 examines the effects of the use of the two weighting schemes across a number of poverty cutoffs ranging from 0.1 to 1.5 standard deviations away from the standardised mean of households' total weighted deprivations.[3] From this figure, two main findings may be identified.

Firstly, the use of the different weighting schemes and different poverty cutoffs lead to significantly different levels of poverty. The poverty headcount is thus shown to be highly sensitive to the weighting scheme and cutoff applied. Secondly, examining the two weighting schemes more closely, even though within sub-section 6.4.2.2 high correlation was found between household ranks when the two schemes are employed, Figure 6.3 shows that these schemes identify significantly different sizes of the poor, regardless of the poverty cutoff used.

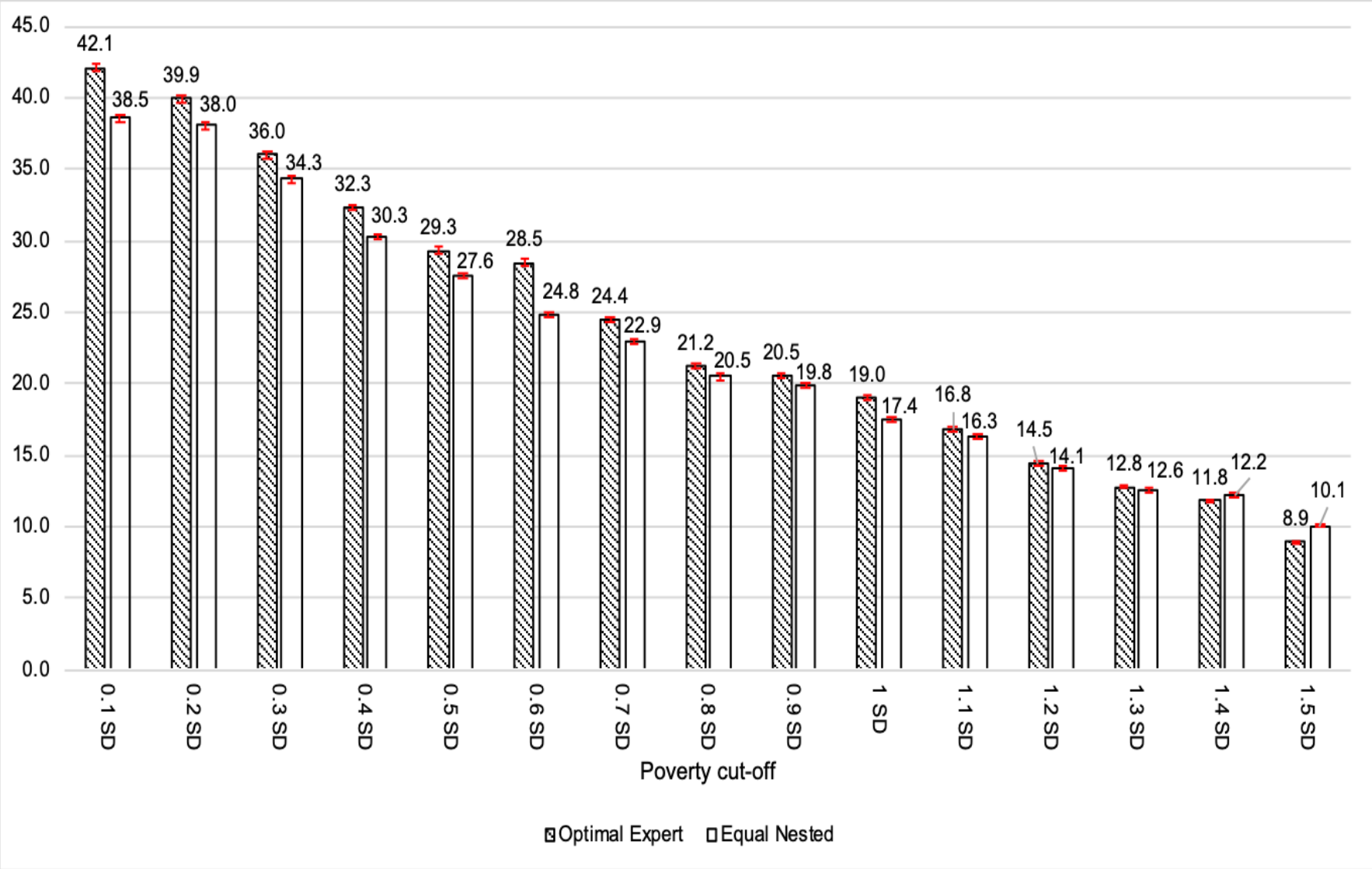
number of households which are deprived in all weighted indicators, is congruent to calculating those with a total weighted deprivation score of one. If this 'intersection' approach was applied, however, no households (zero percent) were counted as poor in Bogor City. This strengthens the need to utilise an intermediate approach.

[1] Graphs for these distributions are available within the Appendix to this chapter.

[2] This approach slightly differs from the approach applied within the Global MPI, which uses a proportion of weighted indicator deprivations as its poverty cutoff (S. Alkire & Santos, 2010a).

[3] This particular range includes the headcount poverty percentages if the three existing measures, examined in Chapter 2, were employed. Referring back to sub-section 2.3.1 of Chapter 2, the poverty headcount in Bogor City found when the Consumption measure is employed (14 percent), is relatively close to the percent poor if the cutoff for the Delphi-CIMP is set at 'a household is poor if it is suffering a total weighted deprivation that is equal to or more than 1.2 standard deviations away from the standardised mean'. The poverty headcount when the Global MPI is employed (16 percent), is relatively close to the percent poor if the cutoff for the Delphi-CIMP is set at 'equal to or more than 1.1 standard deviations away from the mean'. Finally, the poverty headcount when the Adjusted MPI is employed (19 percent), is relatively close to the percent poor if the cutoff for the Delphi-CIMP is set at 'equal to or more than 0.9 standard deviations away from the mean'.

Figure 6. 2 Percent of poor households using Optimal Expert vs. Equal Nested weights and a range of poverty cutoffs[(a)(b)]



Note: [(a)] Poverty cutoff is defined as: a household is poor if its total weighted deprivation is equal to or above n standard deviations (SDs) away from the standardised mean; [(b)]95 % confidence intervals are shown in red.
Source: Author's computations (Bogor City, 2013)

To illustrate, when using the 'equal to or more than 0.1 to 1.3 standard deviations away from the mean' range of cutoffs, the 'optimal expert' weighting scheme is shown to identify significantly more poverty, compared to the 'equal nested' weights.[1] For this range of poverty cutoffs, employing 'equal nested' weights may thus, *underestimate* the degree of poverty, when compared to if 'optimal expert' weights are used.[2] When the more stringent cutoff of 'equal to or above 1.4 SD above the mean', is used, the 'equal nested' scheme is shown to *overestimate* the number of the poor, when compared to the 'optimal expert' weights.[3]

As noted within previous chapters, Indonesia places high importance on targeting policies as a way to address poverty. For these policies, a significant constraint to 'better' targeting is accountable information on household poverty (Lipton & Ravallion, 1995; Martin Ravallion, 1993). Furthermore, 'inadequate information' on poverty may lead to 'type I errors', i.e. 'incorrectly classifying a household as poor' and/or 'type II errors', i.e. 'incorrectly identifying a household as non-poor' (Cornia & Stewart, 1993, pp. 460-462).[4] Looking at the results within Figure 6.3, for the range of cutoffs 'equal to or more than 0.1 to 1.3 standard deviations from the mean', employing the 'equal nested' scheme may thus lead to significant type II errors, when compared to if 'optimal expert' weights are utilised. For this range of poverty cutoffs, aggregating poverty using 'equal nested' weights, may thus, increase the possibility of failing to deliver social assistance to those who deserve government support.

In practice, however, Martin Ravallion and Datt (1995), point toward the possibility that the most 'cost effective' policy may not have the 'lowest errors of targeting'. This is the case as given the high costs of targeting policies, the 'deliberate introduction' of imperfect coverage, could allow a reduction in these costs, thus resulting in a greater impact on poverty for a given budget (Ibid). Particularly considering the high correlation found within sub-section 6.4.2.2, even though as shown within this current section, employing 'equal nested' weights may lead to significant errors in targeting, the parsimony it allows, could still support its relevance. Despite this, Alatas

[1] This finding still holds if the range of the poverty cutoffs is extended to include cutoffs below the standardised mean.

[2] Within the '0.1 to 1.3 SD poverty cutoff' range, the disparities between the two weighting schemes range from a difference of 0.2 percent when the '1.3 SD cutoff' is used, to close to 4 percent when the '0.1 SD cutoff' is applied. These percentages correspond to the 'optimal expert' scheme identifying 430 to 7,000 more households in Bogor City as poor, when compared to when 'equal nested' weights are used.

[3] This conclusion also holds for all cutoffs above the '1.5 SD' cutoff.

[4] For a summary of targeting policies in Indonesia, please refer to Daly and Fane (2002), among others. For detailed discussion on type 1 and type 2 errors of existing cash transfers in practice in Indonesia, please refer to Alatas et al. (2012).

et al. (2013) and Alatas et al. (2012) note the importance of 'acceptability' and 'community satisfaction', as additional desirable outcomes when it comes to assessing the effectivity of targeting policies. Employing a more accountable and transparent poverty measure, which bases not only its choice of indicators, but also its weights, on local views, may thus be crucial, not only to avoid targeting errors, but also to ensure the fairness and acceptability of the policy.

To summarise, both the weighting scheme and poverty cutoff employed, significantly influences the size of the poor identified by the Delphi-CIMP. For a substantial range of poverty cutoffs, however, the 'optimal expert' scheme is shown to consistently identify more households as poor, when compared to the 'equal nested' scheme.

6.4.2.4 Composition of the Poor

As with the section above, in addition to analysing the effect of using different weighting schemes, within this section, the effect of using a range of poverty cutoffs on the composition of the poor is also examined. Considering this, to analyse the composition of the poor, multivariate logistic regressions are conducted. These regressions aim to assess the effects of different household characteristics on the probability of experiencing poverty, when the two different weighting schemes are applied.[1] Conscious of space, results for the 'equal to or above 0.9 – 1.2 standard deviations away from the mean' range of poverty cutoffs, is displayed.[2] Table 6.4 reports the log odds of these regressions. Two main findings are suggested:

Firstly, for both weighting schemes, although the magnitude of the coefficients slightly differs, the direction of the relationship between the different statistically significant household characteristics, with a household's probability of being poor, are relatively similar, regardless of the poverty cutoff utilised. Looking more closely at these household characteristics, the 'age' and 'educational level' of the household head, alongside the 'composition' of the household are shown to significantly influence the household's probability of being poor in similar ways, for both the 'optimal expert' and 'equal nested' weighting schemes. Delving deeper into the effects of each characteristic, with regard to the 'age' of the household head, an 'inverted U relationship' is observed for both weighting schemes, i.e. poverty decreases as a household head reaches middle age and increases as he or she approached old age.

[1] These household characteristics are also utilised to describe the composition of the poor for the existing measures examined in Chapter 2.

[2] As noted within sub-section 6.4.2.3, this range of cutoffs, includes the percentages of poor households when the Consumption, Global and Adjusted MPIs are employed. Similar results to that found within this sub-section, however, hold if a larger range of poverty cutoffs is used.

Table 6. 4 Multivariate Logistic Regression Results Analysing the Composition of the Poor (Log odds)[(a)(b)]

	Optimal Expert				Equal Nested			
	≧0.9 SD	≧1 SD	≧1.1 SD	≧1.2 SD	≧0.9 SD	≧1 SD	≧1.1 SD	≧1.2 SD
Age	-0.27***	-0.25***	-0.25***	-0.21***	-0.32***	-0.31***	-0.29***	-0.23***
Age Squared	0.002***	0.002***	0.002***	0.002***	0.003***	0.003***	0.002***	0.002**
Gender (ref: Male)								
Female	0.97	0.82	0.43	0.42	1.08	1.06	1.09	0.89
Marital Status (ref: Single, Divorced or Widowed)								
Married	1.08	0.94	0.65	0.51	1.01	1.15	0.91	0.57
Household Head Education Level (ref: No qualifications or never been to school)								
Primary	0.21	0.17	0.22	0.44	-0.52	-0.27	-0.06	0.34
Secondary	-0.81	-0.81	-0.90*	-0.44	-1.90***	-1.68***	-1.64***	-1.70**
Higher Education	-1.47***	-1.61***	-1.60***	-1.54**	-2.49**	-2.62***	-2.85***	-2.61***
Household Head Work Sector (ref: Extractive Industries)								
Economically Inactive	0.15	-0.01	0.42	0.10	-0.14	-0.46	-0.76	-0.94
Manufacturing and Construction	-0.75	-0.83	-0.34	-0.66	-0.57	-0.66	-1.06	-1.17
Trade and Services	-0.63	-0.85	-0.39	-0.88	-0.68	-0.95	-1.26	-1.66*
Community Service and Public Administration	-0.33	-0.64	-0.34	-0.70	-0.71	-0.83	-1.11	-1.52
Other	0.28	0.28	0.88	0.83	-0.65	-0.68	-0.98	-1.21

Table 6.4 (continued) [(a)(b)]

	Optimal Expert				Equal Nested			
	≧0.9 SD	≧1 SD	≧1.1 SD	≧1.2 SD	≧0.9 SD	≧1 SD	≧1.1 SD	≧1.2 SD
Household Composition								
Number of adults in household (ref: 2 Adults)								
1 adult	0.73	0.66	0.32	0.76	0.73	0.95	0.97	0.85
More than 2 adults	-1.04***	-1.08***	-0.95**	-0.69*	-0.59*	-0.61*	-0.56	-0.74*
Number of children in household (ref: No Children)								
1 child	1.47***	1.50***	1.58***	1.45**	1.34***	0.99*	1.12*	1.10*
2 children	1.45**	1.52**	1.50**	1.50**	1.17*	0.77	0.90	0.78
More than 2 children	2.54***	2.39***	2.38***	2.29***	1.99***	1.77***	1.95***	1.88***
Constant	4.32*	4.25*	3.94*	3.06	6.73***	6.75***	6.84**	5.64**
Log likelihood	-222.88	-213.01	-202.17	-182.09	-216.44	-201.81	-185.01	-167.08
Number of obs	573	573	573	573	573	573	573	573
LR chi2(21)	117.58	110.72	100.59	84.32	130.46	126.96	131.58	118.08
Prob > chi2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Pseudo R2	0.21	0.21	0.20	0.19	0.23	0.24	0.26	0.26

Notes: [(a)] The second row shows the poverty cutoffs, i.e. a household is poor if its standardised total weighted deprivation is 'x' standard deviations (SDs) away from the mean. 'x' in this case ranged from 0.9 – 1.2 SDs away from the mean.
[(b)] *, ** and *** denote significance at the 5%, 1% and 0.1% level

Source: Author's computations (Bogor City, 2013)

As for the household head's 'level of education', higher levels of education are associated with a decrease in the log-odds of being poor for both schemes. The 'number of adults' within a household is shown to have a significantly negative effect on the probability of being poor. As argued for within Chapter 2, the 'number of adults' could indicate the number of 'working individuals' within the household, which may point toward the household's ability to increase its earnings and thus decrease the household's probability of being poor.[1] Finally, the 'number of children' within the household increases the probability of being poor, as determined by both weighting schemes. This may point toward additional child costs, particularly toward education, households may face (Bessell, 2009, p. 538). [2]

Secondly, on further examination, however, despite the above similarities, household composition, particularly the 'number of children' within the household, seemed to matter more when the 'optimal expert' scheme was utilised. The 'education level' of the household head then played a more significant role when the 'equal nested' weights were applied. Referring back to Table 6.2,, the 'optimal expert' weight for 'Children's Education' (at 0.1) is shown to be higher than its 'equal nested' weight (at 0.05). This higher weight may have led to households with 'more children' having a higher likeliness of being poor when 'optimal expert' weights are used. On the other hand, looking at the weight for 'Adult Education', the 'equal nested' scheme was shown to assign a higher weight (at 0.05) compared to that given by the 'optimal expert' scheme (at 0.03). This then, could have contributed to the 'education level of household head' variable playing a larger role in explaining household poverty if 'equal nested' weights are utilised. These differences, although subtle, highlight the possible impacts of applying the different weighting schemes toward policy. To illustrate, employing 'optimal expert' as opposed to 'equal nested' weights, point toward the need for additional social assistance to be delivered to households with children, in particular, to ensure that these children attend and have access to adequate schooling.

In sum, regardless of the poverty cutoff chosen, the two weighting schemes seemed to identify households with relatively similar characteristics, as poor. These households are generally headed by older individuals, who are less educated and live within households with additional

[1] 'Work' in this case, may refer to 'formal' or 'informal' employment. This distinction between 'formal' and 'informal' work is particularly important as the informal sector is shown to contribute significantly to Indonesia's economy (Cuevas, Rosario, Barcenas, & Christian, 2009). Although this distinction is less clear within the Susenas, this may offer a possible explanation of the in-significance of the 'work sector' variables within Table 6.4, as these variables tend to record 'formal work' status within the sectors of employment considered.

[2] Sub-section 7.3.3 of Chapter 7 further explores the composition of the poor for the Delphi-CIMP.

caring responsibilities toward children. Although this was the case, differences in the size of the coefficients found when the two schemes were applied, may still, affect policy.

6.4.2.5 Auxiliary Robustness Check

This sub-section acts as a final step in the construction of the Delphi-CIMP. It aims to supplement the above analyses, particularly on the impact of employing different poverty cutoffs on poverty levels. Following Saltelli (2007) a composite index is often used as a tool to generate advocacy and policy based on the information it provides on 'relative positions' of countries or areas. The analysis within this sub-section is thus, of particular relevance, when seeking to compare the poverty level within Bogor City, to that in other districts within West Java itself or to the situation in Indonesia overall. Results within this sub-section are also of relevance to the discussion within Chapter 7 of this thesis, where the Delphi-CIMP is compared with the three existing measures, i.e., Consumption, Global and Adjusted MPIs. Conducting this comparative exercise necessitates the selection of a specific poverty cutoff for the Delphi-CIMP.

As shown within the above sub-sections, the size of the poor was particularly sensitive to changes in these poverty cutoffs. This was, however, less the case, when the composition of the poor was examined. To further assess this finding, the method suggested within S. Alkire and Santos (2010a), S. Alkire, Santos, Seth, and Yalonetzky (2010), Yalonetzky (2011) and Despotis (2005), is followed.[1] This method examines whether country or area ranks differ when different poverty cutoffs are employed. These 'ranks' are determined by the poverty levels within said countries or areas. Correlations coefficients between these ranks when the different cutoffs are applied, are then calculated.

Although Bogor City consists of six sub-districts and sixty-eight sub-sub-districts, data within the Susenas are not representative for these levels. Considering this, 'area' ranks within this sub-section refers to 'district' and 'provincial' ranks. Particular focus is given to 'district' ranks for the twenty-seven districts (including Bogor City), within West Java. For provincial ranks, the positions of all Indonesia's thirty-three provinces (including West Java) are examined. To determine these ranks, poverty levels according to the Delphi-CIMP, using the two weighting schemes and different poverty cutoffs, are calculated. Tables 6.5 - 6.8 within the Appendix of this chapter illustrate the results of this exercise. As shown, high levels of correlation exist between district ranks when different poverty cutoffs are applied (an average of 0.99), regardless of weighting scheme applied. Similarly, significant correlation (an average of 0.99)

[1] J. E. Foster, McGillivray, and Seth (2012) provide a discussion of other alternative methods.

was also found when provincial ranks were examined. Following Despotis (2005), S. Alkire and Santos (2010a) and S. Alkire et al. (2010), 'high correlation' indicates that these area ranks are relatively robust to the range of poverty cutoffs applied.[1] Choosing an arbitrary poverty cutoff is thus, less likely to lead to significant differences with regard to how areas are ranked. Although at times it is necessary for this arbitrary choice to be made, to note S. Anand and Sen (1997), this decision should be made as clear as possible and open to public debate (p. 6).

6.5 Conclusion

This section aims to summarise findings with regard to the two research questions addressed within this chapter, i.e. (1) What are the optimal set of 'expert' weights to utilise to combine the indicators chosen by Delphi participants? and (2) To what extent do these weights differ from the commonly applied 'equal nested' weights, when they are used to create a 'new' composite measure of poverty based on Delphi results? This 'new' measure was dubbed the Delphi Composite Index of Multidimensional Poverty (Delphi-CIMP).

To tackle Research Question (1) this chapter proposed the use of a 'linear optimisation' method, rooted within Data Envelopment Analysis (DEA) and the 'benefit-of-the-doubt' approach to weighting. This method was chosen as it enabled the incorporation of weights resulting from the Delphi exercise described within Chapters 3 and 4 of this thesis. Furthermore, the application of the 'benefit-of-the-doubt' approach within this chapter, enabled the weights resulting from this 'linear optimisation' exercise, to reflect the levels of importance Delphi participants placed on each dimension and indicator, their policy priorities and to a degree, the levels of trade-offs, participants placed between dimensions and indicators. The weights resulting from this exercise were labelled 'optimal expert' weights. When examining these weights, following the constraints set within the 'linear optimisation' exercise, it was found that their values lay within the lower and upper bound weights chosen by Delphi participants. These 'optimal expert' weights, thus, aimed to act as an objective solution to the disagreement evident within Delphi results (as indicated by range of weights suggested) on the exact values of weights to use within the final composite measure. The 'optimal expert' weights were also shown to vary across the different dimensions and indicators, reflecting a hierarchy of importance and possible policy priorities with regard to these aspects of poverty.

[1] S. Alkire and Santos (2010a) and S. Alkire et al. (2010) also found high correlation between country ranks for a range of poverty cutoffs applied to the Global MPI. This high level of correlation was utilised to support their argument for selecting an arbitrary poverty cutoff.

To address Research Question (2), comparative analysis was conducted, contrasting the effect of applying the Delphi based 'optimal expert' weights with the commonly utilised 'equal nested' weights. Three main findings were subsequently discovered. Firstly, with regard to changes in household rank when the two weighting schemes were applied. A household's total weighted deprivation was utilised to determine household rank. Spearman rank correlation coefficients between household rank when the two weighting schemes are used, were then calculated. A high level of correlation was found, indicating that households ranked as poorest when the 'optimal expert' scheme was utilised, were likely the same households as when the 'equal nested' weights were used. Policy wise, given the importance of targeting within poverty alleviation projects in developing nations, including Indonesia, the use of both weighting schemes is likely to identify a similar set of households as eligible to receive social assistance. If importance is only placed on how households are ranked, given the resource constraints governments face, the use of the more parsimonious, 'equal nested' weights, thus, may suffice.

The above in mind, the second finding, describes whether differences exist with regard to the size of the poor, when different weighting schemes are utilised. For a range of poverty cutoffs, it was found that the 'optimal expert' weighting scheme identified significantly larger proportions of the poor, compared to the 'equal nested' scheme. Even though, household ranks did not differ largely, this significant difference with regard to the size of the poor, could however, lead to substantial miss-targeting, when one approach to weighting is used as opposed to the other. Using the 'optimal expert' weighting scheme as a benchmark, the use of the 'equal nested' weights, could underestimate the level of poverty within the case study area. If targeting policies are in use, basing identification on the measure using 'equal weights' could likely increase the possibility of eligible poor households not receiving assistance.

Thirdly, with regard to the composition of the poor. It was found that both weighting schemes identify households with relatively similar characteristics. These households were found to be headed by older individuals, who are less educated and live within households with additional caring responsibilities toward children. This result seems to strengthen the results of the household rank-correlation exercise summarised above. This also confirms prior analysis within existing literature conducted to contrast the use of 'expert' weights generated through a DEA benefit-of-the-doubt approach vs. equal weights (L. Cherchye et al., 2008; Despotis, 2005; Mahlberg & Obersteiner, 2001). Despite these similarities, difference in the *sizes* of the effects of the household characteristics considered, particularly those of the 'education level of the household head' and the 'number of children within the household', on the probability of being

'optimal expert' Delphi-CIMP poor vs. being 'equal nested' Delphi-CIMP poor, may still lead to differing policy conclusions.[1] Thus, following Despotis (2005), the 'optimal expert' scheme seem to offer a level of superiority as it is 'less arbitrary' and thus, renders the composite measure resulting from its use 'less contestable'. With regard to this, Alatas et al. (2013) and Alatas et al. (2012) highlight the importance of 'acceptability' when applying poverty alleviation policies, particularly those involving targeting methods in Indonesia. The use of the Delphi method, created a 'sense of ownership' of the resulting measure, since suggestions made by participants were incorporated into the measure in a visible way; this in turn bolstered confidence in the measure's utility.[2]

A final note is now given on the auxiliary robustness checks conducted within this chapter. As found above, the use of the different weighting schemes led to differences when the size of the poor was examined. Less pronounced differences were, however, present when household rank and composition were examined. Other than the effect of weights, the impact of an additional source of uncertainty within measurement, i.e., poverty cutoffs, was examined. It was found that different poverty cutoffs led to differences in the size of the poor. Significant differences were, however, *not present* when examining the composition of the poor. Furthermore, analysis of district and provincial ranks based on their headcount poverty when a range of poverty cutoffs were employed, confirmed that similar ranks when assigned, even though different cutoffs were employed. Thus, in general, except for a significant effect on the size of the poor, the results within this chapter were shown to be robust to a substantial range of poverty cutoffs.

To conclude, this chapter illustrated the feasibility of incorporating results from a Delphi study into a composite measure of poverty. Acknowledging the limitations of the Delphi proposed within this thesis, it is hoped that discussion on the use of this method, will generate additional research on the involvement of local stakeholders within the building of poverty measures. As Nardo et al. (2005) state, 'composite measures should never be seen as a goal per se; they should be seen, instead, as a starting point for initiating discussion and attracting public interest and concern' (p. 7).

[1] Similar results are found if data from West Java and Indonesia is used; i.e. household ranks determined by the two weighting schemes are highly correlation, however, the extent of poverty identified by the two schemes significantly differ.

[2] Bogor City's Mayor, Vice Mayor and Head of the District Statistical and Development Body came to Trinity College, University of Oxford in October 2018, to outline the policy steps and concrete action toward poverty alleviation, taken based on the Delphi CIMP (Trinity College, 2018).

Chapter 7: Comparing the Delphi Composite Index with Existing Measures

> *The diversity of possible judgments should be taken into account in the measurement [of poverty]. While this may permit less all-embracing conclusions, it does offer the prospect of unambiguous conclusions in certain circumstances. Equally important, results that do depend on the choices made on the course of measurement can be recognized as such*
>
> *(Nolan & Whelan, 1996, p. 2)*

7.1 Introduction

As a continuation to the analysis within previous chapters, this chapter forms the final step in fulfilling the objective of illustrating how the Delphi may be utilised as a novel methodological approach to address the challenge of defining and measuring poverty in a way that is 'valid and meaningful for the context in which measurement is needed'; a methodology, which ultimately aims to support sound and reliable policy-making (Nolan & Whelan, 1996, p. 1).

Within Chapter 2 of this thesis, we saw that when examining three existing measures, Consumption, the Global and Adjusted MPIs, differences were found with regard to how these measures identified the size and composition of the poor. Reliance on any one of these measures as a single yardstick of policy effectiveness, may thus be misleading. Responding to these results, within Chapter 3, a Delphi study was proposed as a novel way of addressing the arbitrariness in the way existing measures of multidimensional poverty selected components within them, i.e. their dimensions, indicators and weights. Although far from offering a tell-all solution to the myriad challenges faced when measuring multidimensional poverty, the Delphi method that was suggested, offered a transparent and robust way to mitigate this challenge. Furthermore, results of the study, presented within Chapter 4, confirmed the importance of utilising a participatory platform, such as that offered by the Delphi, in uncovering the diversity of possible judgments behind this selection process.

Chapters 5 and 6 of this thesis, then aimed to translate the Delphi's results into concrete measurement, with Chapter 5 focusing on incorporating these results within a dashboard measure, whilst Chapter 6 examined the effects of incorporating Delphi opinions, particularly those on weights, within a composite measure. Findings within Chapter 6 established that differences, primarily with regard to the extent of poverty identified, existed when 'expert' weights, which aimed to reflect opinions generated through the Delphi process, as opposed to the commonly applied 'equal nested' weights, were used.

Considering the above, the goal of this chapter is to further illustrate the implications of utilising the Delphi method within the measurement of poverty. To continue the discussion within Chapter 2, this chapter conducts a comparative exercise, comparing the Delphi-CIMP with the three existing measures, the Consumption, Global and Adjusted MPIs, to further analyse the possible impact of the choices made within measurement, i.e. the 'menu' of dimensions and indicators, alongside their weights, on the size of the poor, how households are ranked according to their total deprivation and the characteristics of households identified as poor. This in mind, this chapter aims to tackle four research questions.

The first question is thus: *Does the Delphi-CIMP identify a different size of the poor compared to the three existing measures?* The second research question asks: *Do household ranks differ when different measures are utilised?* The third then questions: *Does the Delphi-CIMP identify a different composition of the poor, when compared to the existing measures?* Finally, the last research question within this chapter, seeks to uncover whether the results of the above questions, hold when the analysis is conducted at three geographic levels, i.e. district, provincial and national: *Do the comparative results with regard to how the measures identify the size and composition of the poor, alongside household ranks, differ when different levels of analysis are considered?* Addressing this final question seeks to shed light on whether policy makers should indeed, consider applying district specific measures, such as that developed within this thesis, when assessing poverty. This in mind, differing from Chapters 4, 5 and 6, which focused discussions on results within the case study, i.e. Bogor City, following Chapter 2, this chapter also conducts analysis using data from West Java and Indonesia as a whole.

The four research questions above are explored using Susenas 2013 data. The household is again, considered as the unit of analysis. The rest of this chapter is thus, organised as follows: Section 7.2 continues the discussion within Chapter 2 and re-examines the possible sources of difference between the four measures analysed within this chapter, Consumption, Global MPI, Adjusted MPI and the Delphi-CIMP. Section 7.3 presents results and Section 7.4 concludes.

7.2 Review of Measure Components and Expectations

Before reviewing the components included in each of the four measures, a note of clarification needs to be made with regard to two aspects of the Delphi-CIMP. Firstly, within Chapter 6, two composite measures were created based on Delphi results, the 'optimal expert' Delphi-CIMP and the 'equal nested' Delphi-CIMP. As a brief overview, the 'optimal expert' Delphi-CIMP included not only the dimensions and indicators selected by Delphi participants, but also took

into account opinions on how these components should be weighed. The 'equal nested' Delphi-CIMP, on the other hand, included the dimensions and indicators resulting from the Delphi, however, ignored information on weights. Considering this, in order to fully account for the opinions gathered through the Delphi, this chapter focuses on the 'optimal expert' Delphi-CIMP.

Secondly, a note needs to be made with regard to the choice of poverty cutoff for the Delphi-CIMP. Chapter 6, particularly within sub-section 6.4.2.5, analysed the effect of utilising a range of these cutoffs on changes in the magnitude of poverty, composition of the poor and whether district and provincial ranks differed when different cutoffs are utilised. Results indicated that, although the poverty cutoff significantly affected the size of the poor identified, the composition and district/provincial ranks were found to be relatively robust to the choice of these cutoffs. This in mind, this chapter utilises an arbitrary intermediate poverty cutoff, situated between the two extremes of the 'intersection' and 'union' cutoffs, to define poverty using the Delphi-CIMP. This cutoff is: *a household is poor if its standardised weighted total deprivation is equal to or above one standard deviation away from its mean*. As will be shown within the results of this chapter, this poverty cutoff also leads to a size of the poor that lies somewhat between that identified by the Consumption measure and the two existing multidimensional measures.

With the above in mind, this section aims to highlight aspects within the Delphi-CIMP, which may lead to differences in how it identifies the poor, compared to the three existing measures. I start this section by comparing Delphi-CIMP with the Consumption measure.

7.2.1 Delphi-CIMP vs. Consumption

This sub-section aims to pinpoint aspects in which these two measures differ, which in turn may offer clarity as to *why* differences, as will be presented within the results of this chapter, emerge with regard to how these measures quantify and identify the poor. As described within Chapter 2, the Consumption measure defines poverty as the inability in monetary terms of a household to fulfil its 'basic needs' in the forms of food and non-food items. A household is considered as poor if its monthly consumption level is below a 'poverty line', which is defined as the level at which the household may adequately fulfil a daily need of 2,100 kilo calories per day per individual within it and meet its needs in non-food items, such as health, education and living arrangements (Central Bureau of Statistics Indonesia (CBS), 2014; Iriana et al., 2012; Martin Ravallion, 1994, p. 75).

Although one may argue that the Consumption measure encompasses multiple aspects describing a household's daily needs, this measure focuses only on one facet of these goods

and services; i.e. their monetary value.[1] As a result, the Consumption measure relies on market prices to aggregate these aspects of daily needs, as opposed to the use of 'weights', whether set by analysts or as the Delphi study proposed, takes into account the views of those involved within poverty alleviation. This, is the area, where consumption and most multidimensional measures differ (Martin Ravallion, 2011b, p. 236). In addition to this, the use of more direct indicators of deprivation, such as the inability to read and write, as opposed to a household's 'potential' of being able to send their children to school as reflected by its monthly consumption expenditure, may generate a different magnitude of poverty and could result in different household characteristics leading to the experience of poverty.

As described within Chapter 6, the Delphi-CIMP includes five dimensions and eighteen indicators, which encompass various areas of deprivation.[2] Different from the Consumption measure, which bases its selection of goods and services to include, on the 'consumption pattern' of a 'representative population', the Delphi-CIMP employed the judgment of 'experts' in selecting its components. Although, within this thesis, the term 'expert' was limited to policy makers from one district, considering these views, may also lead to differences with regard to the extent and identity of those who are regarded as poor. Little is known, notably for the context of Indonesia, of the possible differences that may arise due to the use of multidimensional measures, specifically those which aim to encompass the views of experts, in particular, opinions on weights versus consumption-based measures. The comparative exercise within this chapter is thus explorative. This in mind, the first few rows of Table 7.1 list expectations (rather than established hypotheses) on the possible relationship between these two measures. Existing literature comparing multidimensional measures, which for the most part employed 'equal weights', versus consumption-based measures, is referred to, in order to form these expectations. In addition to this, the findings of Chapter 2 are also drawn upon.

This sub-section focuses on existing evidence available concerning multidimensional poverty in Indonesia. Although as noted within previous chapters, literature on this issue is limited, evidence found, points toward a 'mismatch' of outcomes when a multidimensional rather than uni-dimensional framework is applied to measure poverty.

[1] Martin Ravallion (2011b) states that 'multidimensionality per se cannot be what distinguishes a multidimensional index of poverty (MIP); indeed, the only truly one-dimensional indices are the rice-based measures once found in some countries in Asia, but no longer in use' (p. 236).
[2] As listed within Table 6.1 of Chapter 6, the Delphi-CIMP also includes 'Household Consumption', however as one out of the five indicators, which proxy deprivation within its 'Living Standards' dimension.

Table 7. 1 Expected Relationship between the Delphi-CIMP and the Existing Measures

Measures	Expected Results
Delphi CIMP vs. Consumption	**Size** Delphi CIMP poor > Consumption poor (dependent on poverty cutoff) **Rank** Significant difference in how measures rank the poor **Composition** Significant differences and similarities in the types of households identified as poor
Delphi CIMP vs. Global MPI	**Size** Delphi CIMP poor > Global MPI poor (dependent on poverty cutoff) **Rank** Significant difference in how measures rank the poor **Composition** Significant differences and similarities in the types of households identified as poor
Delphi CIMP vs. Adjusted MPI	**Size** Adjusted MPI poor > Delphi CIMP poor (dependent on poverty cutoff) **Rank** Significant difference in how measures rank the poor **Composition** Significant differences and similarities in the types of households identified as poor

Source: Author's summary

As touched upon within Chapter 1, Sumarto and de Silva (2014) discovered low levels of overlap between their proposed multidimensional measure, a measure aiming to replicate the Global MPI, and Indonesia's Consumption measure.[1] Furthermore, these authors concluded that the Consumption measure tended to underestimate the extent of poverty in Indonesia, when compared to their 'newly' constructed multidimensional index. These authors found that this conclusion remained valid for the entirety of the time period analysed, i.e. years 2004 to 2013.

[1] Sumarto and de Silva (2014)'s MPI included the Global MPI's three dimensions; i.e. 'Education', 'Health' and 'Living Standards'. Differences were, however, found within the indicators included in their version of the Global MPI. To illustrate, within their 'Health' dimension, 'Vaccination' and 'Birth Attendance' were utilised to replace the Global MPI's 'Nutrition' indicator, and within their 'Living Standards' dimension, 'Employment' was utilised to replace the Global MPI's 'Assets' indicator. Their measure, following the Global MPI, utilised equal nested weights.

The above results were further strengthened within Hanandita and Tampubolon (2015). These authors formed another 'new' index of multidimensional poverty, which included dimensions that mirrored those within the HDI, i.e. income, health and education.[1] They then utilised five indicators, which included 'per capita consumption', to proxy deprivation within these dimensions. Findings from this study confirmed tendency of the consumption measure to under-report poverty levels. In addition to this, they revealed that 'consumption expenditure' was only 'weakly correlated' with deprivations within other domains of wellbeing. A point of particular interest within Hanandita and Tampubolon (2015)'s analysis, was that even though the multidimensional measure they proposed included 'consumption' as one of its five indicators, the inconsistencies found between the 'sole use' of consumption as a measure of poverty, and a measure, which encompassed other aspects, continued to persist. They emphasised the possible impact of these inconsistencies by stating that if 'targeting policies' were solely based on consumption without considering the multidimensional nature of poverty, some '4.5 million adult Indonesians' would be miss-categorised as non-poor, thus deeming them ineligible for social assistance (Hanandita & Tampubolon, 2015, p. 574).

Finally, evidence within Chapter 2 of this thesis, which compared the Global and Adjusted MPIs, with Consumption, further confirmed the findings summarised above. For the 'expectations' noted within Table 7.1, these results indicate that despite the different configurations and components utilised to measure multidimensional poverty in Indonesia, when it comes to the *size of the poor*, Consumption is consistently shown to possess the tendency of identifying a lower number of the poor when compared to multidimensional measures. This chapter thus, expects that this relationship will hold, if the Delphi-CIMP is compared against consumption.[2] Similar to this, on the *household ranks*, this chapter also expects the conclusions within existing literature, to persist; i.e. given the low correlation found between deprivation in consumption and other types of deprivation, *household ranks* may be expected to be only modestly correlated, when the Delphi-CIMP and the consumption measures are considered.

As for the composition of the poor, the results of the comparative exercise conducted within Chapter 2 are specifically drawn upon to form expectations for this issue. Following these

[1] Within Hanandita and Tampubolon (2015), these three dimensions were assigned equal weights.
[2] With regard to the size of the poor, however, as noted at the start of this section, an 'arbitrary' intermediate poverty cutoff was utilised to calculate the Delphi-CIMP. As described within Chapter 6, even though household composition and area ranks, were found to be robust to this choice of cutoff, the size of the poor, was highly dependent on this cutoff. This expectation is thus, also dependent on the cutoff utilised. Section 7.3 of this chapter discusses this further.

findings, the Delphi-CIMP is expected to identify relatively different households as poor, compared to the Consumption measure. Noting this however, similarities between these two measures may also be found; as concluded within Chapter 2, regardless of the measure, households headed by those with lower levels of education had a persistently high likelihood of being considered as poor.

7.2.2 Delphi-CIMP vs. Existing Multidimensional Measures

This section contrasts the Delphi-CIMP with one of the most widely applied multidimensional measures within the developing world, i.e., the Global MPI, alongside a localised 'national' application of this measure, the Adjusted MPI. As noted within Chapter 2, little or no previous literature exists, which compares the use of locally developed multidimensional measures with international measures, such as the Global MPI, particularly for the context of Indonesia. This in mind, alongside noting the results within Chapter 2, this section primarily places attention on a re-examination of the components within the multidimensional measures analysed in this chapter, seeking to highlight possible sources of variation between them.

There are three main sources of variation between these measures, i.e. differences with regard to (1) the number and types of dimensions and indicators included within them, (2) the choice and level of stringency of the indicator deprivation cutoffs utilised and (3) the weights used to aggregate these measures. Table 7.2 examines the first source of possible difference. This table presents a list of *all* dimensions and indicators included within the three multidimensional measures examined. The "+" sign indicates the presence of a dimension or indicator within an index, whilst the "-" indicates that the said dimension or indicator is not included within a specific measure.

As can be seen, the two existing measures include three dimensions, whilst the Delphi-CIMP includes five. Thus, compared to the Delphi-CIMP, the two existing multidimensional measures are shown to 'miss' the inclusion of the 'Employment' and 'Contraception' dimensions. Looking at indicators, the Delphi-CIMP incorporates nearly twice the number of indicators included within the Global MPI, with eighteen indicators, as opposed to the Global MPI's ten and the Adjusted MPI's eleven.

Looking at the types of the indicators, Table 7.2 shows that even within similar dimensions, for each measure, a number of different indicators are used to proxy deprivation within the said dimensions.

Table 7. 2 Dimensions and Indicators Included within all Three Multidimensional Measures [(a)]

Dimension/Indicator	Delphi-CIMP	Global MPI	Adjusted MPI
Education	**+**	**+**	**+**
Children's' education	+	-	-
Adult education	+	-	-
Literacy	+	-	+
Ownership of birth certificate	+	-	-
Years of schooling	-	+	-
School attendance	-	+	-
Access to preschool education	-	-	+
Education continuity	-	-	+
Employment and Assets	**+**	**-**	**-**
Employment	+	-	-
Ownership of small assets	+	-	-
Ownership of house/land	+	-	-
Living Standards	**+**	**+**	**+**
Consumption (household consumption)	+	-	-
Sanitation	+	+	-
Drinking water	+	+	-
Cooking and washing water	+	-	-
Government help	+	-	-
Cooking fuel	-	+	+
Electricity	-	+	+
Housing (adequate floor, wall or roof)	-	+	-
Housing (adequate floor, wall and roof)	-	-	+
Assets	-	+	-
House/Land ownership	-	-	+
Health and Safety	**+**	**-**	**-**
Healthy life	+	-	-
Access to immunization	+	-	-
Safety	+	-	-
Child Health and Contraception	**+**	**-**	**-**
Birth attendant	+	-	-
Child mortality	+	-	-
Contraception	+	-	-
Health	**-**	**+**	**+**
Nutrition	-	+	-
Child mortality	-	+	-
Sanitation	-	-	+
Clean water	-	-	+
Under 5 nutritional intake	-	-	+
Birth attendant	-	-	+

Notes: [(a)] "+" indicates presence of dimension/indicator within measure; "-" indicates non-presence of a dimension/indicator.
Source: Author's summary

To illustrate, within the 'Education' dimension, the Global and Adjusted MPIs, are shown to exclude the Delphi-CIMP's 'Ownership of Birth Certificate' indicator. In addition to this, the Delphi-CIMP's indicator measuring 'Safety' is excluded by the 'Health' dimension within both the Global and Adjusted MPIs.

The indicators measuring 'Employment' are also excluded within the two existing measures. In contrast, for the 'Living Standards' dimension, a number of indicators, which were thought of as less crucial by Delphi participants, such as 'Cooking fuel', 'Electricity' and 'Types of floor, wall and roof', are included within the two existing measures. Likewise, within the 'Health' dimension, the Delphi-CIMP does not include the Adjusted MPI's 'Under 5 Nutritional Intake' indicator.

When comparing the Delphi-CIMP with the Adjusted MPI, in particular, it is interesting to note that despite both aiming to be specifically relevant for the context of Indonesia, albeit one at district and the other at national level, differences in the types of indicators included continued to be present. This may be due to the fact that the Delphi-CIMP was formed to be relevant within a predominantly urban area, whereas the Adjusted MPI aimed to reflect poverty within Indonesia overall, which encompasses areas which are predominantly rural. These differences in the types of indicators included, could point toward the relevance of considering district specific measures, particularly for a nation as diverse as Indonesia.[1]

In addition to the types of indicators, as noted above the second point of difference between the multidimensional measures examined in this chapter, is the severity of cutoffs utilised to define deprivation within these indicators. Table 7.3 below summarises these cutoffs. Within this table, only the indicators included in the Delphi-CIMP are examined. The '+' symbol indicates the presence of an indicator within the measures examined. When an indicator is present within more than one measure, the '+' symbol specifies the cutoff that is the *least* stringent. The *most* stringent cutoff is then depicted with the '+ + +' symbol.[2]

[1] As an additional consideration, results from the post-Delphi semi-structured interviews described within Section 4.7 of Chapter 4, confirmed the importance of considering district specific, ground level, measures of poverty in Indonesia.

[2] Arguments behind the selection of these cutoffs are presented within Chapter 2, for the Global and Adjusted MPIs and within Chapter 5, for the Delphi-CIMP.

Table 7. 3 Comparison of cutoffs employed to measure deprivation within Delphi-CIMP indicators **(a) (b) (c) (d) (e)**

	Delphi CIMP	Global MPI	Adjusted MPI
Indicator	Level of stringency of indicator cutoff	Level of stringency of indicator cutoff	Level of stringency of indicator cutoff
Children's Education [(f)]	+ +	+	+ + +
Adult's Education[(g)]	+ +	+	N/A
Literacy	+	N/A	+
Ownership of Birth Certificate	+	N/A	N/A
Employment	+	N/A	N/A
Ownership of Assets	+	+ +	N/A
House/Land Ownership	+	N/A	+
Consumption	+	N/A	N/A
Sanitation	+	+	+
Drinking Water	+ +	+	+ +
Cooking and Washing Water	+	N/A	N/A
Government Help	+	N/A	N/A
Healthy life	+	N/A	N/A
Access to immunizations.	+	N/A	N/A
Safety	+	N/A	N/A
Birth attendant	+	N/A	+
Child mortality	+	+	N/A
Contraception	+	N/A	N/A

Note:
[(a)] The reference point for this table is the indicator cutoffs employed within the Delphi-CIMP.
[(b)] Only Delphi-CIMP indicators are included within this table. As shown within Table 7.2, the other two measures include other indicators not listed within this table.
[(c)] '+' indicates lowest level of stringency within the cutoff; '+ + +' indicates highest level of stringency within the cutoff; and 'N/A' indicates that an indicator is not included within a measure.
[(d)] If levels of stringency are the same the '+' symbol is used for all measures.
[(e)] The '+' symbol is also used to indicate presence of an indicator, regardless of stringency, when indicator is not present within other measures.

[(f)] The Adjusted MPI labels its 'Children's education' indicator, 'Education continuity'. The Global MPI labels its 'Children's education' indicator, 'School attendance'.

[(g)] The Global MPI measures 'Adult education' within its 'Years of schooling' indicator.

Source: Author's summary

Considering these levels of stringency, Table 7.3 shows that the main difference between the three measures, lies within cutoffs for the 'Education' and 'Assets' indicators. The discussion below thus, places particular focus on these indicators.

Despite all acknowledging the importance of 'Education', as noted within Table 7.2, the three measures defined and labelled indicators concerning 'Children's Education' differently. For this indicator, the Delphi-CIMP aimed to capture 'access to pre-school education' for those aged three to six, alongside, whether school aged children (aged seven to sixteen) attended nine years of compulsory education. As a comparison, the Global MPI's indicator relating to 'Children's Education', i.e. 'School attendance', aimed to examine whether children up to the age at which they should be attending Grade 8, were indeed, undergoing schooling. On the other hand, the Adjusted MPI includes two indicators to measure 'Children's education', i.e. the first measuring access to 'pre-school education' and the second, i.e. 'Education continuity', assessing whether children within the household complete high school.

As described within Table 7.3 the three measures also assigned *different thresholds* defining deprivation within these various 'Children's Education' indicators. Compared to the Delphi-CIMP, the Global MPI employs a *less* stringent cutoff for its 'Children's Education' indicators, i.e. that all children within the household should have access to education until they are in Grade 8, as opposed to the Delphi-CIMP's criteria, which necessitates that all children within the age groups specified above, must have access to pre-school education and must at least attend 9 years of compulsory education until they complete Grade 9 (Alkire et al., 2016; Republic of Indonesia Ministry of Education and Culture, 2008). The Adjusted MPI is then shown to assign a *more stringent* criterion compared to the Delphi-CIMP, stipulating that all children should have access to pre-school education and must complete their high school education (Budiantoro et al., 2020; Prakarsa, 2016).[1]

As noted within Chapter 2, enrolment rates, particularly at secondary and high school levels continue to present major challenges for Indonesia (G. W. Jones & Pratomo, 2016; Ministry of Education and Culture Republic of Indonesia, 2019; Suryadarma & Jones, 2013). Chapter 2 showed that even subtle differences in stringency with regard to children's school enrolment may lead to significant differences with regard to the number of deprived households recorded

[1] As described within Chapter 2, this criterion was likely set as a result of the Government of Indonesia's renewed commitment toward achieving universal access to twelve years of education (Republic of Indonesia Ministry of Education and Culture, 2016). This commitment however currently acts as a development goal, unlike the 'nine-year compulsory education' requirement, which is mandatory by law (Republic of Indonesia Ministry of Education and Culture, 2008).

within these indicators.[1] Indeed, the significant difference in stringency with regard to the Adjusted MPI's cutoff for its 'Education continuity' indicator, compared to the cutoff employed for the Global MPI's 'School Attendance' indicator, contributed to the Adjusted MPI identifying nearly twice the amount of poor households, compared to the Global MPI.

Another area of difference between the three measures, is found when the cutoff for 'Adult Education' is examined. As listed within Table 7.2, the 'Adult education' indicator is missing within the Adjusted MPI. Whereas within the Global MPI, it is measured within the measure's 'Years of schooling' indicator. Compared to the Delphi-CIMP, the Global MPI, again, assigns a less stringent cutoff for this indicator, only requiring at least one member of the household to obtain at least *six years* of education. For the Delphi-CIMP, however, at least one adult within the household must complete their *nine years* of compulsory education.[2] Again, this difference in stringency, may affect the number of poor households identified by these measures.

Another point of difference with regard to indicator cutoff stringency can be found for the 'Asset' indicator. Table 7.3 shows that the Delphi-CIMP assigns a less stringent cutoff when compared to the Global MPI, requiring households to have at least one type of asset valued at IDR500,000, whereas the Global MPI requires households to own at least two assets and a car/truck.[3] This difference could also contribute toward how these two measures identify the size of the poor.

The final point of difference between the three measures lies within how dimensions and indicators are *weighed*. Table 7.4 lists all the dimensions and indicators included within the three measures alongside their weights. As can be seen, the Global and Adjusted MPIs utilised 'equal nested' weights, i.e. their dimensions are weighted equally. Indicators *within* each dimension are also given equal weights, however, across the dimensions, these weights differ.

[1] Table 7.7 of the Appendix to this chapter notes that the number of deprived households within the 'Children's education' indicator as defined by the Delphi-CIMP was 24 percent in Bogor City. The Global MPI's 'School attendance' indicator identified 3 percent; whilst the Adjusted MPI's 'Education continuity' indicator identified 71 percent of total households in Bogor City as deprived.

[2] Referring to Table 7.7 in the Appendix to this chapter, the Delphi-CIMP's 'Adult education' indicator identified 16 percent of total households in Bogor City as deprived in this indicator. The Global MPI's 'Years of schooling' identified 2.5 percent of total number of households as deprived.

[3] Again referring to Table 7.7, this difference in stringency led the Global MPI's definition of 'Asset Ownership' to identify 24 percent of total households in Bogor City to be deprived; whereas the Delphi CIMP's definition identified 4 percent.

Table 7. 4 Weights for All Dimensions and Indicators Included within the Three Multidimensional Measures [a] [b]

Dimension/Indicator	Dimension/Indicator Weights		
	Delphi CIMP	Global MPI	Adjusted MPI
Education	**0.29**	**0.33**	**0.33**
Children's' education[c]	0.1	0	0
Adult education[c]	0.03	0	0
Literacy[c], [e]	0.15	0	0.11
Ownership of birth certificate[c]	0.02	0	0
Years of schooling[d]	0	0.17	0
School attendance[d]	0	0.17	0
Access to pre-school education[e]	0	0	0.11
Education continuity[e]	0	0	0.11
Employment and Assets	**0.18**	**0**	**0**
Employment[c]	0.08	0	0
Ownership of small assets[c]	0.06	0	0
Ownership of house/land[c]	0.04	0	0
Living Standards	**0.14**	**0.33**	**0.33**
Household consumption level[c]	0.05	0	0
Sanitation[c], [d]	0.02	0.06	0
Drinking water[c], [d]	0.02	0.06	0
Cooking and washing water[c]	0.02	0	0
Government help[c]	0.03	0	0
Cooking fuel[d], [e]	0	0.06	0.08
Electricity[d], [e]	0	0.06	0.08
Housing (adequate floor, wall or roof) [d]	0	0.06	0.08
Housing (adequate floor, wall and roof) [e]			
Assets[d]	0	0.06	0
House/Land ownership[e]	0	0	0.08
Health and Safety	**0.26**	**0.33**	**0.33**
Healthy life[c]	0.09	0	0
Access to immunization[c]	0.03	0	0
Safety[c]	0.14	0	0
Nutrition[e]	0	0.17	0.08
Child mortality[d]	0	0.17	0
Sanitation[e]	0	0	0.08
Clean water[e]	0	0	0.08
Birth attendant[e]	0	0	0.08
Child Health and Contraception	**0.12**	**0**	**0**
Birth attendant[c]	0.07	0	0
Child mortality[c]	0.03	0	0
Contraception[c]	0.03	0	0
Sum of Dimension weights	1	1	1
Sum of Indicator weights	1	1	1

Note: [a] Dimensions and indicators listed within this table are those included within all three measures; i.e. Delphi-CIMP, Global MPI and Adjusted MPI.

[b] The dimensions listed are Delphi CIMP dimensions. Some indicators are mentioned twice, as they are included within different dimensions by each different measure.

[c] Indicates that the indicator is included within the Delphi-CIMP.

[d] Indicates that the indicator is included within the Global MPI.

[e] Indicates that the indicator is included within the Adjusted MPI.

Source: Author's summary

As noted within Chapter 6, this difference in indicator weights somewhat questions the theoretical underpinnings behind the use of equal weights, i.e. that all aspects of poverty should be given equal importance. This issue is even more evident when comparing these weights with those utilised within the Delphi-CIMP. The Global and Adjusted MPIs are shown to assign zero weights to the dimensions and indicators included in the Delphi-CIMP, which have been excluded within their measurement. Compared to the two existing measures, the Delphi-CIMP employed 'expert' generated weights resulting from an application of the 'benefit-of-the-doubt' approach to weighting, which took into account opinions on weights, as expressed by Delphi participants. As stated within Chapters 3, 4 and 6 of this thesis, these weights aimed to reflect the degrees of importance and possible trade-offs between the dimensions and indicators included within the Delphi-CIMP.

The above in mind, within Table 7.4, it can be seen that the Global and Adjusted MPIs assigned higher weights to their 'Education' dimensions. As a result, except for 'Literacy', indicators within 'Education' are also given higher weights when compared to those used within the Delphi-CIMP.

Particularly so for the Adjusted MPI, these higher weights, coupled with the higher level of stringency of the cutoffs employed for these indicators, may likely lead this measure to identify more households as poor when compared to the Delphi-CIMP. For the Global MPI, as the cutoffs were less stringent, thus, even though higher weights were assigned to its 'Education' indicators, this measure may still tend to identify less households as poor compared to the Delphi-CIMP.

Looking within the 'Employment and Assets' dimension, the Global and Adjusted MPI's exclusion of the 'Employment' indicator, could contribute to a decrease in the number of the poor identified by these measures when compared to the Delphi-CIMP.[1] Furthermore, the Global MPI, unlike the Adjusted MPI, assigns a weight of zero to the Delphi-CIMP's 'House/Land Ownership' indicator.[2] This may also lead this measure to identify less poor, compared to the Delphi-CIMP. On the other hand, when examining 'Asset' indicators more closely within Table 7.4, the Global MPI is shown to assign a similar weight to its 'Ownership of Assets' indicator, compared to that assigned by the Delphi-CIMP. Even though these similar weights are used, the more stringent cutoff used by the Global MPI for this indicator noted within

[1] Referring to Table 7.7 in the Appendix to this chapter, five percent of households in Bogor City exhibit deprivation in this indicator.
[2] Thirty-two percent of households in Bogor City exhibit deprivation in this indicator.

Table 7.3, may still be a factor, which acts to increase the number of poor households identified by the Global MPI compared to the Delphi-CIMP.[1]

As for the 'Living Standards' indicators, the Delphi-CIMP is shown to assign zero weights to a number of indicators considered important within the Global and Adjusted MPIs, i.e. 'Cooking Fuel', 'Electricity' and 'Quality of Housing Material'. When examining levels of deprivation within these excluded indicators, particularly for Bogor City, Table 7.7 within the Appendix, shows that relatively low numbers of households are deprived within these indicators.[2] The exclusion of these indicators within the Delphi-CIMP, may thus, not lead to a significant increase in the number of poor households identified by this measure.

Within the 'Health' dimension, the Delphi-CIMP's exclusion of the 'Nutrition' indicator, which is included within the Adjusted MPI, could lead to a decrease in the number of poor households identified as poor by this measure, when compared to the Adjusted MPI; as Table 7.7 shows, close to 36 percent of households in Bogor City exhibit deprivation in this indicator. The Global and Adjusted MPIs, are also shown to exclude a substantial number of indicators in 'Health', particularly those within 'Child Health and Contraception'. This exclusion of indicators may also contribute to these measures identifying fewer poor households, when compared to the Delphi-CIMP.

Considering the above, coming back to Table 7.1, three expectations are made with regard to the relationship between the Delphi-CIMP and the two existing multidimensional measures examined. Firstly, with regard to the size of the poor, as noted within Table 7.2, the Delphi-CIMP encompasses more aspects of well-being in which households may be deprived in compared to the two existing measures; households may thus, be more likely to be identified as poor by this measure, compared to the Global and Adjusted MPIs. The level of indicator cutoff stringencies listed within Table 7.3 and weights noted within Table 7.4, particularly with regard to the 'Education' indicators, however point to the possibility of the Adjusted MPI identifying the largest size of the poor, when compared to the Delphi-CIMP and Global MPI. Given these

[1] Again, referring to Table 7.7, the Global MPI's cutoff for 'Ownership of Assets' identified six times more deprived households, compared to the Delphi-CIMP's.
[2] Only 0.1 percent of households in Bogor City are deprived in 'Electricity', 0.2 percent are deprived in 'Roof, Floor and Wall', 3.5 percent are deprived in 'Cooking Fuel'.

considerations, one may expect that the Delphi-CIMP would identify more households as poor compared to the Global MPI, however, less, when compared to the Adjusted MPI.[1]

Secondly, with regard to how these measures rank the poor, as noted within the discussion above, results from Chapter 2 show that the Adjusted MPI assigned significantly different ranks when compared to the Global MPI. Differences were also present with regard to the composition of the poor identified by these two existing measures. This was present even though, as summarised within Tables 7.2-7.4 above, the Adjusted MPI, for the most part, included similar dimensions and indicators, and applied a similar weighting scheme to that within the Global MPI. Considering this, given the notable differences between the components of the Delphi-CIMP and those within the two existing measures, one could thus, also expect differences with regard to how households are ranked to arise when comparing these measures. Thirdly, these differences in measure components, could also lead us to expect these three measures to identify different types of households as poor.

7.3 Results

7.3.1 Size of the Poor

This sub-section addresses this chapter's first research question; i.e. do poverty levels differ when using the Delphi-CIMP compared to the three existing measures examined within this thesis? As noted within the Introduction of this chapter, to address this question, data from the district, provincial and national levels are analysed. Considering this, this sub-section thus, also aims to shed light on the fourth research question: whether conclusions found with regard to Research Questions (1) hold when analysis is conducted at these three levels. Table 7.5 illustrates the results. Three main findings are suggested:

Firstly, examining within and across each geographical level, Consumption poverty is shown to identify the *lowest* level of poverty, compared to all the multidimensional measures examined, regardless of geographical level. This result is in line with that found within existing literature, and the expectations listed within Table 7.1 of this chapter.

Secondly, the Adjusted MPI is shown to identify the largest number of poor households compared to the other measures, regardless of the geographical level of analysis. Noting the discussion within sub-section 7.2.2, the Adjusted MPI's high stringency level of its 'Education'

[1] Although, as noted at the start of sub-section 7.2.1, the size of the poor identified by the Delphi-CIMP is dependent on the poverty cutoff employed..

indicator cutoffs may have contributed to this.[1] In addition to this, the Adjusted MPI's inclusion of the 'Under 5 Nutritional Intake' indicator, an area where a significant number of households are shown to experience deprivation, could also have led to a significant increase in the number of poor households, which this measure identifies compared to the Delphi-CIMP.[2]

The third finding points to differences in results, when comparing the Delphi-CIMP with the two existing multidimensional measures, if different geographical levels of analysis are considered. Table 7.5 shows that the Delphi-CIMP identified *more* poverty compared to the Consumption measure and *less* compared to the Adjusted MPI, regardless of geographical level. This is, however, not the case when the Delphi-CIMP is compared with the Global MPI. Within Bogor City (at district level), the Delphi-CIMP identifies a higher number compared to the Global MPI, yet at the provincial and national levels, the Delphi-CIMP reports lower numbers of poverty compared to the Global MPI.

Table 7. 5 Proportions of the Poor According to the Four Measures

Poverty Measure	Lower Bound[(a)]	Poor Households	Upper Bound[(a)]
Bogor City			
Delphi CIMP	18.79%	18.96%	19.13%
Consumption Poverty	13.96%	14.09%	14.23%
Global MPI	15.47%	15.62%	15.77%
Adjusted MPI	19.28%	19.44%	19.60%
West Java			
Delphi CIMP	17.56%	17.58%	17.60%
Consumption Poverty	10.73%	10.75%	10.76%
Global MPI	20.47%	20.50%	20.52%
Adjusted MPI	29.58%	29.61%	29.63%
Indonesia			
Delphi CIMP	14.12%	14.13%	14.14%
Consumption Poverty	11.65%	11.66%	11.67%
Global MPI	21.40%	21.41%	21.42%
Adjusted MPI	29.03%	29.04%	29.05%

Note: [(a)] The 'lower' and 'upper bound' values are for the 95 percent confidence intervals.
Source: Author's computations

[1] As noted within previous footnotes, referring to Table 7.7 in the Appendix to this chapter, the Adjusted MPI's 'Education Continuity' indicator, with its highly stringent cutoff, identified close to three times more deprived households, compared to the Delphi-CIMP's 'Children's Education' indicator.
[2] Referring again to the Appendix of this chapter, 36 percent of households are deprived in this indicator.

These differences may again be due to the types of indicators included within these measures. As noted within sub-section 7.2.2, the Global MPI includes indicators, such as 'Quality of Housing' and 'Type of Cooking Fuel'. These indicators are excluded within the more urban specific Delphi-CIMP. Referring to Table 7.7 in the Appendix of this chapter, it can be seen that in Bogor City, those lacking in the quality of the 'roof, floor or wall', which comprise the house they live within, was found to be relatively low at 1.5 percent deprived households.[1]

In West Java and Indonesia, however, households lacking in the quality of their homes, amounted to more than ten times that found in Bogor City (at 16 and 19 percent respectively). As discussed within Chapter 5, although the issue of slum settlements is still a significant concern within Bogor City, the houses, which slum-residents occupy, are made of relatively sturdier materials (as an example, corrugated iron), compared to those specified to define a deprived household within the Global MPI. As for 'Cooking Fuel', only 3.5 percent of households in Bogor City are deprived in this indicator, compared to 21 percent in West Java and 33 percent in Indonesia.

In sum, differences in the size of the poor were found when the different measures were utilised. In particular, Consumption was shown to report the lowest number of poor households compared to the multidimensional measures, regardless of geographical level of analysis. This confirms the findings within literature and Chapter 2 of this thesis. Furthermore, it reaffirms the need to consider the use of multidimensional measures, alongside the Consumption measure, when assessing poverty. In addition to this, when examining the size of the poor across the three levels of analysis, differences were also present with regard to which multidimensional measure identified the most poverty. As noted above, the Delphi-CIMP was shown to identify more poverty compared to the Global MPI in Bogor City; this was not the case when data from West Java and Indonesia were considered.[2] Conclusions with regard to the relative position of these areas according to the proportion of the poor within them, would then differ, when the Delphi-CIMP is employed as opposed to the Global MPI. This in turn, may significantly affect policy, as within Indonesia specifically, the amount of funds disbursed by the central to district

[1] As described within Chapter 2, the Adjusted MPI also includes an 'Adequate Housing' indicator, however, employs a different cutoff compared to that within the Global MPI. The Adjusted MPI specifies that each household has to be deprived in all components; i.e. floor, wall and roof to be considered as deprived, whereas for the Global MPI, a household is deprived if it is lacking in any one of these components. The Adjusted MPI's cutoff leads to 0.2 percent of households in Bogor City being deprived.
[2] These conclusions with regard to the Delphi-CIMP are however, somewhat tentative, as noted at the start of sub-section 7.2, the size of the poor identified by this measure, is sensitive to the poverty cutoff chosen.

governments, depends on the relative positions of these districts with regard to their poverty levels (Brodjonegoro, 2002, 2003, 2004).[1]

7.3.2 Household Rank

This sub-section aims to answer the second research question of this chapter, i.e. whether household ranks differ when the different measures are used. As with the size of the poor, analysis of differences in household rank is conducted at three levels, district, provincial and national. This in mind, this sub-section also seeks to address Research Question (4).

The analysis within this sub-section follows the steps taken within sub-section 2.3.2 of Chapter 2. Firstly, for the Consumption measure, households were ranked according to their per capita daily consumption. Then, for the three multidimensional measures, a household's rank was determined by its total weighted deprivation. Considering the above steps, unlike for the size of the poor, a poverty cutoff was not needed in order to determine household ranks. Finally, Spearman's rank correlation coefficients, comparing these household ranks when the different measures are used, were calculated. Table 7.6 illustrates the results of this exercise. From these results, three main conclusions can be made.

Firstly, relatively modest levels of correlation (an average of 0.3) can be found between the ranks assigned to households if their levels of consumption per-capita were used to base these ranks, compared to that if their total weighted deprivation according to all three multidimensional measures, was used. Households, which are shown to exhibit high levels of consumption, may thus, not necessarily be those, which experience low levels of deprivation, and vice versa.[2] This confirms the findings within literature summarised in sub-section 7.2.1 of this chapter; in particular, within Hanandita and Tampubolon (2015), who highlight the low levels of correlation (an average of 0.1), which exist between levels of consumption and deprivation in other indicators of poverty, such as health and schooling.[3]

[1] When correlation between district ranks, as determined by the levels of poverty evident within them when the four measures are applied, were computed, an average correlation of 0.44 was found. This confirms that the use of the four different measures may lead to differences with regard to which district may be most eligible to receive the highest amount of central government transfers.
[2] Modest tetrachoric correlation between a household's poverty status when the four measures are used confirms this finding. Table 7.8 within the Appendix of this chapter presents these correlations.
[3] Section 5.4 within Chapter 5 of this thesis also found modest levels of correlation (an average of 0.38) between household deprivation status in consumption and deprivation within the other seventeen indicators included in the Delphi-CIMP.

Table 7. 6 Spearman Rank Correlation Coefficients between Household Rank as Determined by the Four Measures [(a)] [(b)]

	Bogor City				West Java				Indonesia			
	Delphi CIMP	Consumption	Global MPI	Adjusted MPI	Delphi CIMP	Consumption	Global MPI	Adjusted MPI	Delphi CIMP	Consumption	Global MPI	Adjusted MPI
Delphi CIMP	1				1				1			
Consumption	0.30	1			0.37	1			0.34	1		
Global MPI	0.37	0.28	1		0.46	0.32	1		0.47	0.35	1	
Adjusted MPI	0.62	0.42	0.38	1	0.65	0.42	0.4	1	0.66	0.43	0.50	1

Note:

[(a)] All coefficients are significant at 1% level.

[(b)] Ranks for the Consumption measure were assigned at descending order, so as to maintain a positive correlation with ranks according to total weighted deprivation, as determined by the multidimensional measures

Source: Author's computations

Secondly, the largest level of correlation can be seen between the Delphi CIMP and the Adjusted MPI. This may be due to the fact that both measures were developed for the context of Indonesia. The level of correlation between these two measures (at around 0.6) is, however, not perfect and still, relatively modest. Following that noted within the conclusion of sub-section 7.3.1, this result further strengthens the need for both national and sub-national measures when measuring multidimensional poverty in a nation as diverse as Indonesia.

Thirdly, to address Research Question (4), low to moderate levels of correlation between the four measures, were found, regardless of the geographical level of analysis.

The above in mind, subtle differences with regard to the extent of these correlations, however, do exist when data from the different levels are examined. To illustrate, a relatively higher correlation between the Delphi-CIMP and the two existing multidimensional measures is present when data from Indonesia and West Java are used, compared to that when data only from Bogor City is used. As the Global and Adjusted MPIs aim to be relevant at international and national levels, this result suggests that these measures could be less suitable when utilised to monitor poverty within a predominantly urban district, such as Bogor City. This again points to the likely need for multidimensional measures, which aim to capture specific district level characteristics.

To summarise, differences in how households are ranked, as indicated by the low to moderate levels of correlation found between the four measures, may lead to opposing conclusions with regard to which households may deserve social assistance, when the different measures are used. This is of particular relevance, since as noted within previous chapters, targeting policies, which aim to support the deliverance of help to those in need, make up a substantial part of poverty alleviation efforts within developing countries, including Indonesia.

7.3.3 Composition of the Poor

To further examine the conclusions found within the previous sections above, this sub-section aims to answer Research Question 3; i.e. whether the different measures identify different types of households as poor. To address this question, multivariate logistic regressions were conducted to analyse the effect of a list of socio-economic household characteristics on the probability of a household being categorised as poor according to the four measures.[1] Predicted

[1] Arguments behind the inclusion of these household characteristics are presented within Chapter 2.

probabilities were then calculated from these results. Figures 7.1 to 7.6 illustrate the trends within these predicted probabilities.[1]

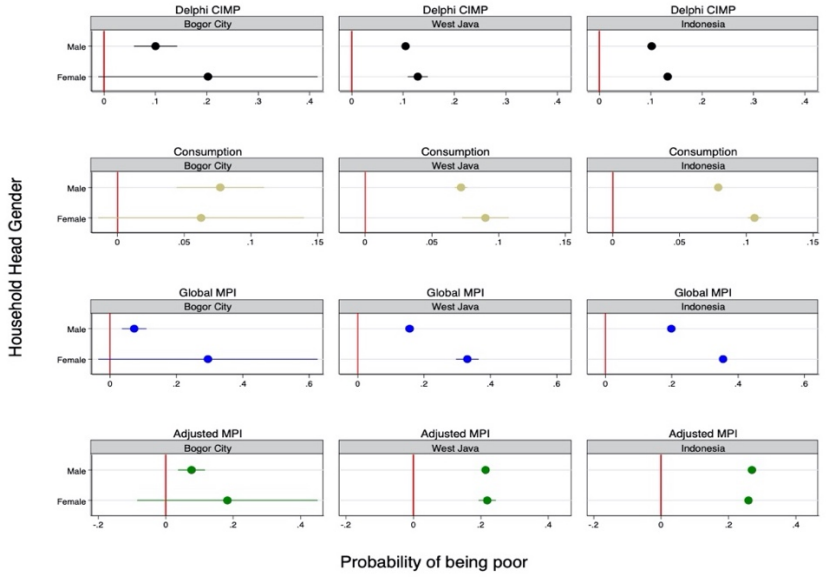
Before discussing the results of the above exercise, a note of clarification is, again, needed with regard to the poverty cutoff utilised for the Delphi-CIMP. As noted within sub-section 6.4.2.4 of Chapter 6, a similar composition of the poor was found for a relatively wide range of poverty cutoffs indicating a level of robustness with regard to the types of households identified as poor by the Delphi-CIMP, regardless of the poverty cutoff chosen. Unlike the size of the poor, the results within this sub-section are thus less sensitive to this cutoff level. This in mind, the arbitrary intermediate cutoff chosen for analysis within this chapter is again utilised within this sub-section, i.e. *a household is poor in the Delphi-CIMP measure if its standardised weighted total deprivation is equal to or above one standard deviation away from its mean*.

The above in mind, Figure 7.1 describes the relationship between the 'gender' of the household head and the probability of the said household being categorised as poor by the four measures, across the three geographical levels of analysis. Firstly, within Bogor City, although noting the wide confidence intervals, the results suggest that compared to the Consumption measure, the Delphi-CIMP, similar to the Global and Adjusted MPIs, identifies 'female' headed households as more likely to be poor. This finding confirms the discussion within sub-section 2.3.3.2 of Chapter 2, which notes that the Consumption measure in Bogor City, may miss the fact that female headed tend to be more vulnerable to shocks and thus, are more susceptible to poverty (Pritchett et al., 2000, p. 21).

As shown within Figure 7.1, the Delphi-CIMP, seems to reflect this 'vulnerability' in Bogor City. Indeed, going back to the types of indicators within the Delphi-CIMP, the inclusion of aspects, such as 'Child Mortality', 'Birth Attendant' and 'Contraception', aimed to better capture deprivation experienced by women in the household.[2] Examining across the three geographical levels, a different conclusion is, however, found, in that when data from West Java and Indonesia are considered, both the Consumption and Delphi-CIMP measures, identify 'female' headed households as more likely to be poor.[3]

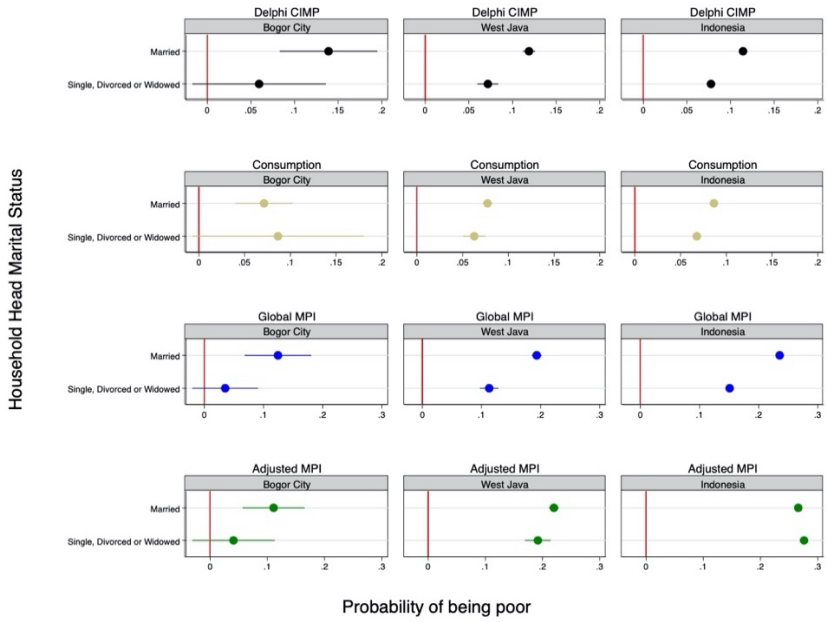
[1] Following Chapter 2, these predicted probabilities were calculated for each socio-economic determinant, whilst keeping other determinants at their mean values. Table 7.9 within the Appendix of this chapter presents the log odds resulting from these regressions.
[2] As noted within Chapter 5, these indicators only apply to women within the household.
[3] Looking at the confidence intervals in Figure 7.2, the effect of 'gender' becomes significant as the sample size increases when data from West Java and Indonesia are used.

Figure 7. 1 Predicted probabilities of household poverty status by gender of household head (with 95 percent confidence intervals)



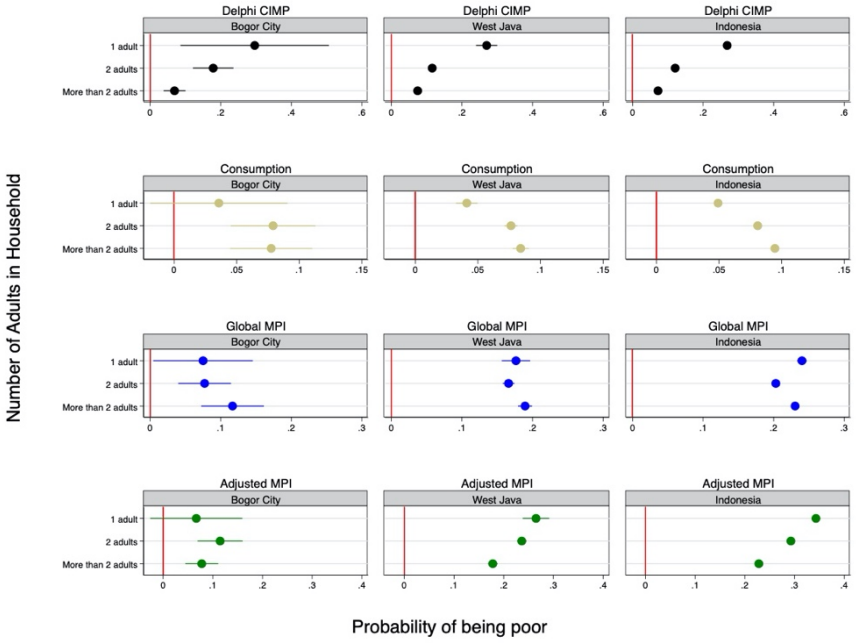
Source: Author's computations

Figure 7. 2 Predicted probabilities of household poverty status by marital status of household head (with 95 percent confidence intervals)



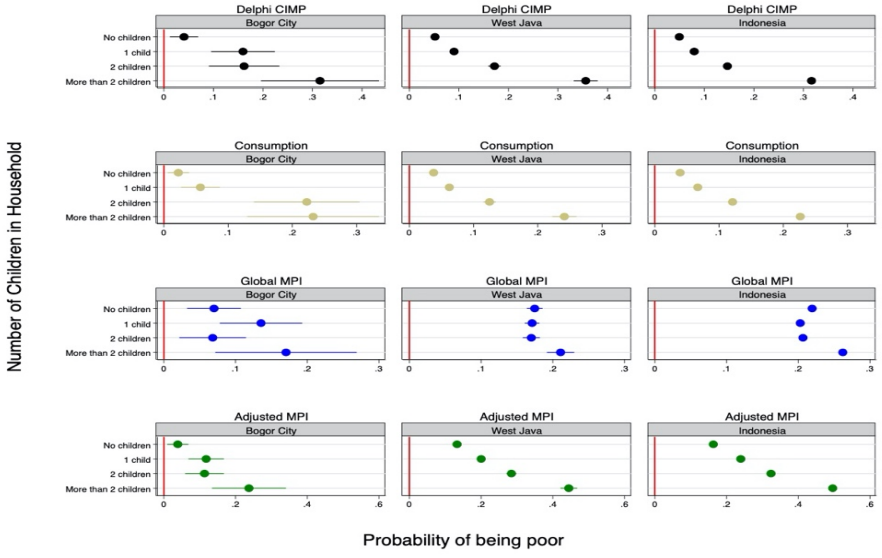
Source: Author's computations

Figure 7. 3 Predicted probabilities of household poverty status by number of adults in household (with 95 percent confidence intervals)



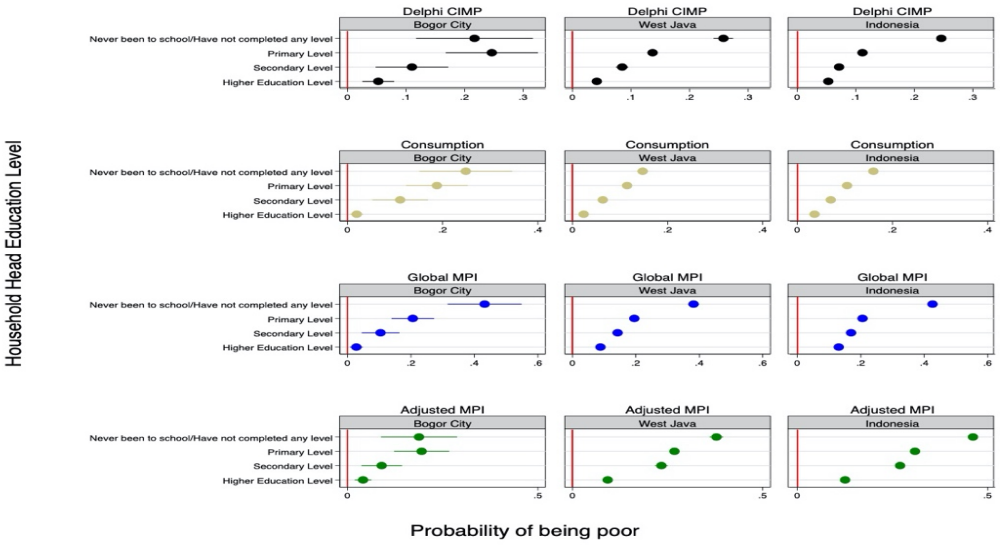
Source: Author's computations

Figure 7. 4 Predicted probabilities of household poverty status by number of children in household (with 95 percent confidence intervals)



Source: Author's computations

Figure 7. 5 Predicted probabilities of household poverty status by education level of the household head (with 95 percent confidence intervals)



Source: Author's computations

This may point toward the Delphi-CIMP being more able to capture 'female' vulnerability when compared to the Consumption measure, if it is applied within Bogor City. This ability is somewhat overridden when larger sample sizes from the higher geographical levels are used.

Figure 7.2 illustrates the relationship between household head's 'marital status' and the probability of being poor. As shown, in Bogor City, the Delphi-CIMP, like the Global and Adjusted MPIs, tends to identify 'married' headed households as more likely to be poor. Looking at the confidence intervals, however, the effects of 'marital status' are shown to be insignificant within this city. Given this result, following the discussion within Chapter 2, living within a 'single, divorced or widowed' headed household may decrease the probability of being poor, as improvements in the economic status of the household, alongside possible welfare and child support transfers, could have occurred (John C. Anyanwu, 2014). The Delphi-CIMP and the two existing multidimensional measures seem to better capture these possible improvements in Bogor City, compared to the Consumption measure, which identifies 'Married' headed households as having a higher probability of being poor.

Figure 7.3 reports the relationship between the probability of being poor, with the 'number of adults' within the household. The Delphi-CIMP, in contrast to the other measures, seems to identify households with 'more adults' as less likely to be poor. In particular, looking at the

confidence intervals, the ‘more than 2 adults in the household’ category is shown to be particularly significant in reducing the probability of experiencing Delphi-CIMP poverty, within Bogor City, West Java and Indonesia. Referring to the discussion within sub-section 2.3.3.1 of Chapter 2, the ‘number of adults’ within a household could indicate the number of ‘working individuals’ within it, which in turn, may play a role in reducing the probability of being poor (Dartanto & Nurkholis, 2013). Looking at the types of indicators included within the measure, the Delphi-CIMP’s ‘Employment’ indicator, may have captured this particular characteristic within the household.[1, 2]

Different from the above discussion where disparities were present, Figures 7.4 and 7.5 highlight similarities between the four measures. Figure 7.4 shows the impact of the ‘number of children’ in the household on the probability of being poor. As can be seen, having more children is shown to lead to higher probabilities of being poor, regardless of the measure and regardless of the geographical level analysed. As discussed in Chapter 2, the additional costs connected to childcare may be a driving factor behind this result. Another area of similarity between measures is found within Figure 7.5, i.e. increases in the household head’s educational level decreases the chances of the said household experiencing poverty. This is true for all four measures and regardless of the geographical level of analysis. In particular for the Delphi-CIMP the importance of ‘Education’ in Indonesia is reflected within the myriad education related indicators included in this measure. In addition, the stringency of the cutoffs utilised to define ‘minimum attainment’ within these indicators, also aims to capture the key role that education plays with regard to mitigating the effects of poverty.

In sum, as stated within Chapter 2, the differences in the composition of the poor, summarised above, indicate the need to consider both consumption and multidimensional measures when assessing poverty. Differences, which come to the fore when the different geographical levels are analysed, although somewhat subtle, also point toward the need to consider contextual, in particular, district level issues when measuring poverty in Indonesia. Of additional importance is that despite these differences, additional child costs were found to be universally significant as a characteristic that leads to more poverty, regardless of measure and geographical level. Increases in ‘education level’ were then shown to universally decrease the probabilities of being

[1] As discussed within sub-section 2.3.3.2 of Chapter 2, for the Adjusted MPI, increases in the ‘number of adults’ within the household also decreases the probability of experiencing poverty.
[2] As also noted within Chapter 2, another area of difference is that for all measures, except the Global MPI, an inverted-U relationship is, for the most part, evident when the impact of ‘age’ of household head on the probability of being poor, is examined.

poor for all measures, across the geographical levels. These two areas of similarity point to the relevance of policies which aim to increase the access and quality of education, in particular for children.

7.4 Conclusion

This chapter aimed to address the questions; (1) Does the Delphi-CIMP identify a different size of the poor when compared to the Consumption, Global and Adjusted MPI measures? (2) Do household ranks differ when different measures are utilised? (3) Does the Delphi CIMP identify a different composition of the poor? And (4) Do different measures need to be used when assessing poverty at different levels of analysis (district, provincial and national)? The findings with regard to these questions are summarised below.

In response to Research Question (1), results within this chapter indicate that differences exist with regard to the extent of poverty as identified by the Delphi-CIMP, when compared to the three existing measures examined. In Bogor City, the Delphi-CIMP identified more poverty compared to the Consumption measure, thus, strengthening arguments within literature on the existence of a 'miss-match' between Consumption and more direct measures of deprivation. Within Bogor City, the Delphi-CIMP also identified more poverty compared to the Global MPI indicating that the indicators and cutoffs within the Delphi-CIMP may better reflect policy priorities and areas of concern within this city. The Delphi-CIMP, however, identified less poverty, when compared to the Adjusted MPI. Looking at the types of indicators, stringency of their cutoffs and values of their weights, the highly stringent cutoff utilised by the Adjusted MPI to define deprivation within its education indicators, compared to those used within the Delphi-CIMP, may have contributed to this measure identifying more households as poor compared to the Delphi-CIMP.

For Research Question (2), household ranks are shown to be only modestly correlated when the four measures are used. A relatively higher correlation, however still at a relatively modest level (at around 0.6), was nevertheless, found between household ranks as determined by the Delphi CIMP and those determined by the Adjusted MPI. Despite the Delphi-CIMP seeking to be of relevance within one particular district, this relatively higher correlation, may be due to both measures aiming to include indicators that are relevant for Indonesia.

To address Research Question (3), a number of differences emerged with regard to the types of households identified as poor by the four measures. These differences were in particular present, when comparing the Delphi-CIMP with the Consumption measure using data from

Bogor City. In Bogor City, the Delphi-CIMP was shown to identify female headed households, where women are married, however, live within households with a lower number of adults, as having a higher probability of experiencing poverty. Whereas for the Consumption measure, poor households in Bogor City were more likely shown to be male headed, where the household head is 'single, divorced or widowed' and a higher number of adults are present within the household. Despite these differences, level of education and additional costs of having more children within the household, were shown to impact the probabilities of being poor, as identified by the four measures in similar ways, with 'Education' shown to decrease the chances of being poor and increases in the 'Number of Children' in the household, shown to increase the possibilities of experiencing poverty.

Finally, to answer Research Question (4), differences were found with regard to the above conclusions, when different geographical levels of analysis were considered. These differences were particularly evident when comparing the effects of using the four measures, with data from Bogor City (at district level) vs. that from higher geographical levels of analysis, i.e. West Java (at provincial) and Indonesia (at the national level). On the size of the poor, the Delphi-CIMP was shown to identify more poverty compared to the Global MPI in Bogor City, however, identified less when national and provincial data were utilised. Subtle differences also emerged when examining the composition of the poor across the geographical levels. Again, these differences were found to be particularly evident when comparing the Delphi-CIMP and the Consumption measure. To illustrate, the Delphi-CIMP tended to identify more 'female' headed households as poor compared to the Consumption measure, when Bogor City was examined. As the level of geographical analysis broadened to include the whole of West Java and Indonesia, however, both these measures were shown to identify 'female' headed households as more likely to experience poverty.

To conclude, the findings above strengthens the arguments within Chapter 2 of this thesis, i.e. that the use of Consumption for the sole measure of poverty, is not enough, as the multidimensional measures examined, were shown to not only identify a different size of the poor, however, also different household ranks and types of households as poor. In addition to this, in particular through examining the impact of using the district-level developed, Delphi-CIMP, contextual issues were found to be of importance when measuring multidimensional poverty. Following the long-established arguments presented within Nolan and Whelan (1996), poverty measurement should strive to reflect the definitions of poverty and the aspects of deprivation thought to be of relevance within the context in which measurement is needed.

Chapter 8: Conclusion

Poverty is a notoriously ill-defined term; there is remarkably little consensus on how best to measure the condition

(Nolan & Whelan, 1996, p. 1)

Poverty and inequality should no longer be safely subordinated to maximising total economic output alone

(Grusky & Kanbur, 2006)

Little is known about the present situation and achievements in relation to multiple deprivations in Indonesia and the empirical evidence is scanty

(Sumarto & de Silva, 2014, p. 2)

Indonesia is seriously lagging in non-income dimensions of poverty, such as health and education; if multidimensional poverty is defined as expenditure poverty and deprivation in at least one of these non-income dimensions, around half of all Indonesians are poor

(Milne, 2006, p. 31)

8.1 Background

One of the crucial factors that needs to be considered with regard to poverty alleviation, is the 'availability of accurate and reliable data on poverty measurements' (Central Bureau of Statistics Indonesia (CBS), 2014). There is unprecedented importance in setting a clear definition of poverty, which guides its measurement, before sound policy and action can be taken to address and eradicate it. Nolan and Whelan (1996), comment on the existing 'confusion associated with the term [poverty]', how definitions should go hand in hand with measurement and ask the poignant question, 'what is it that one is actually trying to measure?' (p. 1). This debate on how best to 'empirically measure' the phenomenon of poverty, still remains relevant today.

Currently the Indonesian government measures poverty using a uni-dimensional monetary poverty measure in which a person is considered poor if the sum of his or her food and basic needs consumption expenditure is below a certain determined level (Central Bureau of Statistics Indonesia (CBS), 2014; J. Priebe, 2014). Since its first inception in the 1980s, however, the legitimacy of the use of this consumption-based monetary poverty line as an adequate basis for measuring poverty, has been questioned. As Sajogyo (1977), founder of Indonesia's first consumption based poverty line, argues 'consumption numbers can only paint half the picture of the complexity of poverty, thus, when technology, knowledge and data permit, a more reliable measure, based on other dimensions of well-being, should be striven towards' (Tempo 1997). Sajogyo's warnings seem to echo the many debates within literature concerning the reliability or otherwise of uni–dimensional monetary-based measurements of poverty.

M. C. Nussbaum (2011) summarises these arguments by stating that, through the sole use of monetary-based measures, ‘for a long time, economists, policy makers and bureaucrats who work on the problems of the world's poorer nations, told people a story that distorted human experience’, as there is no guarantee that each and every person in a society is able to convert and enjoy his or her income as much as the next person.

This potential miss-match between uni-dimensional monetary measures, such as income and consumption, and measures based on multiple areas of poverty, has led to the proliferation of numerous theories and empirical applications, which aim to encompass the multi-faceted nature of poverty, applied both in the developed and developing worlds (see e.g. Amartya Sen (1976b), Amartya Sen (1979a), Streeten et al. (1981), Mayer and Jencks (1989), Tsui (2002), Bourguignon and Chakravarty (2003); Nolan and Whelan (1996), Whelan et al. (2014) and S. Alkire and Santos (2010a)). With regard to Indonesia, Sumarto and de Silva (2014) point to the fact that little is still known on the ‘achievements’ of Indonesia’s households, provinces and districts when multiple dimensions of poverty are considered; furthermore, insufficient empirical literature exists on how best to measure these multiple deprivations. This issue is of great concern as Milne (2006) notes, if multiple dimensions are utilised to measure poverty and poverty is defined as expenditure poverty and deprivation in at least one other dimension, around ‘half’ of Indonesia’s population will be considered poor.

The above in mind, the core issue within this thesis is, how to define and measure poverty in a way that is valid and meaningful in the area, in which measurement is needed, so as to enable the formation of sound and reliable policies. This thesis analyses this issue within the context of the developing world; using Indonesia, as a case study. In light of this, the following questions emerge: Is a measure of poverty based on consumption enough to describe the various forms of deprivations households may experience? If not, how would we go about measuring these multiple areas of deprivations? Upon the suggestion of an appropriate method, to what extent would this ‘new’ multidimensional measure identify different sizes and compositions of the poor, when compared to a measure based solely on consumption?

The above questions acted as the starting point of this thesis.

The following sections of this chapter are organised as follows. Section 8.2 aims to revisit the main objectives and key research questions this thesis sets out to address. Section 8.3 summarises this thesis’ main findings. Section 8.4 discusses the main contributions of this

thesis; both to literature and to guide possible policy applications. Section 8.5 notes suggestions for future research and Section 8.6 briefly concludes.

8.2 Research Goals Revisited

This thesis has the following two main objectives: (1) evaluating and comparing existing measures, both uni and multidimensional, before (2) suggesting a novel approach to the measurement of multidimensional poverty, which aims to mitigate the challenges faced within these existing measures. The overarching research question of this thesis is thus: to what extent does this 'new' measure of multidimensional poverty differ from the widely used consumption measure, within the context of a developing country? In other words, do we need a multidimensional measure? This thesis proposes to address this overarching question by examining differences with regard to how these measures identify the size and composition of the poor, alongside evaluations on how households are ranked based on their poverty level as defined by these measures. These are crucial issues to investigate, for, 'what we measure affects what we do; and if our measurements are flawed, decisions may be distorted' (Stiglitz, Sen, & Fitoussi, 2010a).

Grusky and Kanbur (2006) call upon the need for 'careful reflection' and 'conceptual ground clearing', before 'safely' moving from the relatively 'easy to interpret and understand' income and consumption-based measures, to an 'operational multidimentionalist' understanding of poverty (p. 2-3). This thesis aims to contribute toward mitigating the challenges faced when embarking on this paradigm shift. It does so by first, suggesting a way of defining and justifying a list of valued dimensions and indicators, which aim to 'capture' the phenomenon of poverty, before proposing a method of incorporating these multiple aspects into concrete measurement which recognises the need for a measure of multidimensional poverty to reflect the preferences, degrees of importance and trade-offs various social groups assign toward each dimension and indicator. Within this thesis, I suggest the use of the Delphi, as such a method.

The Delphi is broadly defined as 'a method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem' (Harold A. Linstone & Turoff, 1977, p. 3). The method 'involves the repeated individual questioning of the experts (by interview or questionnaire) and avoids direct confrontation of the experts with one another' (N. Dalkey & Helmer, 1963, p. 458). This use of iterative rounds and the presentation of 'group feedback' between rounds is argued to be 'more

conducive to independent thought' and leads to 'the gradual formation of opinions' (N. Dalkey & Helmer, 1963, p. 2).

Before implementing the Delphi to advance on the challenges outlined above, it was important to understand the conceptual bases and methodological assumptions made within existing measures of poverty, alongside how they were operationalised, particularly to assess poverty in Indonesia. As Townsend states, 'any attempt to justify a new approach towards the definition and measurement of poverty, must begin with previous definitions and evidence' (Townsend, 1979b, pp. 32-33). In sum, studying these past measures allowed me to identify areas in which advancements were needed.

The above in mind, five specific research questions were framed around the overarching research question. These questions are:

i. By replicating the Consumption, Global and Adjusted MPI measures: Are household ranks, the size and composition of the poor sensitive to these different measures?
ii. Through the use of the Delphi method: Can consensus be achieved with regard to which dimensions, indicators and respective weights, to include within a 'new' multidimensional poverty measure, which reflects preferences of key policy actors within a specific district level case study area in Indonesia?
iii. Through the use of the Delphi: Which measurement form, 'dashboard' or 'composite', was thought to best describe poverty within the case study area? If both were found to be relevant, before constructing a composite measure, what would the dashboard measure reveal with regard to deprivation within the case study area?
iv. Using results from the above suggested Delphi method and attempting to address the arbitrariness in which existing measures assign weights to dimensions and indicators: Does applying the 'expert' generated weights resulting from the Delphi within a composite measure of multidimensional poverty, identify different households as poor, when compared to the use of 'equal' weights?
v. Finally: Do the different dimensions, indicators and weights resulting from the Delphi, lead to differing conclusions with regard to identifying the poor, when compared to the Consumption, Global and Adjusted MPI measures?

Although a degree of attention can be found within current literature on the possible mismatch between monetary based measures versus a more multidimensional framework of understanding poverty within the developing world, to date, particularly in Indonesia, few have systematically compared the effects of applying these measures on the *types* of households they identify, alongside how these households are ranked. Even fewer studies can be found focusing on comparing a globally recognised multidimensional measure, such as the Global MPI, with a *nationally* developed multidimensional measure, i.e. the Adjusted MPI, which mirrors the global measure in almost every aspect *but* the number and kinds of indicators included, alongside the relatively participatory way in which these components were selected. The few pieces of research which do exist, have generally focused on comparing the Consumption measure, with a newly suggested multidimensional measure, which bases the selection of its components on 'data availability' and 'informed guesses' made by researchers. This, in turn, introduces a degree of arbitrariness within these newly suggested measures, thus rendering them to fall into the same critiques, which afflict globally recognised measures, such as the Global MPI and the Human Development Index (HDI). Furthermore, focus within these past studies on multidimensional poverty was for the most part given to how these measures identify different sizes of the poor; less so, on differences in composition.

More importantly, however, even though a wide range of methods can now be found with regard to how multidimensional measures may be constructed, particularly within the context of Indonesia, no published work exists on the use of the Delphi method in mitigating the challenges faced when attempting to achieve a level of consensus with regard to which dimensions and indicators to include within a measure of multidimensional poverty. Moreover, as this thesis illustrates, the Delphi method played a pivotal role in eliciting opinions on how these aspects should be *weighed*; aiming to address crucial concerns regarding how the multiple dimensions and indicators should be aggregated. As Norman C Dalkey and Rourke (1971) argue the Delphi is found to be a particularly favourable method 'for subject matters where the best available information is the *judgments* of knowledgeable individuals' (emphasis added, p. 1). Poverty is certainly such a topic. Following Nolan and Whelan (1996), 'the diversity of possible judgments should be taken into account in the measurement [of poverty] procedure; while this may permit less all-embracing conclusions, it does offer the prospect of unambiguous conclusions in certain circumstances; equally important, results that do depend on the choices made on the course of measurement can be recognized as such' (p. 2). I argue that the Delphi method significantly contributes to these goals.

In addition to the above, applying the Delphi method at the *district* level in Indonesia, also allowed the 'new' measure based on the method's results, to reflect the relative needs and standards thought to be of particular value within the said district. In addition to the 'sense of ownership' toward the 'new' Delphi-based measure this created, given the diversity of Indonesia as a nation and its decentralised nature, additional significance was found in uncovering whether, a nationally developed multidimensional measure would suffice, or rather, if district level measures are better suited as bases for policy.

The above in mind, this thesis has thus moved beyond the limitations of the Consumption measure, to exploring the merits of applying a multidimensional framework to analyse poverty within the developing world, whilst tackling what Thorbecke (2007) calls, the 'thorny problem', of the selection of 'attributes' within these quantitative multidimensional poverty measures. For, albeit the considerable progress that has been made with regard to identifying the related 'theoretical issues' in their application, many 'unresolved questions' continue to be evident with regard to how to *utilise* these measures in a 'truly operational sense' (Ibid, p. 18).

8.3 Key Findings

This section is organised so as to address the five research questions stated within Section 8.2; starting with results from the first exercise of examining the existing measures of poverty considered within this thesis.

8.3.1 Examining Existing Measures

In order to address the first research question listed within Section 8.2 above, Chapter 2 of this thesis provided a comparative overview of the use of a number of existing poverty measures currently employed to measure poverty in Indonesia. By 'existing measures', this thesis limits its analysis to three measures; i.e. (1) the Consumption poverty measure, which is the official standard used by the Government of Indonesia; (2) the widely applied Global MPI; and (3) a national multidimensional measure, developed by a local NGO, which I call the Adjusted MPI. Three levels of analysis are considered; i.e. national (Indonesia overall), provincial (West Java) and district levels (Bogor City). The household is then, considered as the unit of analysis. The results of this chapter may be summarised into the three main points below.

Firstly, the three measures identified significantly different extents of poverty at all three geographical levels of analysis. Using the Consumption measure as a benchmark, the Global MPI identified twice as many households as poor in Indonesia. The difference in the size of the

poor was even more apparent when the Adjusted MPI was employed; as it identified three times more poverty in Indonesia, when compared to the Consumption measure. This amounts to some additional 9.5 million households in Indonesia identified as poor by this measure, compared to the Consumption measure. These results somewhat go hand in hand with the argument made by Milne (2006), quoted at the start on this chapter, on the extent and importance of multidimensional poverty in Indonesia.

Secondly, household rank was examined. Within this exercise, for the Consumption measure, households are ranked according to their level of monthly consumption per capita. For the multidimensional measures, households are ranked according to their total weighted deprivation in all the dimensions and indicators included within these measures. Rank correlation coefficients between these 'poverty scores' were then calculated. Moderate levels of correlation; with an average of 0.37, were found between the Consumption measure and the multidimensional measures, across the three levels of geographical analysis. These modest correlations suggest that household ranks would very likely shift somewhat substantially when a different concept of poverty, which takes into account multidimensional aspects of deprivation is applied, as opposed to utilising Consumption as the single yardstick to measure poverty. In comparing the two existing multidimensional measures, i.e. the internationally applied Global MPI, with the nationally generated Adjusted MPI, a relatively higher average correlation of 0.42 was found. This level of correlation, however, remains modest and far from indicating complete equivalence between how these measures rank households. This result somewhat supports arguments made on the need for locally developed multidimensional measures, thus further strengthening the necessity for the 'new' Delphi-based measure proposed within this thesis, particularly within a nation as diverse as Indonesia.

Thirdly, household composition was studied. Analysing differences in the characteristics of households categorised as poor according to the three measures was crucial as a way of understanding the possible reasons behind the differences outlined within the first and second findings above. In addition to this, differences in composition may act as an indication of the extent to which one measure, if employed alone, is likely to mislead policy on the types of households needing social assistance. With this in mind, predicted probabilities were calculated from multinomial regression models, which included key household characteristics found to be relevant within literature, particularly for the context of Indonesia. Considering this, a number of differences were found with regard to the composition of the poor, when the three measures were employed.

The following differing characteristics were found when examining certain types of household headships. When examining data from Bogor City, with regard to the ‘gender’ of the household head, ‘female’ headed households are shown to be more likely to experience poverty for the Global and Adjusted MPI measures. This is not evident for the Consumption measure, where ‘male’ headed households exhibit a higher likelihood of being poor in Bogor City. When household head ‘marital status’ was examined in Bogor City, being ‘single, divorced or widowed’, contrary to established empirical and theoretical arguments, is shown to decrease the probability of being Global and Adjusted MPI poor. As argued for within Chapter 2, after adjusting for household size, this may be due to an improvement in the economic status of divorced or separated men, whilst, for women, moving back to live with parents, alongside possible child support transfers, may reduce the probability of experiencing poverty. Indeed, when household size is examined in Bogor City, for the Global MPI, living within smaller households was shown to reduce the probability of being poor. For Consumption, being ‘single, divorced or widowed’, however, increases the probability of being poor in Bogor City. In addition to this, when looking at the pattern and extent of differences in the predicted probabilities of being poor, disparities were also found between the Global and Adjusted MPIs, particularly when the sample was expanded to include data from the whole of Indonesia. This further reinforced the need for a ‘new’ locally specific multidimensional measure, which aims to address the level of arbitrariness, particularly with regard to the choice of weights, found within these existing multidimensional measures.

Despite the above differences, similarities between all three measures were also found. For all measures, increases in the level of education of the household head was shown to decrease the probability of experiencing poverty. Furthermore, households facing additional responsibilities toward childcare were also found as more likely to experience poverty, regardless of the geographical level of analysis.

Given the above results, this chapter concluded that reliance on any one of the existing measures examined as a single standard to guide policy and assess policy effectiveness may be misleading, for each measure not only identified a different extent of poverty, but also, despite some similarities found, categorised different types of households as poor. These differences thus, illustrated the importance of employing both uni and multidimensional measures to inform policy. Furthermore, as the need for a multidimensional approach was established, the necessity for advancements toward improving how multidimensional poverty is measured, became additionally apparent.

8.3.2 Dimensions, Indicators and Weights from a Delphi Exercise

Studying the existing multidimensional poverty measures within Chapter 2, a number of shortcomings were found, particularly related to the arbitrariness in the way the measures' components, i.e. dimensions, indicators and weights, were selected. Although the availability of data may partly drive these decisions, transparent discussion and satisfactory justification stemming from both theoretical and sound empirical work, which act as background to these methodological choices, are crucially needed. This however, posed a challenge, as Grusky et al. (2006) state, 'consensus has not been reached on *which* dimensions [of poverty] matter, nor even on *how* to decide what matters' (emphasis added, p. 12). Chapters 3 and 4 of this thesis, were located within literature targeted to advance progress in addressing this issue; thus, seeking to address the second research question noted within Section 8.2 of this chapter. In particular, Chapter 3 outlined how the Delphi method may act as a robust way of gathering opinions and achieving consensus on the dimensions, indicators and weights to include within a context specific multidimensional measure; whilst, Chapter 4 described the results of this Delphi study, which can be summarised into the two main points below.

The Delphi within this thesis was conducted within a specific district level case study area and included key policy makers within its process. The first step of the Delphi prompted participants to state a relevant definition of poverty, which should guide their choices throughout the exercise. This was followed by an exercise aimed at clustering indicators into dimensions, before finally introducing three iterative questionnaire rounds, which required participants to select relevant dimensions and indicators to include within a measure of multidimensional poverty. In addition, these iterative rounds were particularly important as they also allowed the collection of data on weights. These weights aimed to reflect the levels of importance and possible trade-offs participants assigned to chosen dimensions and indicators; thus, aiming to specifically address the arbitrariness of the 'equal nested weights' utilised within the Global and Adjusted MPIs. The results of this Delphi exercise showed that a degree of consensus was achieved with regard to these choices.

Firstly, with regard to how poverty should be defined, participants unanimously agreed upon the need for a multidimensional framework to understand and measure poverty. They also acknowledged the importance of the Capability Approach, i.e. that poverty measurement should be based on the valued freedoms and capabilities individuals are able to enjoy. Furthermore, analysing questionnaire responses throughout the three iterative Delphi rounds, a convergence of opinion was observed, indicating a level of consensus among key policy actors included

within the study, with regard to a specific list of dimensions and indicators. This list consisted of five dimensions; i.e. (1) Education, (2) Employment and Assets, (3) Living Standards, (4) Health and Safety and (5) Child Health and Contraception; and eighteen indicators to proxy these dimensions.

Secondly, on the crucial issue of weights, differences in opinion continued to be present on the *exact value* of these weights. Participants, however, agreed upon a range of values for these weights. Robustness checks conducted following the Delphi exercise, confirmed the above choices of dimensions and indicators, alongside the relevance of this range of weights. Furthermore, post-Delphi robustness check interviews highlighted the need for district measures, as national level measures may not reflect the specific characteristics and customs that are relevant for their community.

8.3.3 Applying Delphi Results into a Dashboard Measure of Poverty

When seeking to aggregate multidimensional poverty two directions are broadly followed; i.e. presenting the multiple dimensions and indicators in dashboard or composite form. In response to research question three within Section 8.2, this thesis argued for the importance of both approaches; Chapter 5 presented results of a dashboard measure, before results of the composite measure are presented within Chapter 6. This sub-section summarises the results of Chapter 5. For Chapter 5, focus was given to analysing data from the case study area chosen, i.e. Bogor City.

The dashboard in Chapter 5 enabled detailed examination of the extent of deprivation within each of the eighteen indicators considered as important by Delphi participants. This, in turn, provided evidence as to which indicators exhibited the highest levels of deprivation, thus, pointing toward areas of well-being which may benefit from urgent action. Monitoring poverty through an eighteen-indicator-dashboard, however, could prove to be a cumbersome exercise. This in mind, deprivation within the five dimensions were also analysed.

Before determining these ‘dimension deprivation levels’, decisions however, needed to be made with regard to specific dimensional cutoffs which define deprivation within each dimension, i.e. the number of indicators within each dimension a household needs to be deprived in, in order to be considered as deprived within the said dimension. On this issue, the ‘union’ approach to defining dimensional cutoffs, identified the largest levels of deprivation within all dimensions, whilst the ‘intersection’ approach, identified the lowest levels for all the five dimensions examined. Considering these extreme levels of deprivation as identified by these two well-

recognised approaches, the impact of a range of ‘intermediate’ cutoffs were also tested. Results of this exercise found that levels of deprivation within dimensions, were not robust to the choice of ‘intermediate’ cutoff applied.

Proponents of dashboard measures argue for their relevance, as among other things, within the construction of this form of measurement, weights are not required, thus, arguably decreasing the amount of arbitrary judgment needed in order to construct them. Chapter 5, however found that despite this being the case, dashboards are far from free of value judgments, as levels of deprivation within dimensions, were shown to be highly sensitive to the dimensional cutoff chosen. Furthermore, although presenting crucial detailed information on deprivation levels across the five dimensions and eighteen indicators selected by Delphi participants, the dashboard did not allow the ability to construct ‘complete comparisons’ of poverty over time and space. These findings strengthened the need for a composite index.

8.3.4 Applying Delphi Results into a Composite Measure of Poverty

Within Chapter 6, a composite index was constructed to supplement the dashboard. Two challenges were faced when attempting to incorporate the range of weights generated from the Delphi process within this composite measure; i.e. (1) that the sum of the weights should equal one and (2) that the degree of disagreement with regard to the exact value of these weights should be incorporated within the final measure. An optimisation exercise based on the Data Envelopment Analysis (DEA) and benefit-of-the-doubt approach was suggested in order to address these challenges. Utilising this approach ensured that the weights generated; dubbed as ‘optimal expert’ weights, reflected implicit policy preferences, degrees of importance placed on each dimension and indicator, alongside possible trade-offs between them.

To test the impact of the use of these ‘optimal expert’ weights on the extent and composition of the poor, the use of these weights within the final Delphi-generated measure, were compared against the commonly applied ‘equal nested’ weighting scheme. Similar to Chapter 5, data for Bogor City is utilised within the analysis of this chapter. This in mind, this comparative exercise aimed to address the fourth research question listed within Section 8.2. Results of this exercise are summarised within the four main points below.

Firstly, the values of ‘optimal expert’ weights, illustrated a hierarchy of importance between each dimension and indicator. Larger weights were assigned to components thought to be of relatively higher importance to poverty alleviation efforts in the case study area. This hierarchy was not evident within the ‘equal nested’ weighting scheme.

Secondly, despite the above, when examining household rank, both weighting schemes seem to order households in a relatively similar way. For this step within the analysis, households were assigned ranks according to their total weighted deprivation using the two weighting schemes. A high correlation coefficient (0.9) was found between these ranks, indicating that households considered as poorest when the 'optimal expert' scheme was utilised, were likely the same households as when the 'equal nested' weights were used. If importance is only placed on how households are ranked, the simpler to apply, 'equal nested' weights, thus, may suffice. Despite being high, however, these correlations were imperfect, thus, some level of miss-match may still occur when 'equal nested' weights are used as opposed to 'optimal expert' weights.

Thirdly, despite similarities in how households were ranked, for a range of poverty cutoffs, the 'optimal expert' weighting scheme identified significantly more households as poor. Using the 'optimal expert' weighting scheme as a benchmark, the use of the 'equal nested' weights, thus tended to underestimate the level of poverty within the case study area. This may in turn lead to significant policy drawbacks, as the use of 'equal nested' weights, could likely increase the number of poor households, which are overlooked by social assistance programmes.

Finally, when examining the composition of the poor, for a range of poverty cutoffs, no significant differences were found between the use of the two weighting schemes; i.e. both weighting schemes tended to identify households with relatively similar characteristics.

In sum, results from this chapter indicate the possibility that, dependant on the type of policy intended and considering the limited resources that governments frequently face, the use of equal weights may suffice within the construction of a composite measure of multidimensional poverty. Despite this, particularly bearing in mind the differences in sizes of the poor identified by these two weighting schemes, the use of one weighting scheme as opposed to the other, could still lead to significant differences with regard to policy formation.

In addition to this, the 'optimal expert' scheme offers a level of superiority by basing itself on empirical results. In other words, the participatory process in which these weights were arrived upon, contributed to a level of acceptance of this weighting scheme. To illustrate, considerable discussion and debate arose within the local government on the impact of the Delphi study on policy. These debates were also covered within local press reports, thus, sparking additional discussion and transparency of the Delphi's results. More importantly, however, the results of the dashboard and composite measures analysed within Chapters 5 and 6, led to concrete

evidence-based policy action, still in practice today (Bogor City Government, 2016; Republic of Indonesia Coordinating Ministry of Administrative and Bureaucratic Reform, 2016).

8.3.5 Comparing the Delphi Generated Measure with Existing Measures

Chapter 7 acted as the penultimate step in proposing the Delphi method as a way of mitigating the arbitrariness still evident within the way existing measures of multidimensional poverty select the dimensions, indicators and weights to include within measurement. The goal of this chapter was to illustrate the implications of utilising the Delphi method within the measurement of poverty; thus, aiming to address the final research question listed within Section 8.2. Focus on this chapter was given to comparing the Delphi-generated composite measure, in particular the measure which applied 'optimal expert' weights, which I dub the Delphi Composite Index of Multidimensional Poverty (Delphi-CIMP), with the Consumption, Global and Adjusted MPI measures.

By conducting this comparative exercise, this chapter aimed to fill an existing gap within current literature on how methodological choices and contextual factors may influence the way multidimensional poverty is measured and recorded. Analysis within this chapter was conducted using national, provincial and district level data, in order to further examine whether district specific measures may be needed to assess poverty in Indonesia. Results of this exercise are summarised within the three main findings below.

Firstly, following the steps of the comparative exercises conducted within previous chapters, household ranks are examined. For the multidimensional measures, similar to previous exercises, household rank was determined by their total weighted deprivation. For the Consumption measure, it was determined by the household's per capita monthly consumption. Results showed that ranks resulting from the use of the Delphi-CIMP, were significantly different from those resulting from applying the three existing measures.

The lowest level of correlation (at an average of 0.3 across the three geographical levels of analysis) was found between Delphi-CIMP ranks and Consumption measure ranks. A slightly higher level of correlation was found (an average of around 0.6 across the three levels of analysis), however, between the Delphi-CIMP and Adjusted MPI ranks. This was expected, as the Adjusted MPI, aimed to reflect national level needs, which hoped to encompass the views of all Indonesia's provinces and districts. Correlation between these two measures was, however, still moderate, indicating the need to supplement a national measure with more localised measures, which reflect the specific customs and needs of Indonesia's districts.

Secondly, examining the size of the poor proved to be a more cumbersome exercise, as Delphi participants were not willing to agree upon a specific poverty cutoff, to act as a threshold defining the level of deprivation a household had to experience in order to be categorised as poor. Keeping this in mind, robustness checks were conducted to assess the impact of this choice of poverty cutoff on how provinces and districts were ranked if the Delphi-CIMP was utilised to measure poverty within these areas. These ranks were found to be highly correlated (with an average correlation of 0.99) across a range of poverty cutoffs. The decision to select an arbitrary cutoff, in order to facilitate complete comparisons between measures, was thus supported. This cutoff was selected so as to result in an intermediate level of poverty; i.e. approximately at the midpoint between the cutoff, which generates the number of poor when the union and intersection approaches were applied.

Using this cutoff, the Delphi-CIMP is shown to identify more poverty In Bogor City, when compared to the Consumption and Global MPI, however, less when compared to the Adjusted MPI. When data from West Java and Indonesia were utilised, however, the Delphi-CIMP was shown to identify more poverty compared to the Consumption measure, however, less, when compared to both the Global and Adjusted MPIs.

The above finding confirms previous conclusions on the tendency of multidimensional measures to identify more poverty compared to the Consumption measure. Policy wise, suggesting a ‘new’ measure, which results in higher poverty levels compared to that of the Consumption measure, which acts as the official measure of poverty in Indonesia, however, carries particularly sensitive implications, specifically toward public perception on the reliability and effectivity of current policy. In this case, the transparency in the way opinions were analysed within the Delphi-CIMP, seeks to reassure potential users of the degree of reliability of the findings resulting from applying this measure within the case study area.

When comparing the multidimensional measures, differences in the size of the poor identified by these measures, may likely be the result of differences in the stringency of the indicator cutoffs, employed within these measures. On this issue, for all three multidimensional measures, indicator cutoffs were defined in a relatively similar way, except, in particular, for the cutoff for education indicators; with the Global MPI setting the least stringent, followed by the Delphi-CIMP and the Adjusted MPI applying the most stringent cutoffs to these indicators. These differences in cutoffs led to significantly different levels of deprivation within these education indicators; with the Adjusted MPI’s indicator cutoff related to ‘Education Continuity’ identifying 71 percent deprived households in Bogor City, the Global MPI’s ‘School Attendance’ indicator

identifying 3 percent deprived households and the Delphi-CIMP's 'Children's Education' indicator identifying 24 percent deprived households in Bogor City.

Finally, with regard to the composition of the poor, a number of different characteristics, particularly related to the impact of 'gender' and 'marital status' on the probability of being poor across the four measures, were found to be evident. What was further of interest was that these areas of difference between the measures somewhat changed when different levels of analysis; i.e. national, provincial and district levels, were considered. This further strengthened the need for context specific measures, particularly at district level given Indonesia's diversity and its decentralised nature.

8.4 Contributions towards Academia and Policy

8.4.1 The use of the Delphi in Multidimensional Poverty Measurement

Although the Delphi method has been applied within studies concerned with the selection of factors affecting the quality of life, to the author's knowledge, this was its first application within research concerned with selecting components of a multidimensional poverty measure. Its incorporation as an integral part of developing a poverty measure, presents a novel alternative to the use of other methods commonly used to gather the opinions of experts, such as Focus Group Discussions (FGDs). The level of anonymity the Delphi offers also served to ensure freedom of discussion within participants and mitigate the possible biases resulting from dominant voices. In addition, the Delphi offered a way of systematically recording not only 'cumulative outcomes', i.e. the final choice of dimensions, indicators and weights, but also, importantly, enabled the documenting of 'comprehensive outcomes', i.e. how these aspects of poverty measurement were chosen, who were included within this choice, how opinions were analysed and how consensus with regard to these choices was determined (Amartya Sen, 2004b).

8.4.2 Expert vs. Equal Weights

The issue of weighting within multidimensional indices of poverty has been a source of longstanding debate in academic research. The use of 'equal' weights, although evident within a substantial number of existing indices of welfare, including the widely applied Global MPI and Human Development Index (HDI), has been a source of major criticism against the relevance of these measures. Advancements on this issue is crucial as these measures are often utilised as benchmarks, against which countries often compare their welfare achievements. Although

results from this thesis can in no way be generalised, the similarity in the way households were ranked when Delphi-generated ‘expert-based’ weights vs. ‘equal nested’ weights were applied, may indicate the possible relevance of utilising equal weights within measurement, particularly as policy makers often face limited resources when it comes to implementing policy. Nevertheless, the differences evident specifically with regard to the extent of poverty identified by these two different weighting schemes, coupled with the level of reliability offered through basing measurement on empirical evidence on weights, served to strengthen the need for the ‘expert’ generated weights. This was even more so, as their use established a sense of ownership of the measure, which in turn prompted concrete action with regard to the formation of policies aimed at mitigating the debilitating effects of poverty within the case study area.

8.4.3 International vs. National vs. Localised Measures

Academic literature has given relatively little attention to comparing localised measures of multidimensional poverty versus globally formulated ones. This is of importance, as shown within this thesis, the relevance of a measure depends on whether it takes into account contextual issues when measuring poverty. Global measures, which aim to be of use for international comparisons of poverty across nations, may thus not be necessarily applicable when aiming to assess poverty at a more localised setting. This is even more the case, as illustrated by this thesis, within a diverse setting, such as Indonesia. As shown, even nationally developed measures may miss the mark and exclude indicators, which are specifically relevant for each district. This in mind, global measures continue to be of use in order to provide a means of comparing performance between different nations, however, local measures should be developed when attempting to assess more localised dimensions of multidimensional poverty, as complements to these global measures.

8.4.4 Policy Implications

The results of the analysis within this thesis have clear policy implications, particularly for the Government of Indonesia and its district governments. Findings from the comparative exercises conducted and the views expressed within the Delphi exercise confirmed the need for a multidimensional measure. Shifting focus on not only the Consumption poor, but also on those who are poor across a broader understanding of poverty which incorporates myriad dimensions and indicators of wellbeing, may be challenging for governments within the developing world. This thesis, however, shows that this can be done. Particularly for Indonesia, existing nationally representative survey data, such as the Susenas used within this thesis, may act as a starting

point, before additional datasets are collected to include further possible indicators of poverty. In illustrating the use of the Delphi, it was also shown that although far from suggesting a perfect method, eliciting opinions on components of a multidimensional measure (dimensions, indicators and weights), in a systematic and transparent way, was possible.

The Delphi conducted within this thesis was the first-time key policy makers within the case study area, gathered together, to discuss the issue of poverty measurement and how each of their sectors may contribute to its alleviation. This Delphi study generated substantial discourse on issues surrounding the importance of a multidimensional framework, both within the government and civil society.

8.5 Limitations and Suggestions for Future Research

8.5.1 Scope of the Study

The major limitation of this research is its scope. The Delphi exercise was only conducted within one district. More importantly, however, participants within the Delphi only included policy makers. A significant area where improvement could be made, is thus, on the issue of expanding the study to include not only additional districts, to assess whether differences in opinion would indeed occur if different settings were considered, however, more crucially, whether the inclusion of additional actors, particularly those experiencing poverty themselves, may alter the results found within this thesis. Conducting additional Delphi surveys at different points in time, whilst employing the same participants, could also provide clarity as to whether time specific contextual issues may influence results.

As described within Chapters 3 and 4, policy makers included within the Delphi found selecting a specific poverty cutoff a highly sensitive issue; expanding the pool of participants to include the civil society, NGOs, academics and people experiencing poverty, may lead to a different conclusion with regard to this issue. If not a particular cutoff, agreement may be achieved on a range of poverty cutoffs. This range of cutoffs could then possibly be applied to what Chiaperro-Martinetti (1994) dub ‘membership functions’ and the use of ‘fuzzy set’ theory in the measurement of poverty. Additional Delphi surveys could also be conducted to test whether there exists what Qizilbash (2003) calls a ‘core dimension’, which will enable the application of a ‘supervaluationalist’ approach to poverty measurement (p. 41).

In addition to the above, the importance Delphi participants placed on the ability to analyse the distribution of total weighted deprivations across households, could enable further analysis on

inequality levels within a multidimensional space; another crucial issue within the developing world.

Also touching upon the scope of the study, limited time, resources and the notable challenge of participant fatigue, necessitated the use of a pre-existing list of indicators available within the Susenas, alongside pre-existing literature on laws and regulations to base the selection of deprivation thresholds for these indicators. As noted within Chapter 4, a small number of Delphi participants (five out of a total of thirty-six participants) reported the need for 'freedom of religion' and 'access to recreational areas' to be included as possible important indicators of poverty. It would thus be beneficial to further assess whether more 'missing indicators' exist if the participant pool is expanded. Furthermore, additional questionnaires on the adequacy of existing internationally and nationally accepted minimum thresholds for indicators, may also add to literature and serve to inform policy on the relevance of these thresholds, particularly within different settings.

Finally, the Delphi exercise conducted within this thesis, although limited in scope, generated a substantial amount of data, the analysis of all however, was beyond the scope of this thesis. Examining how individuals' views changed across the iterative Delphi rounds; whether gender, levels of experience or educational attainment, played a role in the likeliness of a participant's probability of changing his or her opinion to converge toward the opinion of the group, could be a fascinating area of further research.

8.5.2 Additional Robustness Checks

In addition to the above limitations, the construction of the Delphi composite measure; i.e. dubbed the Delphi-CIMP within this thesis, utilised similar normalisation and aggregation methods as those employed within the Global MPI. Particularly with regard to the aggregation function used within the Delphi-CIMP, as noted within Chapter 6, the use of 'linear aggregation' implies a number of questionable assumptions; specifically, with regard to the level of 'elasticity of substitution' that is assumed between each of the indicators included in the measure. Within 'linear aggregation', elasticity of substitution is assumed to be infinite, thus, achievements within indicators are assumed to be perfect substitutes. In addition to this, 'linear aggregation' assumes a fixed rate at which transformed achievements can be exchanged. As Bourguignon and Chakravarty (2003) argue, 'these may not be very satisfactory assumptions'; as indicators may not be perfect substitutes. Furthermore, 'the substitutability between indicators may change with the extent of deprivation within each of these attributes' (Ibid, p. 40-41). Analysis of more

complex aggregation functions, alongside whether the choice to utilise different transformation functions may lead to differences in the identification of poverty, could thus act as an additionally important area for further research

8.6 Brief Concluding Thoughts

This thesis showed that how one defines and measures poverty, significantly affects poverty outcomes. For a nation as diverse as Indonesia, with a close to 270 million population, spread across around 6,000 inhabited islands, 34 provinces and 514 districts, recognising that utilising Consumption poverty, as the single yardstick to assess progress in welfare levels, may not be enough, is crucial. To follow the nation's decentralised nature, acknowledging the need for additional multidimensional measures, which reflects the diversity of the nation, is also of paramount importance. Policy wise, results from this thesis, strengthens the need to measure multidimensional poverty and for a decentralised approach to poverty alleviation, so as to make way for locally initiated programmes that seeks to address specific areas of deprivation relevant within each district.

Appendix to Chapter 3

Table 3. 2 List of Delphi Participants' Government Sectors and Divisions

Bogor City Government Sectors and Divisions	Number of Delphi Participants
The Secretariat for the Administration of Social Welfare Policy	2
The Secretariat for the Administration of Economic Development	1
The Sub-National Development Planning Agency	1
Sector for the Development of Social Welfare	1
Body for Community Development and Family Planning	2
Body for the Development of Social Welfare	1
Body for City Health	1
Body for the Empowerment of Community Health (a Subsidiary body of the Body of City Health)	1
Office of Building and Settlement Supervision	1
Sub-office of the Promotion of Adequate Housing (Sub-division of the Office of Building and Settlement Supervision)	1
Body for City Education	1
Sub-division of Elementary Education (Body for City Education)	1
Sub-division of Secondary Education (Body for City Education)	1
Sub-division of Higher Education (Body for City Education)	1
Body for Population and Demography	1
Body for Agriculture	1
Office for Food Security	1
Sub-office for Diversification of Food and Food Security	1
Body for Community Service	1
Body for Cooperatives	1
The City Government Agency for Labour, Employment and Transmigration	1
Office for the Promotion of Small Medium Enterprises	1
Sub-office of the Promotion of Small Medium Enterprises	1
Office for the Development and Promotion of Social, Cultural and Governmental Activities	1
Office for the Development and Promotion of Local Enterprises	2
Office for Tourism and Creative Economy	1
Sub-division Office of Creative Economy (Office for Tourism and Creative Economy)	1
Bogor City's Main Micro Lending Institution	1
People's Representative of Bogor City	1
Representative Head of the Sub-district of West Bogor	1
Representative Head of the Sub-district of Central Bogor	1
Representative Head of the Sub-district of Eastern Bogor	1
Representative Head of the Sub-district of South Bogor	1
Total	36

Source: Author’s compilations (Bogor City, 2016)

Table 3. 3 List of Robustness Check Interviewees' Government Sectors and Divisions

Bogor City Government Sectors and Divisions	Number of Interviewees (Delphi participants)	Number of Interviewees (non-Delphi participants)
The Secretariat for the Administration of Social Welfare Policy	2	2
The Secretariat for the Administration of Economic Development	1	1
The Sub-National Development Planning Agency	2	2
Sector for the Development of Social Welfare	1	1
Body for Community Development and Family Planning	1	1
Body for the Development of Social Welfare	2	2
Body for Community Development and Family Planning	1	1
Body for City Health	0	1
Body for the Empowerment of Community Health	1	0
Office of Building and Settlement Supervision	0	0
Sub-office of the Promotion of Adequate Housing	0	0
Body for City Education	2	2
Sub-division of Elementary Education (Body for City Education)	0	0
Sub-division of Secondary Education (Body for City Education)	0	0
Sub-division of Higher Education (Body for City Education)	0	0
Body for Population and Demography	0	0
Body for Agriculture	0	0
Office for Food Security	0	0
Sub-office for Diversification of Food and Food Security	0	0
Body for Community Service	1	1
Body for Cooperatives	0	0
The City Government Agency for Labour, Employment and Transmigration	0	0
Office for the Promotion of Small Medium Enterprises	1	1
Sub-office of the Promotion of Small Medium Enterprises	0	0
Office for the Development and Promotion of Social, Cultural and Governmental Activities	0	0
Office for the Development and Promotion of Local Enterprises	0	0
Office for Tourism and Creative Economy	0	0
Sub-division Office of Creative Economy	0	0
Bogor City's Main Micro Lending Institution	0	0
People's Representative of Bogor City	0	0
Representative Head of the Sub-district of West Bogor	0	0
Representative Head of the Sub-district of Central Bogor	0	0
Representative Head of the Sub-district of Eastern Bogor	0	0
Representative Head of the Sub-district of South Bogor	0	0
Total	15	15

Source: Author's compilations (Bogor City, 2016)

Appendix to Chapter 4

Table 4. 10 Poverty Definitions, Keywords, Theoretical Categories and Uni/Multidimensional Classifications[(a)(b)]

Participant	Poverty Definition	Keywords	Category	Uni or multidimensional
1	Poverty is a person or household or community which does not have adequate material means and cannot fulfil their basic needs	Material needs, basic needs	MD, BN	Multidimensional
2	A person who is poor lacks fulfilment of their basic needs.	Basic needs	BN	Multidimensional
3	For me, poverty in Bogor City is best described as communities in Bogor which do not have adequate means to lead a fulfilled life according to what is valued by that community.	Societal standards	RP	Multidimensional
4	Poverty is when a person does not have stable employment leading to not being able and enjoy the freedom to live adequately	Being able	CA	Multidimensional
5	Poverty is experienced by those who are below the standard of living considered adequate in Bogor City. Those who are poor do not have the capability to fulfil these valued standards.	Societal standards, capability	RP, CA	Multidimensional
6	Poverty is those who do not have the capability and freedom to fulfil the living conditions deemed sufficient by society to lead a fulfilled life.	Societal standards; capability	RP, CA	Multidimensional
7	Poverty is defined as the situation when a family does not have the ability to fulfil its basic needs; these basic needs include adequate food, housing, education and health. Alongside these needs, lack of a stable income also contributes to poverty.	Income, Basic needs, Ability	I, BN, CA	Multidimensional
8	Poverty is when a person or household's standard of living is below that deemed necessary in society.	Societal standards	RP	Multidimensional
9	Poverty is absent when people can fulfil their daily needs.	Daily needs	BN	Multidimensional
10	Poverty is when income is low and employment is absent or unstable thus leading to un-freedom and the inability to lead a fulfilled life.	Income, Ability	I, CA	Multidimensional

Table 4.10 (continued)[(a)(b)]

Participant	Poverty Definition	Keywords	Category	Uni or multidimensional
11	Poverty in Bogor City is a problem of lack income, lack of education, health; hence people are not free to become and enjoy valued freedoms.	Income; valued freedoms	I, CA	Multidimensional
12	Poverty is when a person cannot take part in the activities considered as vital within their community.	Societal standards	RP	Multidimensional
13	Poverty is the lack of capability of a household to fulfil material and mental needs.	Material needs, Ability	MD CA	Multidimensional
14	Poverty is the situation when a household does not have the ability to fulfil its daily needs.	Ability, Daily Needs	CA, BN	Multidimensional
15	Poverty is when income earned cannot cover the cost of daily needs.	Income, Basic Needs	I, BN	Multidimensional
16	Poverty is when household members cannot afford and do not have access to basic welfare needs.	Income, Basic needs	I, BN	Multidimensional
17	Poverty is when there are people left behind within a community; when they are not capabable to enjoy valued aspects of life.	When somebody is left behind; valued capability	RP, CA	Multidimensional
18	Poverty is the main factor that inhibits development in Bogor City. It is a condition when people within a community suffer from lack of basic needs. We have to work together to alleviate poverty and create an effective strategy to combat poverty.	Basic needs	BN	Multidimensional
19	Poverty is when a person lacks the ability to maintain their life.	Ability	CA	Multidimensional
20	Poverty is when communities do not have the ability to maintain an adequate life.	Ability	CA	Multidimensional

Table 4.10 (continued)[(a)(b)]

Participant	Poverty Definition	Keywords	Category	Uni or multidimensional
21	Poverty is the absence of the ability to independently fulfil basic needs, without the dependency on the assistance of others.	Basic Needs, Ability	BN, CA	Multidimensional
22	Poverty is when the head of the household does not have the ability to fulfil the basic needs of his household.	Ability, Basic needs	CA, BN	Multidimensional
23	Poverty is the inability of a person to have an adequate income and adequate education.	Income, Ability	I, CA	Multidimensional
24	Poverty is sufferings resulting from having a low education, leading to not fitting in to society	Societal standards	RP	Multidimensional
25	Poverty is when a person or household are not able to be free from problems of money, lack of education, health problems and lack of welfare.	Income, Basic needs, Freedom	I, BN, CA	Multidimensional
26	Poverty is when an individual, household or part of the community do not have adequate material means to fulfil their basic needs and take part in activities necessary in their society.	Basic needs, Material needs, Relative poverty	BN, MD, RP	Multidimensional
27	Poverty is when the household head does not have a job thus cannot fulfil basic needs	Basic needs	BN	Multidimensional
28	Poverty is the inability of a person to access adequate food, a healthy life, the inability of a person to be productive, to work within the community.	Ability	CA	Multidimensional
29	Poverty is when a person has low buying power and cannot fulfil the basic needs deemed necessary in society.	Societal standards, Basic Needs	RP, BN	Multidimensional

Table 4.10 (continued) [(a)(b)]

Participant	Poverty Definition	Keywords	Category	Uni or multidimensional
30	Poverty is when a person is unable to fulfil health needs and basic needs, thus lives in suffering and discomfort.	Basic needs	BN	Multidimensional
31	Poverty is first and foremost the lack of capabilities.	Ability	CA	Multidimensional
32	Poverty is when a person or family does not have the ability to earn an adequate income and live nicely.	Income, Ability	I, CA	Multidimensional
33	Poverty is the inability to access adequate education and health facilities.	Ability	CA	Multidimensional
34	Poverty is being ostracized by society because of your sufferings and inabilities, because you are not able to fulfil your basic needs.	Societal standards, Abilities, Basic needs	RP, CA, BN	Multidimensional
35	Poverty is when a community does not have the ability to live the life, fulfil daily needs of education, health, housing.	Ability, Basic needs	CA, BN	Multidimensional
36	Poverty is when a person cannot take part in social and cultural activities within his or her community because of lack of income.	Income, Relative poverty	I, RP	Multidimensional

Note: [(a)] The question of 'how poverty is defined', was asked of participants at the start of Day 1 and Day 2 of the Delphi process. As described within Section 4.2 of Chapter 4, a degree of iteration was evident in that two participants from Day 1 were replaced on Day 2. This table presents results from Day 2, as the selection of dimensions, indicators and weights, were based on these definitions.
[(b)] BN: Basic Needs, CA: Capability Approach, RP: Relative Poverty (Townsend), I: Income Poverty, MD: Material Deprivation

Source: Author's summary (Bogor City, 2016)

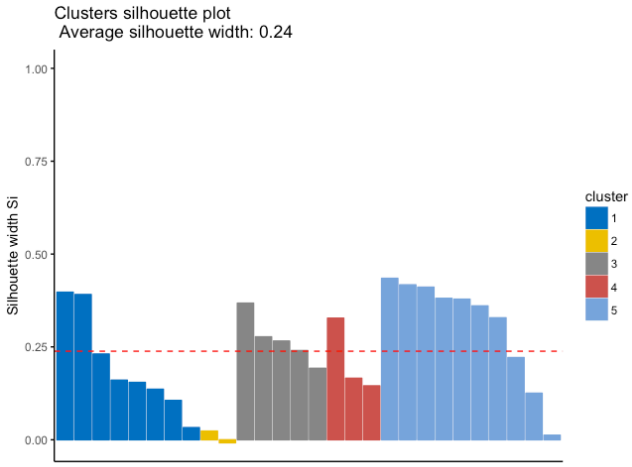
Internal Measures for Cluster Validation

1. Silhouette Analysis Plot

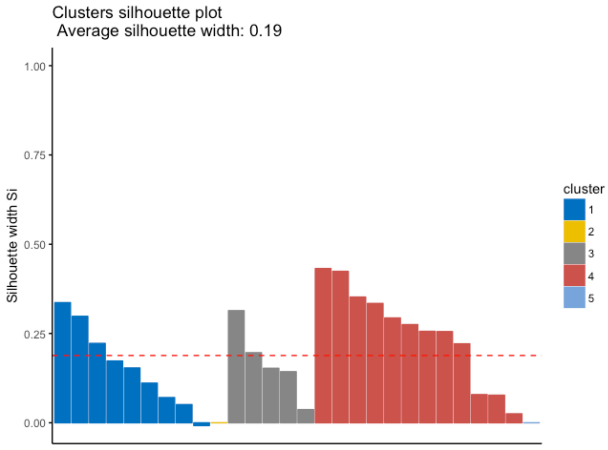
Silhouette analysis (SA) measures how well an observation is clustered and it estimates the average distance between clusters (Kassambara, 2017). SA plot displays a measure of how close each point in one cluster is to points in the neighbouring clusters.

- A value closer to 1 indicated that the object is well clustered. In other words, the object is similar to the other objects in its group.
- A value closer to -1 indicates that the object is poorly clustered, and that assignment to some other cluster would probably improve the overall results.

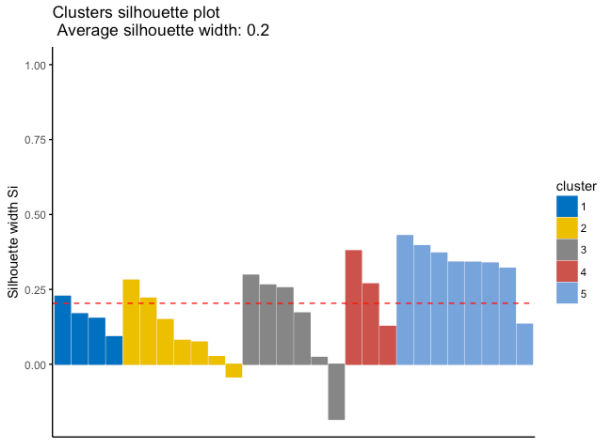
Silhouette Plot for the Complete Method



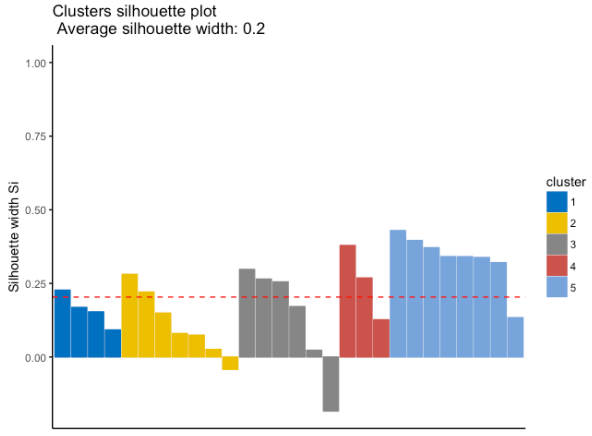
Silhouette Plot for the Average Method



Silhouette Plot for the Ward.D Method



Silhouette Plot for the Ward.D2 Method



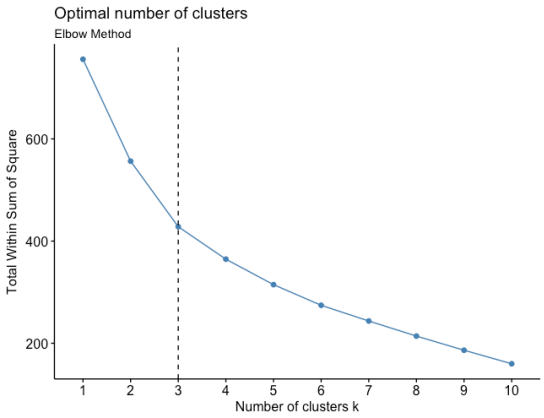
Source: Author's computations (Bogor City, 2016)

Conclusion: The complete clustering method presents the most reasonable value of average silhouette width.

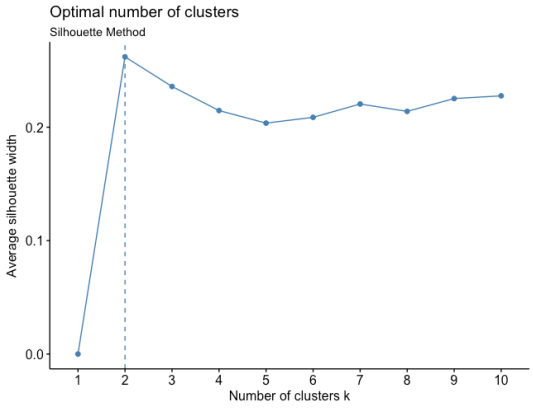
2. Optimal Number of Clusters

Kassambara (2017) states that 'the optimal number of clusters is somehow subjective and depends on the method used for measuring similarities and the parameters used for partitioning'. There are three popular approaches in determining the optimal number of clusters. These three methods were applied to analyse results from the Delphi clustering exercise. The graphs below illustrate the outcomes of the three methods.

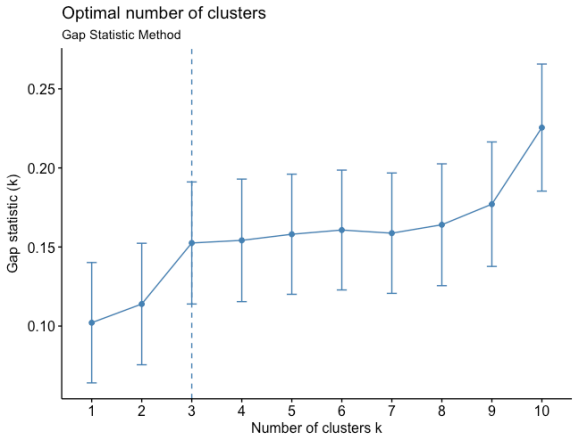
Elbow Method



Average Silhouette Method



Gap Statistic Method

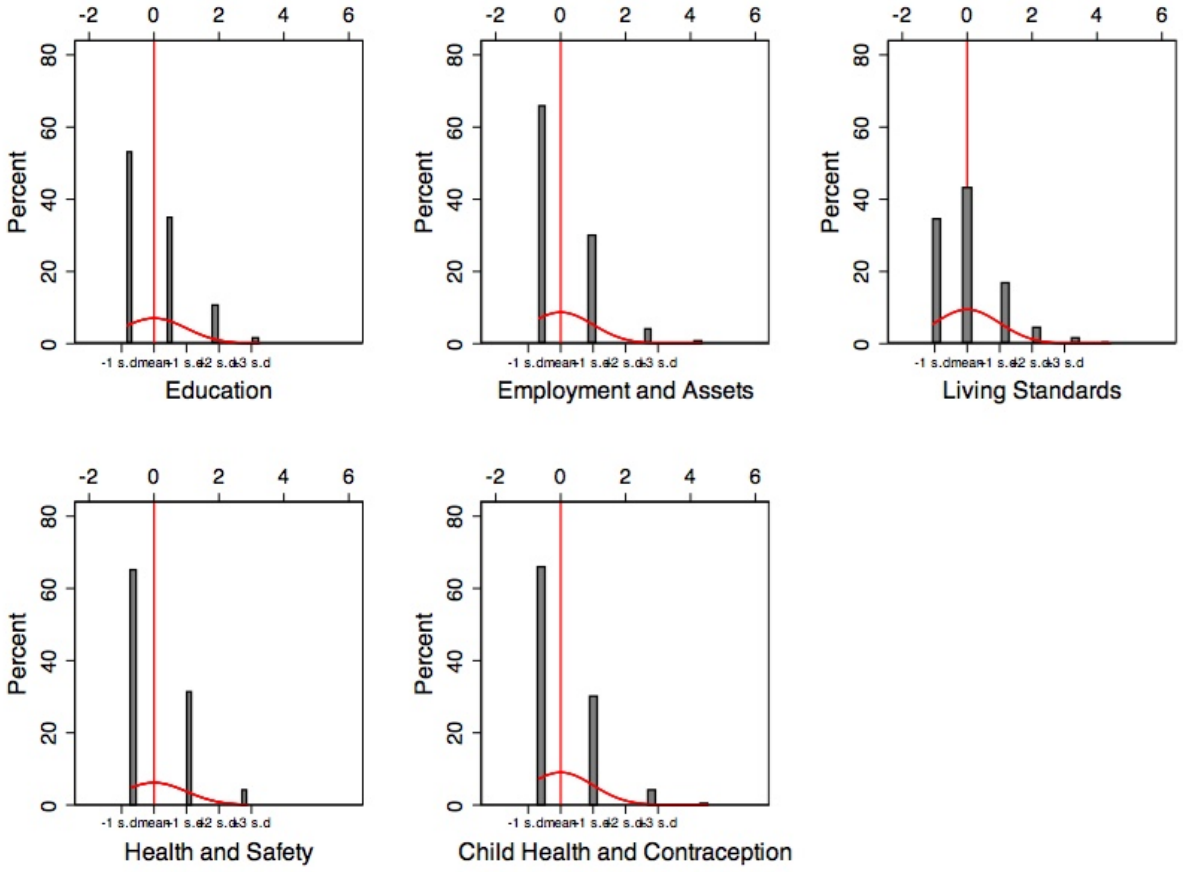


Source: Author's computations (Bogor City, 2016)

Although the results above indicate the optimal number of clusters as 'three', as described within sub-section 4.4.1 of this chapter, opinions from Delphi participants, which stated the optimal number of clusters as *five,* were used. Only one Delphi particpant chose an optimal number of dimensional clusters of 'three'. Thus, following Kassambara (2017)'s argument that 'setting the number of clusters is ultimately a highly subjective issue, which depends on the type of research conducted', the calculated decision of following the results of the Delphi participants was made.

Appendix to Chapter 5

Figure 5. 4 Standardized distribution of total deprivation in each dimension (Bogor City, 2013)



Source: Author's computations

Table 5. 3 Percent of total households who are deprived in each dimension according to different dimensional cutoffs [(a) (b)]

Dimension	Number of deprived indicators	Bogor City			West Java			Indonesia		
		Lower Bound	Deprived	Upper Bound	Lower Bound	Deprived	Upper Bound	Lower Bound	Deprived	Upper Bound
Education	At least 1 (Union)	48.6%	48.7%	48.9%	63.6%	63.6%	63.7%	59.2%	59.2%	59.2%
	At least 2	15.5%	15.6%	15.7%	24.4%	24.4%	24.4%	21.8%	21.8%	21.8%
	At least 3	3.8%	3.9%	4.0%	5.6%	5.6%	5.6%	4.6%	4.6%	4.6%
	4 (Intersection)	.	0.0%	.	0.1%	0.1%	0.1%	0.3%	0.3%	0.3%
Employment and Assets	At least 1 (Union)	62.9%	63.1%	63.3%	62.9%	62.9%	62.9%	55.3%	55.3%	55.3%
	At least 2	24.3%	24.5%	24.7%	20.5%	20.5%	20.5%	16.3%	16.3%	16.3%
	3 (Intersection)	1.5%	1.5%	1.6%	1.1%	1.1%	1.1%	0.7%	0.7%	0.7%
Living Standards	At least 1 (Union)	65.9%	66.1%	66.3%	82.6%	82.6%	82.7%	76.0%	76.0%	76.0%
	At least 2	24.3%	24.5%	24.7%	41.3%	41.3%	41.4%	38.2%	38.2%	38.2%
	At least 3	4.9%	5.0%	5.1%	14.5%	14.6%	14.6%	13.8%	13.8%	13.8%
	At least 4	1.0%	1.0%	1.0%	2.4%	2.4%	2.4%	2.5%	2.5%	2.5%
	5 (Intersection)	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.3%	0.3%	0.3%
Health and Safety	At least 1 (Union)	37.8%	37.9%	38.1%	37.9%	37.9%	38.0%	37.7%	37.7%	37.7%
	At least 2	5.1%	5.2%	5.3%	6.1%	6.1%	6.1%	6.0%	6.0%	6.0%
	3 (Intersection)	.	0.0%	.	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%
Child Health and Contraception	At least 1 (Union)	32.0%	32.2%	32.4%	37.5%	37.5%	37.5%	36.6%	36.6%	36.6%
	At least 2	3.8%	3.9%	4.0%	7.1%	7.1%	7.1%	6.7%	6.7%	6.7%
	3 (Intersection)	0.2%	0.2%	0.3%	0.5%	0.5%	0.5%	0.4%	0.4%	0.4%

Note: [(a)] Results are calculated using survey weights available within the SUSENAS 2013
[(b)] The upper and lower bounds are for the 95 % Confidence Intervals

Source: Author's computations

Appendix to Chapter 6

Computing Optimal Expert Weights for the Delphi-CIMP

Given the deprivation-aggregation function, a household's total deprivation in the five Delphi dimensions and eighteen indicators can be expressed as:

$$D_q = w_1^D\left[w_1^i D_{q,1} + w_2^i D_{q,2} + w_3^i D_{q,3} + w_4^i D_{q,4}\right] + w_2^D\left[w_5^i D_{q,5} + w_6^i D_{q,6} + w_7^i D_{q,7}\right] + w_3^D\left[w_8^i D_{q,8} + w_9^i D_{q,9} + w_{10}^i D_{q,11} + w_{12}^i D_{q,12}\right] + w_4^D\left[w_{13}^i D_{q,13} + w_{14}^i D_{q,14} + w_{15}^i D_{q,15}\right] + w_5^D\left[w_{16}^i D_{q,16} + w_{17}^i D_{q,17} + w_{18}^i D_{q,18}\right] \quad \ldots (1)$$

Note that, within the Delphi exercise, participants were asked to first choose dimension weights. For each participant, the sum of these dimension weights had to equal one. They were then asked to assign weights to indicators within dimensions. For each participant, the sum of these indicator weights within each chosen dimension, had to equal one.

To calculate 'Optimal-expert' weights, the minimisation problem is thus:

$$\min D_q = w_1^D\left[w_1^i D_{q,1} + w_2^i D_{q,2} + w_3^i D_{q,3} + w_4^i D_{q,4}\right] + w_2^D\left[w_5^i D_{q,5} + w_6^i D_{q,6} + w_7^i D_{q,7}\right] + w_3^D\left[w_8^i D_{q,8} + w_9^i D_{q,9} + w_{10}^i D_{q,11} + w_{12}^i D_{q,12}\right] + w_4^D\left[w_{13}^i D_{q,13} + w_{14}^i D_{q,14} + w_{15}^i D_{q,15}\right] + w_5^D\left[w_{16}^i D_{q,16} + w_{17}^i D_{q,17} + w_{18}^i D_{q,18}\right] \quad \ldots (1)$$

Subject to:

Bounding and non-negativity constraint:

$$\text{Lower Bound Weight} \leq w_1^D, w_2^D, \ldots, w_5^D \leq \text{Upper Bound Weight}$$
$$\text{Lower Bound Weight} \leq w_1^i, w_2^i, \ldots, w_{18}^i \leq \text{Upper Bound Weight} \quad \ldots (1.1)$$

Bounding constraint:

$w_1^D, + w_2^D + \ldots + w_5^D = 1$

$w_1^i, + w_2^i + \ldots + w_4^i = 1$

$w_4^i, + w_2^i + \ldots + w_7^i = 1$

$w_8^i, +w_2^i + \ \ldots + \ w_{12}^i = 1$

$w_{13}^i, +w_2^i + \ \ldots + \ w_{15}^i = 1$

$w_{16}^i, +w_2^i + \ \ldots + \ w_{18}^i = 1$ *…(1.2)*

Where:

D_q: total deprivation of household q

$D_{q,1}, D_{q,2}, \ldots, D_{q,18}$: a household's deprivation score in indicators 1 to 18

If a household is deprived within an indicator, it is given a score of "1". If it is not deprived, it is assigned a score of "0". As an example, if it is deprived in indicator 1, $D_{q,1}$=1.

$w_1^D, w_2^D, \ldots, w_5^D$: weights for dimensions 1 to 5

$w_1^i, w_2^i, \ldots, w_{18}^i$: weights for indicators 1 to 18

Given the above minimisation problem, optimal expert weights for each of the five *dimensions,* can be expressed mathematically as:

$$w_1^D = \left[\left(v_{d1,1}\right)\left(w_{d1,1}\right) + \left(v_{d1,2}\right)\left(w_{d1,2}\right) + \cdots + \left(v_{d1,36}\right)\left(w_{d1,36}\right)\right]$$
$$\vdots$$
$$w_5^D = \left[\left(v_{d5,1}\right)\left(w_{d5,1}\right) + \left(v_{d5,2}\right)\left(w_{d5,2}\right) + \cdots + \left(v_{d5,36}\right)\left(w_{d5,36}\right)\right]$$

where:

w_1^D : optimal-expert weight for dimension 1

$w_{d1,1}$: the weight for dimension 1 chosen by participant 1

$w_{d1,2}$: the weight for dimension 1 chosen by participant 2

$w_{d1,36}$: the weight for dimension 1 chosen by participant 36

$v_{d1,1}$: DEA generated weight to re-weight the weight for dimension 1 chosen by participant 1

$v_{d1,2}$: DEA generated weight to re-weight the weight for dimension 1 chosen by participant 2

$v_{d1,36}$: DEA generated weight to re-weight the weight for dimension 1 chosen by participant 36

w_5^D : optimal-expert weight for dimension 5

$w_{d5,1}$: the weight for dimension 5 chosen by participant 1

$w_{d5,2}$: the weight for dimension 5 chosen by participant 2

$w_{d5,36}$: the weight for dimension 5 chosen by participant 36

$v_{d5,1}$: DEA generated weight to re-weight the weight for dimension 5 chosen by participant 1

$v_{d5,2}$: DEA generated weight to re-weight the weight for dimension 5 chosen by participant 2

$v_{d5,36}$: DEA generated weight to re-weight the weight for dimension 5 chosen by participant 36

Optimal expert weights for each of the eighteen *indicators* can be expressed mathematically as:

$$w_1^i = \left[\left(v_{i1,1}\right)\left(w_{i1,1}\right) + \left(v_{i1,2}\right)\left(w_{i1,2}\right) + \cdots + \left(v_{i1,36}\right)\left(w_{i1,36}\right)\right]$$

$$\vdots$$

$$w_{18}^i = \left[\left(v_{i18,1}\right)\left(w_{i18,1}\right) + \left(v_{i18,2}\right)\left(w_{i18,2}\right) + \cdots + \left(v_{i18,36}\right)\left(w_{i18,36}\right)\right]$$

where:

w_1^i : optimal-expert weight for indicator 1

$w_{i1,1}$: the weight for indicator 1 chosen by participant 1

$w_{i1,2}$: the weight for indicator 1 chosen by participant 2

$w_{i1,36}$: the weight for indicator 1 chosen by participant 36

$v_{i1,1}$: DEA generated weight to re-weight the weight for indicator 1 chosen by participant 1

$v_{i1,2}$: DEA generated weight to re-weight the weight for indicator 1 chosen by participant 2

$v_{i1,36}$: DEA generated weight to re-weight the weight for indicator 1 chosen by participant 36

w_{18}^i : optimal-expert weight for indicator 18

$w_{i18,1}$: the weight for indicator 18 chosen by participant 1

$w_{1i8,2}$: the weight for indicator 18 chosen by participant 2

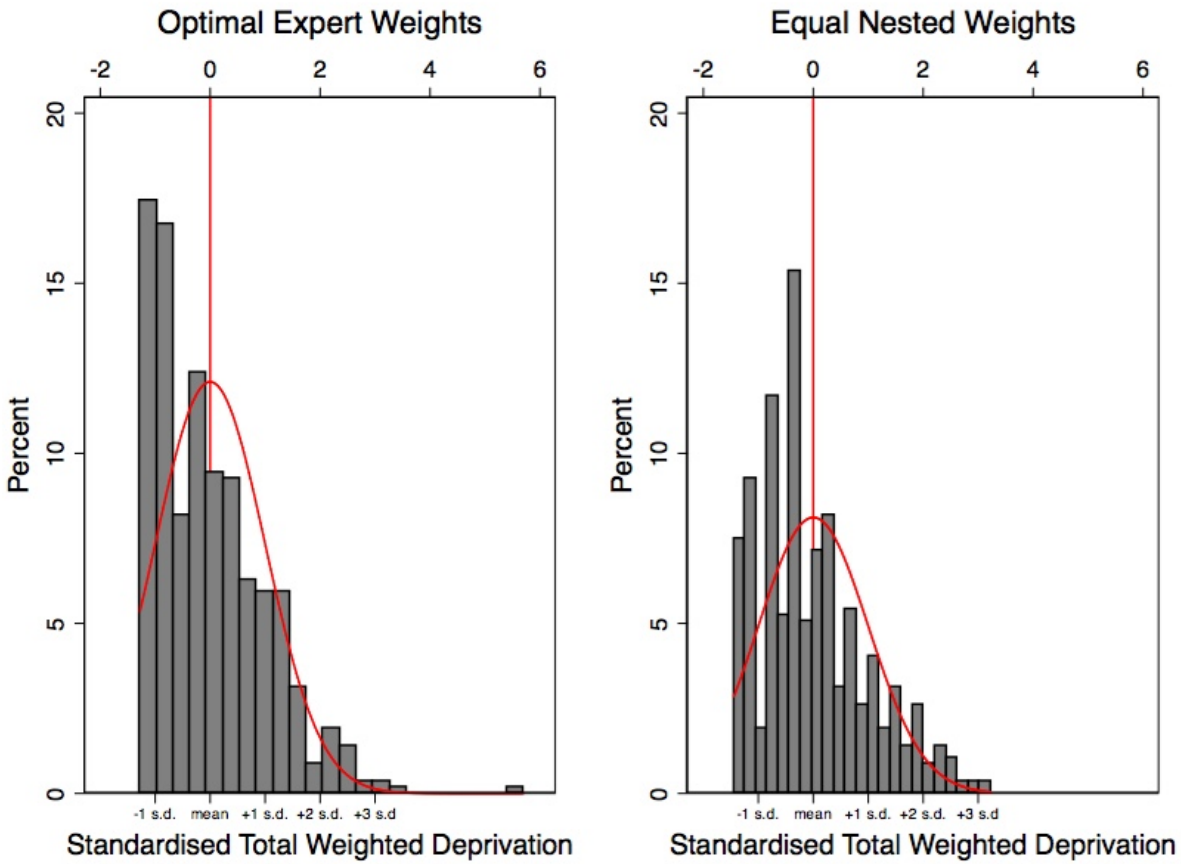
$w_{i18,36}$: the weight for indicator 18 chosen by participant 36

$v_{i18,1}$: DEA generated weight to re-weight the weight for indicator 18 chosen by participant 1

$v_{i18,2}$: DEA generated weight to re-weight the weight for indicator 18 chosen by participant 2

$v_{i18,36}$: DEA generated weight to re-weight the weight for indicator 18 chosen by participant 36

Figure 6. 3 Standardized distribution of household total weighted deprivation (Bogor City, 2013)



Source: Author's Computations

Table 6. 5 Correlations Between District Ranks if Different Poverty Cutoffs are used for the Optimal Expert Delphi-CIMP

	≧ 0.1	≧ 0.2	≧ 0.3	≧ 0.4	≧ 0.5	≧ 0.6	≧ 0.7	≧ 0.8	≧ 0.9	≧ 1	≧ 1.1	≧ 1.2	≧ 1.3	≧ 1.4	≧ 1.5
≧ 0.1	1														
≧ 0.2	0.98	1.00													
≧ 0.3	0.97	0.99	1.00												
≧ 0.4	0.97	0.97	0.98	1.00											
≧ 0.5	0.97	0.97	0.97	0.99	1.00										
≧ 0.6	0.96	0.96	0.97	0.99	1.00	1.00									
≧ 0.7	0.96	0.95	0.95	0.98	0.99	0.99	1.00								
≧ 0.8	0.95	0.95	0.95	0.98	0.99	0.99	1.00	1.00							
≧ 0.9	0.96	0.95	0.96	0.98	0.99	0.99	0.98	0.99	1.00						
≧ 1	0.96	0.95	0.96	0.97	0.98	0.98	0.98	0.98	0.99	1.00					
≧ 1.1	0.97	0.96	0.97	0.98	0.97	0.98	0.97	0.97	0.99	0.99	1.00				
≧ 1.2	0.95	0.95	0.96	0.97	0.96	0.97	0.96	0.97	0.98	0.98	0.98	1.00			
≧ 1.3	0.95	0.95	0.96	0.97	0.97	0.97	0.96	0.97	0.98	0.97	0.98	0.98	1.00		
≧ 1.4	0.94	0.94	0.95	0.96	0.96	0.97	0.95	0.96	0.97	0.97	0.98	0.98	1.00	1.00	
≧ 1.5	0.93	0.94	0.96	0.96	0.94	0.96	0.94	0.94	0.95	0.95	0.97	0.97	0.99	0.99	1.00

Table 6. 6 Correlations Between Provincial Ranks if Different Poverty Cutoffs are used for the Optimal Expert Delphi-CIMP

	≧ 0.1	≧ 0.2	≧ 0.3	≧ 0.4	≧ 0.5	≧ 0.6	≧ 0.7	≧ 0.8	≧ 0.9	≧ 1	≧ 1.1	≧ 1.2	≧ 1.3	≧ 1.4	≧ 1.5
≧ 0.1	1.00														
≧ 0.2	1.00	1.00													
≧ 0.3	1.00	0.99	1.00												
≧ 0.4	0.99	0.99	1.00	1.00											
≧ 0.5	0.99	0.99	0.99	1.00	1.00										
≧ 0.6	0.99	0.99	0.99	1.00	1.00	1.00									
≧ 0.7	0.99	0.99	0.99	0.99	0.99	1.00	1.00								
≧ 0.8	0.98	0.98	0.99	0.99	0.99	0.99	1.00	1.00							
≧ 0.9	0.98	0.98	0.99	0.99	0.99	0.99	1.00	1.00	1.00						
≧ 1	0.98	0.98	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00					
≧ 1.1	0.97	0.98	0.98	0.98	0.98	0.98	0.99	0.99	0.99	1.00	1.00				
≧ 1.2	0.97	0.97	0.98	0.98	0.98	0.98	0.99	0.99	0.99	0.99	1.00	1.00			
≧ 1.3	0.96	0.97	0.97	0.97	0.97	0.98	0.99	0.99	0.99	0.99	1.00	0.99	1.00		
≧ 1.4	0.96	0.97	0.97	0.97	0.97	0.97	0.98	0.99	0.99	0.99	1.00	0.99	1.00	1.00	
≧ 1.5	0.96	0.96	0.96	0.97	0.96	0.97	0.98	0.98	0.99	0.99	1.00	0.99	1.00	1.00	1.00

Source: Author's Computations

Table 6. 7 Correlations Between District Ranks if Different Poverty Cutoffs are used for the Equal Nested Delphi-CIMP

	≧ 0.1	≧ 0.2	≧ 0.3	≧ 0.4	≧ 0.5	≧ 0.6	≧ 0.7	≧ 0.8	≧ 0.9	≧ 1	≧ 1.1	≧ 1.2	≧ 1.3	≧ 1.4	≧ 1.5
≧ 0.1	1.00														
≧ 0.2	0.99	1.00													
≧ 0.3	0.99	0.99	1.00												
≧ 0.4	0.96	0.97	0.97	1.00											
≧ 0.5	0.97	0.97	0.97	0.99	1.00										
≧ 0.6	0.97	0.95	0.95	0.96	0.98	1.00									
≧ 0.7	0.95	0.94	0.93	0.95	0.97	0.99	1.00								
≧ 0.8	0.93	0.92	0.91	0.93	0.95	0.98	0.99	1.00							
≧ 0.9	0.93	0.92	0.92	0.93	0.94	0.98	0.98	0.99	1.00						
≧ 1	0.93	0.92	0.92	0.94	0.95	0.98	0.99	0.99	0.99	1.00					
≧ 1.1	0.92	0.91	0.90	0.93	0.93	0.97	0.97	0.99	0.98	0.98	1.00				
≧ 1.2	0.91	0.91	0.90	0.92	0.92	0.96	0.97	0.98	0.97	0.98	0.99	1.00			
≧ 1.3	0.92	0.92	0.91	0.92	0.93	0.95	0.96	0.95	0.96	0.96	0.96	0.97	1.00		
≧ 1.4	0.92	0.92	0.92	0.93	0.93	0.95	0.95	0.94	0.95	0.95	0.94	0.96	0.99	1.00	
≧ 1.5	0.95	0.95	0.95	0.94	0.95	0.97	0.97	0.97	0.97	0.97	0.96	0.97	0.98	0.98	1.00

Table 6. 8 Correlations Between Provincial Ranks if Different Poverty Cutoffs are used for the Equal Nested Delphi-CIMP

	≧ 0.1	≧ 0.2	≧ 0.3	≧ 0.4	≧ 0.5	≧ 0.6	≧ 0.7	≧ 0.8	≧ 0.9	≧ 1	≧ 1.1	≧ 1.2	≧ 1.3	≧ 1.4	≧ 1.5
≧ 0.1	1														
≧ 0.2	1.00	1.00													
≧ 0.3	0.99	0.99	1.00												
≧ 0.4	0.99	0.99	1.00	1.00											
≧ 0.5	0.99	0.99	1.00	1.00	1.00										
≧ 0.6	0.99	0.99	1.00	1.00	1.00	1.00									
≧ 0.7	0.99	0.99	1.00	1.00	1.00	1.00	1.00								
≧ 0.8	0.99	0.99	1.00	0.99	0.99	0.99	1.00	1.00							
≧ 0.9	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00						
≧ 1	0.98	0.98	0.99	0.99	0.99	0.99	0.99	0.99	1.00	1.00					
≧ 1.1	0.98	0.98	0.99	0.98	0.99	0.99	0.99	0.99	0.99	1.00	1.00				
≧ 1.2	0.97	0.97	0.98	0.98	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00			
≧ 1.3	0.98	0.98	0.99	0.98	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00		
≧ 1.4	0.97	0.97	0.98	0.98	0.98	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00	
≧ 1.5	0.97	0.97	0.98	0.98	0.98	0.98	0.98	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00

Source: Author's Computations

Appendix to Chapter 7

Table 7. 7 Percent of Households Deprived in Each Indicator

		Bogor City 95% Confidence Interval			West Java 95% Confidence Interval			Indonesia 95% Confidence Interval		
Dimension	Indicator	% deprived	Lower Bound	Upper Bound	% deprived	Lower Bound	Upper Bound	% deprived	Lower Bound	Upper Bound
DELPHI-CIMP										
EDUCATION	Children's Education	23.8%	23.6%	23.9%	20.2%	20.2%	20.2%	19.1%	19.1%	19.1%
	Adult's Education	15.8%	15.7%	16.0%	33.4%	33.4%	33.4%	29.4%	29.4%	29.4%
	Literacy	0.4%	0.3%	0.4%	1.8%	1.7%	1.8%	2.9%	2.9%	2.9%
	Ownership of Birth Certificate	28.3%	28.1%	28.5%	38.3%	38.3%	38.3%	34.6%	34.5%	34.6%
EMPLOYMENT AND ASSETS	Employment	4.5%	4.4%	4.6%	7.2%	7.2%	7.2%	5.9%	5.9%	5.9%
	Assets Ownership	4.0%	3.9%	4.1%	9.9%	9.9%	9.9%	8.4%	8.4%	8.4%
	House/Land Ownership	32.3%	32.1%	32.5%	21.8%	21.8%	21.9%	20.5%	20.5%	20.5%
LIVING STANDARDS	Consumption	14.1%	14.0%	14.2%	10.7%	10.7%	10.8%	11.7%	11.6%	11.7%
	Sanitation	27.2%	27.0%	27.3%	40.1%	40.0%	40.1%	35.0%	35.0%	35.0%
	Drinking Water	48.1%	47.9%	48.3%	67.8%	67.8%	67.8%	57.6%	57.6%	57.6%
	Cooking/Washing Water	9.6%	9.5%	9.7%	23.4%	23.3%	23.4%	26.6%	26.5%	26.6%
	Government Help	0.9%	0.9%	0.9%	1.0%	1.0%	1.1%	1.6%	1.6%	1.6%
HEALTH AND SAFETY	Healthy life	17.8%	17.7%	18.0%	17.4%	17.4%	17.4%	17.2%	17.2%	17.2%
	immunizations	23.6%	23.4%	23.7%	23.9%	23.9%	24.0%	23.8%	23.8%	23.8%
	Safety	1.8%	1.7%	1.8%	2.9%	2.9%	2.9%	2.9%	2.9%	2.9%
CHILD HEALTH AND CONTRACEPTION	Birth Attendant	6.1%	6.0%	6.2%	8.6%	8.5%	8.6%	6.4%	6.4%	6.4%
	Child mortality	15.4%	15.2%	15.5%	18.4%	18.4%	18.4%	18.1%	18.1%	18.1%
	Contraception	15.0%	14.9%	15.2%	17.1%	17.1%	17.1%	18.7%	18.7%	18.7%

Table 7.7 (continued)

		Bogor City 95% Confidence Interval			West Java 95% Confidence Interval			Indonesia 95% Confidence Interval		
Dimension	Indicator	% deprived	Lower Bound	Upper Bound	% deprived	Lower Bound	Upper Bound	% deprived	Lower Bound	Upper Bound
GLOBAL MPI										
EDUCATION	Child School Attendance	2.9%	2.9%	3.0%	2.6%	2.6%	2.6%	2.4%	2.4%	2.4%
	Years of Schooling	2.5%	2.5%	2.6%	7.5%	7.5%	7.5%	8.6%	8.6%	8.6%
LIVING STANDARDS	Electricity	0.1%	0.1%	0.1%	0.3%	0.3%	0.3%	3.5%	3.5%	3.5%
.	Sanitation	27.2%	27.0%	27.3%	40.1%	40.0%	40.1%	35.0%	35.0%	35.0%
	Drinking Water	48.1%	47.9%	48.3%	67.8%	67.8%	67.8%	57.6%	57.6%	57.6%
	Housing	1.5%	1.5%	1.6%	18.8%	18.8%	18.9%	16.2%	16.2%	16.2%
	Cooking Fuel	3.5%	3.4%	3.6%	21.1%	21.1%	21.2%	32.6%	32.6%	32.6%
	Asset Ownership	24.4%	24.2%	24.6%	30.0%	29.9%	30.0%	23.7%	23.7%	23.7%
HEALTH	Child Mortality	15.4%	15.2%	15.5%	18.4%	18.4%	18.4%	18.1%	18.1%	18.1%
ADJUSTED MPI										
EDUCATION	Access to Pre-school Education	21.3%	21.2%	21.5%	18.0%	18.0%	18.0%	17.1%	17.1%	17.1%
	Education Continuity	70.7%	70.5%	70.8%	81.8%	81.8%	81.8%	81.3%	81.3%	81.3%
	Literacy	0.4%	0.3%	0.4%	1.8%	1.7%	1.8%	2.9%	2.9%	2.9%
LIVING STANDARDS	Electricity	0.1%	0.1%	0.1%	0.3%	0.3%	0.3%	3.5%	3.5%	3.5%
.	Cooking Fuel	3.5%	3.4%	3.6%	21.1%	21.1%	21.2%	32.6%	32.6%	32.6%
	Roof, floor and wall	0.2%	0.2%	0.2%	3.6%	3.6%	3.6%	4.2%	4.2%	4.2%
	House Ownership	32.3%	32.1%	32.5%	21.8%	21.8%	21.9%	20.5%	20.5%	20.5%
HEALTH	Sanitation	27.2%	27.0%	27.3%	40.1%	40.0%	40.1%	35.0%	35.0%	35.0%
	Drinking Water	48.1%	47.9%	48.3%	67.8%	67.8%	67.8%	57.6%	57.6%	57.6%
	Birth Attendant	6.1%	6.0%	6.2%	8.6%	8.5%	8.6%	6.4%	6.4%	6.4%
	Under 5 nutritional intake	35.5%	35.3%	35.6%	32.3%	32.3%	32.3%	32.4%	32.4%	32.4%

Source: Author's Computations (using Susenas 2013)

Table 7. 8 Tetrachoric Correlations Between Household Poverty Status in the Four Measures

	Bogor City				West Java				Indonesia			
	Delphi CIMP	Consumption	Global MPI	Adjusted MPI	Delphi CIMP	Consumption	Global MPI	Adjusted MPI	Delphi CIMP	Consumption	Global MPI	Adjusted MPI
Delphi CIMP	1				1				1			
Consumption	0.56	1			0.49	1			0.47	1		
Global MPI	0.31	0.2	1		0.40	0.18	1		0.40	0.18	1	
Adjusted MPI	0.78	0.3	0.1	1	0.66	0.37	0.21	1	0.65	0.37	0.21	1

Source: Author's Computations

Table 7. 9 Logistic Regression Results Examining the Composition of the Poor Across Four Difference Poverty Measures

	Bogor City				West Java				Indonesia			
	Delphi	C	GMPI	AMPI	Delphi	C	GMPI	AMPI	Delphi	C	GMPI	AMPI
Age	-0.25***	-0.02	0.06	-0.31***	-0.15***	-0.08***	0.05***	-0.18***	-0.14***	-0.07***	0.01***	-0.16***
Age Squared	.002***	0.0002	-0.0002	0.003***	0.001***	0.001***	-0.0001	0.002***	0.001***	0.001***	0.0002***	0.001***
Gender (ref: Male)												
Female	0.82	0.22	-1.67	-0.99	-0.23*	-0.25	-0.98***	-0.03	-0.31***	-0.33***	-0.80***	0.05*
Marital Status (ref: Single, Divorced/ Widowed)												
Married	0.94	0.21	-1.36	-1.08	-0.56***	-0.22	-0.63***	-0.17*	-0.43***	-0.27***	-0.55***	0.05*
Education Level (ref: No qual)												
Primary	0.17	-0.36	-1.08***	0.04	-0.78***	-0.29***	-0.94***	-0.51***	-0.96***	-0.49***	-1.06***	-0.65***
Secondary	-0.81	-0.98*	-1.88***	-0.85	-1.32***	-0.93***	-1.31***	-0.69***	-1.45***	-0.93***	-1.29**	-0.84***
Higher Ed	-1.61***	-2.83***	-3.29***	-1.70***	-2.08***	-1.97***	-1.85***	-1.79***	-1.77***	-1.63***	-1.60***	-1.79***
Work Sector (ref: Extractive Industries)												
Economically Inactive	-0.01	0.69	-0.97	0.6	0.81***	-0.13	-0.14*	-0.28***	0.48***	-0.13***	-0.18***	-0.51***
Manufacturing & Construction	-0.83	0.45	-1.01	0.8	-0.24***	-0.45***	-0.1	-0.21***	-0.32***	-0.32***	-0.26***	-0.39***
Trade& Services	-0.85	0	-0.85	0.45	-0.26***	-0.54***	-0.12*	-0.36***	-0.44***	-0.72***	-0.33***	-0.57***
Community Service& Public Administration	-0.64	0.6	-1.23	0.85	-0.22*	-0.26**	-0.06	-0.25***	-0.33***	-0.47***	-0.23***	-0.50***
Other	0.28	0.83	(empty)	(empty)	0.14	0.19	-0.12	-0.12	-0.16*	-0.17*	-0.23***	-0.18***

Table 7.9 (continued)

	Bogor City				West Java				Indonesia			
	Delphi	C	GMPI	AMPI	Delphi	C	GMPI	AMPI	Delphi	C	GMPI	AMPI
Household Composition												
Number of adults in household (ref: 2 Adults)												
1 adult	0.66	-0.84	-0.03	-0.59	1.03***	-0.66***	0.07	0.15*	0.98***	-0.53***	0.22***	0.23***
More than 2 adults	-1.08***	-0.02	0.46	-0.43	-0.49***	0.10*	0.16***	-0.36***	-0.57***	0.17***	0.16***	-0.34***
Number of children in household (ref: No Children)												
1 child	1.50***	0.96*	0.73*	1.2*	0.60***	0.52***	-0.03	0.49***	0.51***	0.56***	-0.1***	0.48***
2 children	1.52**	2.52***	-0.03	1.15*	1.34***	1.29***	-0.03	0 .95***	1.19***	1.21***	-0.08***	0.90***
More than 2 children	2.39***	2.58***	1.01*	2.04***	2.32***	2.09***	0.23**	1.65***	2.19***	1.97***	0.23***	1.62***
Urban Rural Classification (ref: Rural)												
Urban					-0.33***	0.41***	-0.23***	-0.52***	-0.42***	-0.12***	-0.41***	-0.60***
Constant	4.25*	-2.41	-0.63	6.76**	2.52***	-0.39	-1.74***	4.61***	2.25***	-0.23**	-0.16*	4.25***
Log likelihood	-213	-212.6	-210.3	-204	-6540.3	-6623	-8771.9	-9811.7	-82147	-89111	-125220	-133085
Number of Obs	573	680	617	617	18187	22441	20020	20021	231022	284005	259105	259115
LR chi2(21)	110.72	132.07	166.87	115.21	2664.5	1838.47	3874.32	3201.85	34831.47	26316.86	44478.5	55903.1
Prob > chi2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Pseudo R2	0.21	0.24	0.28	0.22	0.17	0.12	0.18	0.14	0.17	0.13	0.15	0.17

Source: Author's Computations (Susenas 2013)

Table 7. 10 Indicator with Missing Values and Treatment of Missing Values

Indicator	Level of Analysis	Sample Size (Number of Households)	Missing Values (Number of Households)	Percent Retained Sample (%)	Treatment of missing values
Child Mortality	Indonesia	284,063	10	99.9	Missing households considered missing.
	West Java	22,441	1	99.9	Missing households considered missing.
	Bogor City	680	0	100	N/A
Nutrition	Indonesia	0	All missing	0	Indicator dropped.
	West Java	0	All missing	0	Indicator dropped.
	Bogor City	0	All missing	0	Indicator dropped.
Drinking Water	Indonesia	284,063	24,896	91	Missing households considered missing.
	West Java	22,441	2,420	89	Missing households considered missing.
	Bogor City	680	57	92	Missing households considered missing.
Contraception	Indonesia	284,063	30,345	89	Missing households considered missing.
	West Java	22,441	3,430	85	Missing households considered missing.
	Bogor City	680	54	92	Missing households considered missing.
Birth Attendant	Indonesia	284,063	3	99.99	Missing households considered missing.
	West Java	22,441	0	100	N/A
	Bogor City	680	0	100	N/A
Immunization	Indonesia	284,063	3	99.99	Missing households considered missing.
	West Java	22,441	0	100	N/A
	Bogor City	680	0	100	N/A

Source: Author's summary (Susenas 2013)

Table 7. 11 Percent Retained Sample for Each Measure

Indicator	Level of Analysis	Sample Size (Number of Households)	Missing Values (Number of Households)	Percent Retained Sample (%)	Treatment of missing values
Consumption	Indonesia	284,063	0	100	N/A
	West Java	22,441	0	100	N/A
	Bogor City	680	0	100	N/A
Global MPI	Indonesia	284,063	10	99.9	Missing households not included in final analytical sample
	West Java	22,441	1	99.9	Missing households not included in final analytical sample
	Bogor City	680	0	99.9	N/A
Adjusted MPI	Indonesia	284,063	24,906	91.2	Missing households not included in final analytical sample
	West Java	22,441	2,421	89.2	Missing households not included in final analytical sample
	Bogor City	680	57	92	Missing households not included in final analytical sample
Delphi-CIMP	Indonesia	284,063	52,989	81.4	Missing households not included in final analytical sample (sensitivity tests conducted confirmed relevance of retained sample)
	West Java	22,441	4,254	81	Missing households not included in final analytical sample (sensitivity tests conducted confirmed relevance of retained sample)
	Bogor City	680	107	84.3	Missing households not included in final analytical sample (sensitivity tests conducted confirmed relevance of retained sample)

Source: Author's summary (Susenas 2013)

References

Alatas, V., Banerjee, A., Hanna, R., Olken, B. A., Purnamasari, R., & Wai-Poi, M. (2013). *Does elite capture matter? Local elites and targeted welfare programs in Indonesia*. Retrieved from

Alatas, V., Banerjee, A., Hanna, R., Olken, B. A., & Tobias, J. (2012). Targeting the Poor: Evidence from a Field Experiment in Indonesia. *American Economic Review, 102*(4), 1206-1240. doi:doi: 10.1257/aer.102.4.1206

Alisjahbana, A. S., & Manning, C. (2006). Labour market dimensions of poverty in Indonesia. *Bulletin of Indonesian Economic Studies, 42*(2), 235-261. doi:10.1080/00074910600873674

Alkire, Conconi, A., Robles, G., Roche, J. M., Santos, M. E., & Vaz, A. (2016). The Global Multidimensional Poverty Index (MPI): 5-Year Methodological Note. *Documento de trabajo*(37).

Alkire, S. (2002). *Valuing Freedoms: Sen's Capability Approach and Poverty Reduction*: Oxford University Press.

Alkire, S. (2007). Choosing Dimensions: The Capability Approach and Multidimensional Poverty. In N. Kakwani; & J. Silber (Eds.), *The Many Dimensions of Poverty*. United Kingdom: Palgrave Macmillan.

Alkire, S. (2011). Multidimensional poverty and its discontents. *OPHI Working Paper*(No. 46).

Alkire, S. (2014). The Capability Approach and Well Being MPhil Lecture. Department of International Development University of Oxford.

Alkire, S., & Deneulin, S. (2009). Introducing the Human Development and Capability Approach. *An Introduction to the Human Development and Capability Approach, London: Earthscan*.

Alkire, S., & Foster, J. (2011). Counting and multidimensional poverty measurement. *Journal of Public Economics, 95*(7–8), 476-487. doi:http://dx.doi.org/10.1016/j.jpubeco.2010.11.006

Alkire, S., & Foster, J. (2011). Understandings and misunderstandings of multidimensional poverty measurement. *The Journal of Economic Inequality, 9*(2), 289-314.

Alkire, S., Foster, J., & Santos, M. E. (2011). Where did identification go? *Journal of Economic Inequality, 9*(3), 501-505.

Alkire, S., Kanagaratnam, U., & Suppa, N. (2019). The Global Multidimensional Poverty Index (MPI) 2019. Retrieved from https://ophi.org.uk/wp-content/uploads/OPHI_MPI_MN_47_2019_vs2.pdf. from Oxford Poverty and Human Development Initiative https://ophi.org.uk/wp-content/uploads/OPHI_MPI_MN_47_2019_vs2.pdf

Alkire, S., Kanagaratnam, U., & Suppa, N. (2019). The Global Multidimensional Poverty Index (MPI) 2019. OPHI MPI Methodological Note, 47. In: Oxford Poverty and Human Development Initiative, University of Oxford. State ….

Alkire, S., Kanagaratnam, U., & Suppa, N. (2020). The global multidimensional poverty index (MPI) 2020. *OPHI MPI Methodological Note*(49).

Alkire, S., & Santos, M. E. (2010a). Acute multidimensional poverty: A new index for developing countries. *United Nations development programme human development report office background paper*(2010/11).

Alkire, S., & Santos, M. E. (2010b). Multidimensional poverty index: 2010 data. *Retrieved August, 14*, 2010.

Alkire, S., Santos, M. E., Seth, S., & Yalonetzky, G. (2010). Is the Multidimensional Poverty Index robust to different weights? *OPHI Research in Progress*(22a).

Anand, P., & Van Hees, M. (2006). Capabilities and achievements: An empirical study. *The Journal of Socio-Economics, 35*(2), 268-284.

Anand, S., & Sen, A. (1997). *Concepts of Human Development and Poverty: A Multidimensional Perspective*. Retrieved from

Anderson, G., Crawford, I., & Leicester, A. (2008). Efficiency analysis and the lower convex hull approach. In J. S. N. Kakwani (Ed.), *Quantitative approaches to multidimensional poverty measurement* (pp. 176-191): Springer.

Anyanwu, J. C. (2005). Rural poverty in Nigeria: profile, determinants and exit paths. *African Development Review, 17*(3), 435-460.

Anyanwu, J. C. (2014). Marital Status, Household Size and Poverty in Nigeria: Evidence from the 2009/2010 Survey Data. *African Development Review, 26*(1), 118-137. doi:10.1111/1467-8268.12069

Armstrong, J. S. (2001). *Principles of forecasting: a handbook for researchers and practitioners* (Vol. 30): Springer Science & Business Media.

Atkinson, A. B. (1987). On the measurement of poverty. *Econometrica: Journal of the Econometric Society*, 749-764.

Atkinson, A. B. (2003). Multidimensional deprivation: contrasting social welfare and counting approaches. *The Journal of Economic Inequality, 1*(1), 51-65.

Atkinson, A. B., & Marlier, E. (2010). *Income and living conditions in Europe*: European Union.

Atkinson, A. B., Rainwater, L., & Smeeding, T. M. (1994). *Income distribution in European countries*. Retrieved from

Atkinson, T., Cantillon, B., Marlier, E., & Nolan, B. (2002). *Social indicators: The EU and social inclusion*. Oxford: Oxford University Press.

Attanayake, N., Fox-Rushby, J., & Mills, A. (2001). Household costs of 'malaria' morbidity: a study in Matale district, Sri Lanka. *Tropical Medicine & International Health, 5*(9), 595-606. doi:10.1046/j.1365-3156.2000.00612.x

Australian National University (ANU). (2011). *Indonesian National Social Economic Survey (2010) Susenas 2010: A User's Guide to Data Files Held at the Australian National University*.

Badan Perencanaan Pembangunan Nasional (BAPPENAS). (2014). Rencana Pembangunan Jangka Menengah Nasional 2015-2019. *Presiden Republik Indonesia*.

Badan Perencanaan Pembangunan Nasional (BAPPENAS). (2018). *Republic of Indonesia Ministry of Development Planning Law Number 7 Year 2018 On The Coordination and Evaluation of The National Long Term Development Plan*. Jakarta, Indonesia

Baird, J., Ma, S., & Ruger, J. P. (2011). Effects of the World Bank's maternal and child health intervention on Indonesia's poor: Evaluating the safe motherhood project. *Social Science & Medicine, 72*(12), 1948-1955. doi:https://doi.org/10.1016/j.socscimed.2010.04.038

Baker, J. L., Burger, N., Glick, P., Perez-Arce, F., Rabinovich, L., Rana, Y., . . . Yoong, J. (2013). *Indonesia-Urban poverty and program review*. Retrieved from

Balestrino, A., & Sciclone, N. (2001). Should we use functionings instead of income to measure well-being? Theory, and some evidence from Italy. *Rivista internazionale di scienze sociali*, 3-22.

Banker, R. D., Cooper, W., Grifell-Tajté, E., Pastor, J. T., Wilson, P. W., Ley, E., & Lovell, C. (1994). Validation and generalization of DEA and its uses. *Top, 2*(2), 249-314.

Beretta, R. (1996). A critical review of the Delphi technique. *Nurse researcher, 3*(4), 79-89.

Berk, R. A., Lenihan, K. J., & Rossi, P. H. (1980). Crime and poverty: Some experimental evidence from ex-offenders. *American Sociological Review*, 766-786.

Berloffa, G., & Modena, F. (2013). Income shocks, coping strategies, and consumption smoothing: An application to Indonesian data. *Journal of Asian Economics, 24*, 158-171. doi:https://doi.org/10.1016/j.asieco.2012.11.004

Bessell, S. (2009). Indonesian Children's views and experiences of work and poverty. *Social Policy and Society, 8*(4), 527-540.

Bhargava, A., & Ravallion, M. (1993). Does household consumption behave as a Martingale? A test for rural South India. *The Review of Economics and Statistics*, 500-504.
Bhatia, A., Ferreira, L. Z., Barros, A. J., & Victora, C. G. (2017). Who and where are the uncounted children? Inequalities in birth certificate coverage among children under five years in 94 countries using nationally representative household surveys. *International journal for equity in health, 16*(1), 148.
Bleich, E., & Pekkanen, R. (2013). How to report interview data. *Interview research in political science, 1*, 84-105.
Bogor City Delphi Non-Participants. (2016, 24 August - 7 October 2016) *Bogor City Delphi Robustness Check: Semistructured Interviews with Non-Participants/Interviewer: P. Natih*.
Bogor City Delphi Participants. (2016, 24 August - 7 October 2016) *Bogor City Delphi Robustness Check: Semistructured Interviews /Interviewer: P. Natih*. Bogor City.
Bogor City Government. (2016). Press Release: Bogor City Government Works Together with Oxford Researcher on Urban Poverty [Press release]. Retrieved from https://kotabogor.go.id/index.php/show_post/detail/4391/oxford-university-lakukan-penelitian-pemetaan-kemiskinan-di-kota-bogor
Booth, A. (2003). Decentralisation and Poverty Alleviation in Indonesia. *Environment and Planning C: Government and Policy, 21*(2), 181-202. doi:10.1068/c0127
Booth, D. (1998). *Participation and combined methods in African poverty assessment: renewing the agenda*: Department for International Development.
Bourguignon, F. (2009). Crime as a Social Cost of Poverty and Inequality: A Review Focusing on Developing countries. In S. Yusuf;, S. Evenett;, & W. Wu; (Eds.), *Facets of Globalization: International and Local Dimensions of Development*.
Bourguignon, F., & Chakravarty, S. R. (2003). The measurement of multidimensional poverty. *The Journal of Economic Inequality, 1*(1), 25-49.
Bowles, N. (1999). The Delphi technique. *Nursing Standard (through 2013), 13*(45), 32.
Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*: Cambridge university press.
Brandolini, A. (2008). On applying synthetic indices of multidimensional well-being: Health and income inequalities in selected EU countries.
Brannen, J. (2005). Mixing methods: The entry of qualitative and quantitative approaches into the research process. *International Journal of Social Research Methodology, 8*(3), 173-184.
Brennan, B. (1997). Reconceptualizing non-formal education. *International Journal of Lifelong Education, 16*(3), 185-200.
Brock, C., & Symaco, L. P. (2011). *Education in south-east Asia*.
Brodjonegoro, B. (2002). Fiscal Decentralization in Indonesia. In H. Soesastro, A. L. Smith, & M. L. Han (Eds.), *Governance in Indonesia: challenges facing the Megawati presidency* (Vol. 239): Institute of Southeast Asian Studies.
Brodjonegoro, B. (2003). The Indonesian Decentralization after law revision: Toward a better future. *Department of Economics, University of Indonesia*.
Brodjonegoro, B. (2004). Three years of fiscal decentralization in Indonesia: its impacts on regional economic development and fiscal sustainability Paper read at International Symposium on Fiscal Decentralization in Asia Revisited. *Hitotsubashi, Chiyoda-ku, Tokyo*, 20-21.
Budiantoro, S. (2016). Multidimensional Poverty Index (MPI): Konsep dan Pengukurannya di Indonesia. *Prakarsa Economic Policy Working Paper*.
Budiantoro, S., Martha, L. F., & Sagala, M. (2020). *Perhitungan Indeks Kemiskinan Multidimensi Indonesia 2012-2014*. Retrieved from Jakarta:
Cahyat, A., Gönner, C., & Haug, M. (2007). *Assessing household poverty and wellbeing: a manual with examples from Kutai Barat, Indonesia*: CIFOR.

Callan, T., Nolan, B., & Whelan, C. T. (1993). Resources, deprivation and the measurement of poverty. *Journal of Social Policy, 22*(02), 141-172.
Cameron, S. (2012). *Education, urban poverty and migration: Evidence from Bangladesh and Vietnam*. Retrieved from Florence:
Carlucci, F., & Pisani, S. (1995). A multiattribute measure of human development. *Social Indicators Research, 36*(2), 145-176.
Central Bureau of Statistics. (2005). *Pelaksanaan Pendataan Rumah Tangga Miskin 2005*. Jakarta
Central Bureau of Statistics. (2013a). *Data dan Informasi Kemiskinan Kabupaten/Kota Tahun 2013*. Jakarta
Central Bureau of Statistics. (2013b). Metadata for the National Socioeconomic Survey (Susenas) 2013. Retrieved from http://sirusa.bps.go.id/index.php?r=sd/view&kd=2181&th=2011
Central Bureau of Statistics. (2018a). The Measurement of Unemployment in Indonesia. Retrieved from https://www.bps.go.id/subject/6/tenaga-kerja.html#subjekViewTab2
Central Bureau of Statistics. (2018b). Persentase Rumah Tangga dengan Air Minum Layak. Retrieved from https://sirusa.bps.go.id/index.php?r=indikator/view&id=478
Central Bureau of Statistics (CBS). (2014). *Statistik Telekomunikasi Indonesia*. Retrieved from Jakarta:
Central Bureau of Statistics (CBS). (2019). Gini Rasio per Provinsi. [Gini Ratio for Provinces]. Retrieved from https://www.bps.go.id/dynamictable/2017/04/26/1116/gini-ratio-provinsi-2002-2019.html
Central Bureau of Statistics (CBS). (2020). Kriteria Air Minum Layak. Retrieved from https://www.bps.go.id/subject/29/perumahan.html
Central Bureau of Statistics Bogor City. (2018). *Kota Bogor Dalam Angka 2018*. Bogor City Retrieved from http://bappeda.kotabogor.go.id/images/slidertabs/90edeb5c500f29f56f65d29a484f6052.pdf
Central Bureau of Statistics Indonesia (CBS). (2014). *Penghitungan dan Analisis Kemiskinan Makro Indonesia Tahun 2014*. Indonesia: Central Bureau of Statistics Indonesia
Chakravarty, S. R., & Majumder, A. (2005). Measuring Human Poverty: A Generalized Index and an Application Using Basic Dimensions of Life and Some Anthropometric Indicators. *Journal of human development, 6*(3), 275-299. doi:10.1080/14649880500287605
Chambers, R. (1994). The origins and practice of participatory rural appraisal. *World Development, 22*(7), 953-969. doi:https://doi.org/10.1016/0305-750X(94)90141-4
Chambers, R. (2013). Participation, pluralism and perceptions of poverty. In N. Kakwani & J. Silber (Eds.), *The Many Dimensions of Poverty* (pp. 140-164): Springer.
Chaplin, S. E. (1999). Cities, sewers and poverty: India's politics of sanitation. *Environment and urbanization, 11*(1), 145-158.
Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*(6), 429-444. doi:https://doi.org/10.1016/0377-2217(78)90138-8
Chen, K.-M., Leu, C.-H., & Wang, T.-M. (2019). Measurement and determinants of multidimensional poverty: Evidence from Taiwan. *Social Indicators Research, 145*(2), 459-478.
Cherchye, L. (2001). Using data envelopment analysis to assess macroeconomic policy performance. *Applied Economics, 33*(3), 407-416. doi:10.1080/00036840122353
Cherchye, L., & Kuosmanen, T. (2004). *Benchmarking sustainable development: a synthetic meta-index approach* (9291906158). Retrieved from
Cherchye, L., Moesen, W., Rogge, N., & Van Puyenbroeck, T. (2007). An introduction to 'benefit of the doubt'composite indicators. *Social Indicators Research, 82*(1), 111-145.

Cherchye, L., Moesen, W., Rogge, N., Van Puyenbroeck, T., Saisana, M., Saltelli, A., . . . Tarantola, S. (2008). Creating Composite Indicators with DEA and Robustness Analysis: The Case of the Technology Achievement Index. *The Journal of the Operational Research Society, 59*(2), 239-251. Retrieved from http://ezproxy-prd.bodleian.ox.ac.uk:2084/stable/30132778

Cherchye, L., Moesen, W., & Van Puyenbroeck, T. (2004). Legitimately diverse, yet comparable: on synthesizing social inclusion performance in the EU. *JCMS: Journal of Common Market Studies, 42*(5), 919-955.

Chiaperro-Martinetti, E. (1994). A new approach to evaluation of well-being and poverty by fuzzy set theory. *Giornale degli economisti e annali di economia, 50*(No. 7/9), 367-388. Retrieved from https://www.jstor.org/stable/23247764

Chiappero-Martinetti, E., & von Jacobi, N. (2012). Light and Shade of Multidimensional Indexes. In *Quality of life in Italy* (pp. 69-103): Springer.

Chinn, P. L., & Kramer, M. K. (1983). Theory and nursing a systematic approach.

Chowdhury, S., & Squire, L. (2006). Setting weights for aggregate indices: An application to the commitment to development index and human development index. *Journal of Development Studies, 42*(5), 761-771. doi:10.1080/00220380600741904

Clark, D. A. (2003). Concepts and perceptions of human well-being: some evidence from South Africa. *Oxford Development Studies, 31*(2), 173-196.

Coase, R. H. (1960). The problem of social cost. In *Classic Papers in Natural Resource Economics* (pp. 87-137): Springer.

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., & Battese, G. E. (2005). *An introduction to efficiency and productivity analysis*: Springer Science & Business Media.

Cohen, A., & Saisana, M. (2014). Quantifying the Qualitative: Eliciting Expert Input to Develop the Multidimensional Poverty Assessment Tool. *The Journal of Development Studies, 50*(1), 35-50. doi:10.1080/00220388.2013.849336

Cole, F. L. (1988). Content analysis: Process and application. *Clinical Nurse Specialist, 2*(1), 53-57.

Cooke, B., & Kothari, U. (2001). *Participation: The new tyranny?* : Zed books.

Cooper, W. W., Seiford, L. M., Thanassoulis, E., & Zanakis, S. H. (2004). DEA and its uses in different countries. *European Journal of Operational Research, 154*(2), 337-344. doi:https://doi.org/10.1016/S0377-2217(03)00173-5

Cornia, G. A., & Stewart, F. (1993). TWO ERRORS OF TARGETING. *Journal of International Development, 5*(5), 459-496. doi:10.1002/jid.3380050503

Cornwall, A. (2008). Unpacking 'Participation': models, meanings and practices. *Community Development Journal, 43*(3), 269-283. doi:10.1093/cdj/bsn010

Couper, M. R. (1984). The Delphi technique: characteristics and sequence model. *ANS Adv Nurs Sci, 7*. doi:10.1097/00012272-198410000-00008

Crisp, J., Pelletier, D., Duffield, C., Adams, A., & Nagy, S. (1997). The Delphi method? *Nurs Res, 46*. doi:10.1097/00006199-199703000-00010

Critcher, C., & Gladstone, B. (1998). UTILIZING THE DELPHI TECHNIQUE IN POLICY DISCUSSION: A CASE STUDY OF A PRIVATIZED UTILITY IN BRITAIN. *Public Administration, 76*(3), 431-449. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=1206331&site=ehost-live&authtype=ip,uid

Crocker-Buque, T., Mindra, G., Duncan, R., & Mounier-Jack, S. (2017). Immunization, urbanization and slums – a systematic review of factors and interventions. *BMC Public Health, 17*(1), 556. doi:10.1186/s12889-017-4473-7

Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process*: Sage.

Cuevas, S., Rosario, A., Barcenas, M. L., & Christian, M. (2009). Informal employment in Indonesia. *Asian Development Bank Economics Working Paper Series*(156).
Custer, R. L., Scarcella, J. A., & Stewart, B. R. (1999). The modified Delphi technique-A rotational modification.
Dajani, J. S., Sincoff, M. Z., & Talley, W. K. (1979). Stability and agreement criteria for the termination of Delphi studies. *Technological Forecasting and Social Change, 13*(1), 83-90. doi:http://dx.doi.org/10.1016/0040-1625(79)90007-6
Dalkey, N., & Helmer, O. (1963). An Experimental Application of the Delphi Method to the Use of Experts. *Management science, 9*(3), 458-467. Retrieved from http://www.jstor.org/stable/2627117
Dalkey, N. C. (2002). VI. B. A Delphi Study of Factors Affecting the Quality of Life.
Dalkey, N. C., Brown, B. B., & Cochran, S. (1969). *The Delphi method: An experimental study of group opinion* (Vol. 3): Rand Corporation Santa Monica, CA.
Dalkey, N. C., Rourke, D., Lewis, R., & Snyder, D. (1972). Massachusetts: Lexington Books.
Dalkey, N. C., & Rourke, D. L. (1971). Experimental Assessment of Delphi Procedures with Group Value Judgments.
Daly, A., & Fane, G. (2002). Anti-Poverty Programs in Indonesia. *Bulletin of Indonesian Economic Studies, 38*(3), 309-329. doi:10.1080/00074910215535
Dartanto, T., & Nurkholis. (2013). The determinants of poverty dynamics in Indonesia: evidence from panel data. *Bulletin of Indonesian Economic Studies, 49*(1), 61-84.
Datt, G., & Jolliffe, D. (2005). Poverty in Egypt: Modeling and Policy Simulations. *Economic Development and Cultural Change, 53*(2), 327-346. doi:10.1086/425224
Day, J., & Bobeva, M. (2005). A generic toolkit for the successful management of Delphi studies. *The Electronic Journal of Business Research Methodology, 3*(2), 103-116.
de Graaf-Zijl, M., & Nolan, B. (2011). Household joblessness and its impact on poverty and deprivation in Europe. *Journal of European Social Policy, 21*(5), 413-431.
de Villiers, M. R., de Villiers, P. J. T., & Kent, A. P. (2005). The Delphi technique in health sciences education research. *Medical Teacher, 27*(7), 639-643. doi:10.1080/13611260500069947
Deaton, A., & Paxson, C. (1998). Economies of Scale, Household Size, and the Demand for Food. *Journal of Political Economy, 106*(5), 897. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=1213199&site=ehost-live
Decancq, K., & Lugo, M. A. (2008). Setting weights in multidimensional indices of well-being. *Oxford Poverty & Human Development Initiative OPHI Working Paper 18*.
Decancq, K., & Lugo, M. A. (2013). Weights in Multidimensional Indices of Wellbeing: An Overview AU - Decancq, Koen. *Econometric Reviews, 32*(1), 7-34. doi:10.1080/07474938.2012.690641
Delbecq, A. L., Van de Ven, A. H., & Guftafson, D. H. (1974). *A Guide to Nominal Group Techniques for Program Planning and Delphi Processes*. Glenview, Illinois: Scott Forman and Co.
Delbecq, A. L., Van de Ven, A. H., & Gustafson, D. H. (1975). *Group techniques for program planning: A guide to nominal group and Delphi processes*: Scott Foresman.
Delp, P., Thesen, A., Motiwalla, J., & Seshardi, N. (1977). Nominal group technique. *Systems tools for project planning*, 14-18.
Despotis, D. K. (2005). A Reassessment of the Human Development Index Via Data Envelopment Analysis. *The Journal of the Operational Research Society, 56*(8), 969-980. Retrieved from http://ezproxy-prd.bodleian.ox.ac.uk:2084/stable/4102069
Dey, I. (2003). *Qualitative data analysis: A user friendly guide for social scientists*: Routledge.
Dietz, T. (1987). Methods for analyzing data from Delphi panels: Some evidence from a forecasting study. *Technological Forecasting and Social Change, 31*(1), 79-85.

DIVA-GIS. (2012). Free Country Level Spatial Data. Retrieved from http://www.diva-gis.org/gdata
Djamaluddin, S. (2014). *Analysis of Poverty Determinant in West Java Province*: Institute of Economics and Social Research, Faculty of Economics ….
Donohoe, H. M., & Needham, R. D. (2009). Moving best practice forward: Delphi characteristics, advantages, potential problems, and solutions. *International Journal of Tourism Research, 11*(5), 415-437.
Douglass, M. (1992). The political economy of urban poverty and environmental management in Asia: access, empowerment and community based alternatives. *Environment and urbanization, 4*(2), 9-32.
Drèze, J., & Sen, A. K. (2002). *India: Development and participation*: Oxford University Press.
Duclos, J.-Y., Sahn, D. E., & Younger, S. D. (2006). Robust Multidimensional Poverty Comparisons. *Economic Journal, 116*(514), 943-968. doi:10.1111/j.1468-0297.2006.01118.x
Duffield, C. (1993). The Delphi technique: a comparison of results obtained using two expert panels. *International Journal of Nursing Studies, 30*(3), 227-237. doi:https://doi.org/10.1016/0020-7489(93)90033-Q
Elizalde, M. d. R. C., Cáceres, F. A. C., Latapí, A. E., Sittón, S. N., Andretta, J. S., & Belismelis, G. M. T. (2010). *Methodology for Multidimensional Poverty Measurement in Mexico*. Mexico
Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing, 62*(1), 107-115.
English, J. M., & Kernan, G. L. (1976). The prediction of air travel and aircraft technology to the year 2000 using the Delphi method. *Transportation Research, 10*(1), 1-8. doi:http://dx.doi.org/10.1016/0041-1647(76)90094-0
Erb, M., Faucher, C., & Sulistiyanto, P. (2013). *Regionalism in Post-Suharto Indonesia*: Routledge.
Espinoza-Delgado, J., & Klasen, S. (2018). Gender and multidimensional poverty in Nicaragua: An individual based approach. *World Development, 110*, 466-491.
Everitt, B. (2011). *Cluster analysis* (5th ed. ed.). Chichester: Wiley.
Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society. Series A (General), 120*(3), 253-290. doi:10.2307/2343100
Faulkner, B., & Valerio, P. (1995). An integrative approach to tourism demand forecasting. *Tourism Management, 16*(1), 29-37.
Ferreira, F. H., & Lugo, M. A. (2013). Multidimensional poverty analysis: Looking for a middle ground. *The World Bank Research Observer*, lks013.
Ferreira, F. H. G., & Lugo, M. A. (2013). Multidimensional Poverty Analysis: Looking for a Middle Ground. *The World Bank Research Observer The World Bank Research Observer, 28*(2), 220-235.
Fink, A., Kosecoff, J., Chassin, M., & Brook, R. H. (1984). Consensus Methods: Characteristics and Guidelines for Use. *American Journal of Public Health, 74*(9), 979-983. doi:10.2105/AJPH.74.9.979
Forming Clusters. (2020). Retrieved from https://hlab.stanford.edu/brian/forming_clusters.htm
Foster, J., Greer, J., & Thorbecke, E. (1984a). A class of decomposable poverty measures. *Econometrica: Journal of the Econometric Society*, 761-766.
Foster, J., Greer, J., & Thorbecke, E. (1984b). A Class of Decomposable Poverty Measures. *Econometrica, 52*(3), 761-766. doi:10.2307/1913475
Foster, J., Greer, J., & Thorbecke, E. (2010). The Foster–Greer–Thorbecke (FGT) poverty measures: 25 years later. *The Journal of Economic Inequality, 8*(4), 491-524. doi:10.1007/s10888-010-9136-1

Foster, J. E., McGillivray, M., & Seth, S. (2012). *Rank robustness of composite indices: Dominance and ambiguity*. OPHI Working Paper. OPHI. Oxford.
Foster, J. E., & Shorrocks, A. F. (1988). Poverty orderings. *Econometrica: Journal of the Econometric Society*, 173-177.
Foster, J. E., & Shorrocks, A. F. (1988). Poverty orderings and welfare dominance. In *Distributive Justice and Inequality* (pp. 91-110): Springer.
Galili, T. (2015). dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics, 31*(22), 3718-3720.
Gang, I. N., Sen, K., & Yun, M.-S. (2002). Caste, ethnicity and poverty in rural India. Retrieved from https://www.econstor.eu/bitstream/10419/21039/1/dp629.pdf
Gibson, J. M. E. (1998). Using the Delphi technique to identify the content and context of nurses' continuing professional development needs. *Journal of Clinical Nursing, 7*(5), 451-459. doi:doi:10.1046/j.1365-2702.1998.00175.x
Goldschmidt, C. (1987). *Economic evaluations of unpaid household work: Africa, Asia, latin America and Oceania*.
Goldstein, K. (2002). Getting in the door: Sampling and completing elite interviews. *Political Science & Politics, 35*(04), 669-672.
Goodman, C. M. (1987). The Delphi technique: a critique. *Journal of Advanced Nursing, 12*(6), 729-734.
Gordon, D. M. (1972). Theories of poverty and underemployment.
Gordon, T., & Pease, A. (2006). RT Delphi: an efficient,“round-less” almost real time Delphi method. *Technological Forecasting and Social Change, 73*(4), 321-333.
Gordon, T. J. (2003). The delphi method. In J.C. Glenn & T. J. Gordon (Eds.), *Futures research methodology Version 2.0* (Vol. 2). Washington: American Council for the United Nations University.
Gordon, T. J. (2009). The real-time Delphi method. *Futures Research Methodology Version, 3*.
Gosakti, K. (2020). What now? Seeking vaccines in light of the 2016 scandal and religious opposition. *The Jakarta Post*. Retrieved from https://www.thejakartapost.com/academia/2020/01/22/what-now-seeking-vaccines-in-light-of-the-2016-scandal-and-religious-opposition.html
Gregg, P., Harkness, S., & Machin, S. (1999). Poor kids: trends in child poverty in Britain, 1968–96. *Fiscal Studies, 20*(2), 163-187.
Gregg, P., & Wadsworth, J. (2001). Everything you ever wanted to know about measuring worklessness and polarization at the household level but were afraid to ask. *Oxford Bulletin of Economics & Statistics, 63*, 777-806. doi:10.1111/1468-0084.63.spe1.9
Gregg, P., & Wadsworth, J. (2011). *The labour market in winter: the state of working Britain*: Oxford University Press.
Griffin, K., Khan, A. R., & Ickowitz, A. (2002). Poverty and the Distribution of Land. *Journal of Agrarian change, 2*(3), 279-330.
Grundy, J., Wang, X., Hirabayashi, K. C., Duncan, R., Bersonda, D., Eltayeb, A. O., . . . Nandy, R. (2019). Health and immunisation services for the urban poor in selected countries of Asia. *Infectious Diseases of Poverty, 8*(1), 26. doi:10.1186/s40249-019-0538-4
Grusky, D. B., & Kanbur, R. (2006). Introduction: The conceptual foundations of poverty and inequality measurement. *Poverty and inequality*, 1-29.
Grusky, D. B., Kanbur, S. R., & Sen, A. K. (2006). *Poverty and inequality*: Stanford University Press.
Guio, A.-C., Fusco, A., & Marlier, E. (2009). A European Union approach to material deprivation using EU-SILC and Eurobarometer data. *Integrated Research Infrastructure in the Socio-economic Sciences (IRISS) Working Paper Series, 19*, 2009.
Guio, A.-C., Gordon, D., & Marlier, E. (2012a). Measuring material deprivation in the EU: Indicators for the whole population and child-specific indicators. *Eurostat: Luxembourg*.

Guio, A.-C., Gordon, D., & Marlier, E. (2012b). Measuring material deprivation in the EU: Indicators for the whole population and child-specific indicators. *Publications Office of the European Union: Luxembourg, EU*.
Guio, A.-C., Gordon, D., Najera, H., & Pomati, M. (2017). Revising the EU material deprivation variables 2017 Edition.
Guio, A.-C., Marlier, E., Gordon, D., Fahmy, E., Nandy, S., & Pomati, M. (2016). Improving the measurement of material deprivation at the European Union level. *Journal of European Social Policy, 26*(3), 219-333. doi:doi:10.1177/0958928716642947
Hadiz, V. R. (2010). *Localising power in post-authoritarian Indonesia: A Southeast Asia perspective*: Stanford University Press Stanford.
Hall, S., Leary, K., & Greevy, H. (2014). *Public Attitudes to Poverty*: Joseph Rowntree Foundation York.
Hallowell, M. R., & Gambatese, J. A. (2009). Qualitative research: Application of the Delphi method to CEM research. *Journal of construction engineering and management, 136*(1), 99-107.
Hallowell, M. R., & Gambatese, J. A. (2010). Qualitative Research: Application of the Delphi Method to CEM Research. *Journal of construction engineering and management, 136*(1), 99-107. doi:10.1061/(ASCE)CO.1943-7862.0000137
Hamilton, L. C. (2012). *Statistics with Stata: version 12*: Cengage Learning.
Hanafin, S. (2004). Review of literature on the Delphi Technique. *Dublin: National Children's Office*.
Hanafin, S., Brooks, A.-M., Carroll, E., Fitzgerald, E., GaBhainn, S. N., & Sixsmith, J. (2007). Achieving consensus in developing a national set of child well-being indicators. *Social Indicators Research, 80*(1), 79-104.
Hanafin, S., xe, ad, Brooks, A.-M., Carroll, E., Fitzgerald, E., . . . Sixsmith, J. (2007). Achieving Consensus in Developing a National Set of Child Well-Being Indicators. *Social Indicators Research, 80*(1), 79-104. Retrieved from http://www.jstor.org/stable/27734527
Hanandita, W., & Tampubolon, G. (2015). Multidimensional Poverty in Indonesia: Trend Over the Last Decade (2003–2013). *Social Indicators Research*. doi:10.1007/s11205-015-1044-0
Hardjo, H. (2010). *Susenas 2011*.
Hardjo, S. H. (2010). *SUSENAS 2011: Pelatihan Calon Innas Susenas dan Sakernas 2011*. Jakarta Retrieved from https://slideplayer.info/slide/2989710/
Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing, 32*(4), 1008-1015. doi:10.1046/j.1365-2648.2000.t01-1-01567.x
Head of the Bogor City District Development Body Statistics Division. (2016) *On The Measurement of Multidimensional Poverty/Interviewer: P. Natih*. Bogor City Delphi (Vol 2).
Hendarman, Hadiat, & Aruan, C. D. (2018). *Indonesia SDGs 4 Monitoring: Challenges and Needs*. Retrieved from Mexico: http://tcg.uis.unesco.org/wp-content/uploads/sites/4/2019/01/05-SDGs-4-Statistics-Indonesia-V3.pdf
Hitch, P., & Murgatroyd, J. (1983). Professional communications in cancer care: a Delphi survey of hospital nurses. *Journal of Advanced Nursing, 8*(5), 413-422.
Holtzappel, C. J., & Ramstedt, M. (2009). *Decentralization and regional autonomy in Indonesia: implementation and challenges*: Institute of Southeast Asian Studies.
Hsu, C.-C., & Sandford, B. A. (2007). The Delphi technique: making sense of consensus. *Practical Assessment, Research & Evaluation, 12*(10), 1-8.
Hwang, G.-J. (2011). *New welfare states in East Asia: Global challenges and restructuring*: Edward Elgar Publishing.

Indonesia Demographic and Health Survey 2017. (2018). Retrieved from Jakarta: https://www.dhsprogram.com/pubs/pdf/FR342/FR342.pdf
International Labour Office (ILO). (1976). *Employment, Growth and Basic Needs: A One-World Problem*. Retrieved from Geneva:
Iriana, N., Avenzora, A., & Abidin, J. (2012). *Indonesia: Poverty and Data Governance*. Paper presented at the Kunming, Peoples' Republic of China October 23-25. http://www.adbi.org/files/2012.10.25.cpp.sess5.10.indonesia.poverty.data.gov.pdf
Jacobs, R., Smith, P. C., & Goddard, M. K. (2004). *Measuring performance: an examination of composite performance indicators: a report for the Department of Health*: Centre of Health Economics, University of York.
Jalan, J., & Ravallion, M. (2000). Is transient poverty different? Evidence for rural China. *The Journal of Development Studies, 36*(6), 82-99.
Johar, M., Soewondo, P., Pujisubekti, R., Satrio, H. K., & Adji, A. (2019). In Data We Trust? An Analysis of Indonesian Socioeconomic Survey Data. *Bulletin of Indonesian Economic Studies, 55*(1), 61-82. doi:10.1080/00074918.2018.1515474
Jones, G. W., & Pratomo, D. (2016). Education in Indonesia: Trends, Differentials, and Implications for Development. In *Contemporary Demographic Transformations in China, India and Indonesia* (pp. 195-214): Springer.
Jones, J., & Hunter, D. (1995). Consensus Methods For Medical And Health Services Research. *BMJ: British Medical Journal, 311*(7001), 376-380. Retrieved from http://ezproxy-prd.bodleian.ox.ac.uk:2084/stable/29728325
Jordan, E., & Javernick-Will, A. (2013). Indicators of Community Recovery: Content Analysis and Delphi Approach. *Natural Hazards Review, 14*(1), 21-28. doi:doi:10.1061/(ASCE)NH.1527-6996.0000087
Joshi, D., Fawcett, B., & Mannan, F. (2011). Health, hygiene and appropriate sanitation: experiences and perceptions of the urban poor. *Environment and urbanization, 23*(1), 91-111.
Kakwani, N., & Silber, J. (2007). Introduction. In N. Kakwani & J. Silber (Eds.), *The Many Dimensions of Poverty*: Palgrave Macmillan.
Kakwani, N., & Silber, J. (2008). Introduction: Multidimensional Poverty Analysis: Conceptual Issues, Empirical Illustrations and Policy Implications. *World Development, 36*(6), 987-991. doi:https://doi.org/10.1016/j.worlddev.2007.10.004
Kanagaratnam, U. (2017). *Macro, Household and Individual Level Explanations for the Geographic Differences in Poverty Levels in Indonesia, 2000–2009.* (DPhil Monograph). University of Oxford, Oxford.
Kassambara, A. (2017). *Practical Guide to Cluster Analysis in R: Unsupervised Machine Learning* (Vol. 1): STHDA.
Kaynak, E., Bloom, J., & Leibold, M. (1994). Using the Delphi technique to predict future tourism potential. *Marketing Intelligence & Planning, 12*(7), 18-29.
Keeney, S., Hasson, F., & McKenna, H. (2011). Analysing data from a Delphi and reporting results. *The Delphi Technique in Nursing and Health Research*, 84-95.
Keeney, S., Hasson, F., & McKenna, H. P. (2001). A critical review of the Delphi technique as a research methodology for nursing. *International Journal of Nursing Studies, 38*(2), 195-200.
Keeney, S., McKenna, H., & Hasson, F. (2010). *The Delphi technique in nursing and health research*: John Wiley & Sons.
Kerlinger, F. N. (1973). Foundation of behavioral science. *New York: Holt, Renehard and Winston*.
Ketkar, N. (2017). *Deep Learning with Python*: Springer.
Klasen, S. (2000). Measuring poverty and deprivation in South Africa. *Review of Income and Wealth, 46*(1), 33-58.

Korhonen, P., Tainio, R., & Wallenius, J. (2001). Value efficiency analysis of academic research. *European Journal of Operational Research, 130*(1), 121-132.
Kuusi, O. (1999). *Expertise in the future use of generic technologies*: Valtion taloudellinen tutkimuskeskus VATT.
Laderchi, C. R., Saith, R., & Stewart, F. (2003). 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.
Landeta, J. (2006). Current validity of the Delphi method in social sciences. *Technological Forecasting and Social Change, 73*(5), 467-482. doi:https://doi.org/10.1016/j.techfore.2005.09.002
Langfelder, P., Zhang, B., & Horvath, S. (2008). Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics, 24*(5), 719-720.
Lanjouw, P., & Murgai, R. (2009). Poverty decline, agricultural wages, and nonfarm employment in rural India: 1983–2004. *Agricultural Economics, 40*(2), 243-263. doi:10.1111/j.1574-0862.2009.00373.x
Lanjouw, P., Pradhan, M., Saadah, F., Sayed, H., & Sparrow, R. (2001). *Poverty, education, and health in Indonesia: who benefits from public spending?* : The World Bank.
Lanjouw, P., & Ravallion, M. (1995). Poverty and Household Size. *The Economic Journal, 105*(433), 1415-1434. doi:10.2307/2235108
Layte, R., Nolan, B., & Whelan, C. T. (2001). Reassessing income and deprivation approaches to the measurement of poverty in the Republic of Ireland. *Economic and Social Review, 32*(3), 239-262.
Leaker, D. (2009). Economic inactivity. *Economic & Labour Market Review, 3*(2), 42-46. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=36604179&site=ehost-live&authtype=ip,uid
Lemmer, B. (1998). Successive surveys of an expert panel: research in decision-making with health visitors. *Journal of Advanced Nursing, 27*(3), 538-545.
Lerman, R., & McKernan, S.-M. (2008). *The effects of holding assets on social and economic outcomes of families: A review of theory and evidence*. Retrieved from Washington DC:
Lewis, B. D. (2014). Urbanization and economic growth in Indonesia: good news, bad news and (possible) local government mitigation. *Regional Studies, 48*(1), 192-207.
Lewis-Beck, M., Bryman, A. E., & Liao, T. F. (2003). *The Sage encyclopedia of social science research methods*: Sage Publications.
Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology, 49*(4), 764-766. doi:https://doi.org/10.1016/j.jesp.2013.03.013
Licona, G. H., Díaz, F. A., Cáceres, F. A. C., Latapí, A. E., Enciso, M. G. F., Dommarco, J. Á. R., & Belismelis, G. M. T. (2010). *Methodology for Multidimensional Poverty Measurement in Mexico*.
Lijphart, A. (1971). Comparative politics and the comparative method. *American political science review, 65*(03), 682-693.
Linstone, H. A., & Turoff, M. (1975). *The delphi method*: Addison-Wesley Reading, MA.
Linstone, H. A., & Turoff, M. (1977). *The Delphi method : techniques and applications*. Reading, Mass: Addison-Wesley Pub. Co., Advanced Book Program.
Linstone, H. A., & Turoff, M. (2002). The Delphi Method. *Techniques and applications, 53*.
Lipton, M., & Ravallion, M. (1995). Chapter 41 Poverty and policy. In *Handbook of development economics* (Vol. Volume 3, Part B, pp. 2551-2657): Elsevier.
Liu, K. (2016). Insuring against health shocks: Health insurance and household choices. *Journal of Health Economics, 46*, 16-32. doi:https://doi.org/10.1016/j.jhealeco.2016.01.002

Lovell, C. K., Grosskopf, S., Ley, E., Pastor, J. T., Prior, D., & Eeckaut, P. V. (1994). Linear programming approaches to the measurement and analysis of productive efficiency. *Top, 2*(2), 175-248.
Lüthi, C., McConville, J., & Kvarnström, E. (2010). Community-based approaches for addressing the urban sanitation challenges. *International Journal of Urban Sustainable Development, 1*(1-2), 49-63.
Lydall, H. (1968). *The structure of earnings*: Clarendon Press.
Mack, J., & Lansley, S. (1985). *Poor Britain*: Allen & Unwin London.
MacNealy, M. S. (1999). Strategies for Empirical Research in Writing. In: New York: Longman.
Mahlberg, B., & Obersteiner, M. (2001). *Remeasuring the HDI by data envelopement analysis*. Retrieved from Laxemburg, Austria:
Maksum, C. (2004). *Official poverty measurement in Indonesia.* Paper presented at the international conference on official poverty statistics.
Matell, M. S., & Jacoby, J. (1971). Is There an Optimal Number of Alternatives for Likert Scale Items? Study I: Reliability and Validity. *Educational and Psychological Measurement, 31*(3), 657-674. doi:10.1177/001316447103100307
Mayer, S. E., & Jencks, C. (1989). Poverty and the distribution of material hardship. *Journal of Human Resources*, 88-114.
McFarlane, C., Desai, R., & Graham, S. (2014). Informal Urban Sanitation: Everyday Life, Poverty, and Comparison. *Annals of the Association of American Geographers, 104*(5), 989-1011. doi:10.1080/00045608.2014.923718
McKenna, H. P. (1994). The Delphi technique: a worthwhile research approach for nursing? *Journal of Advanced Nursing, 19*(6), 1221-1225.
Meinzen-Dick, R. (2009). *Property rights for poverty reduction?* . Retrieved from New York:
Melyn, W., & Moesen, W. (1991). Towards a synthetic indicator of macroeconomic performance: unequal weighting when limited information is available. *Public economics research papers*, 1-24.
Micklewright, J., & Stewart, K. (1999). Is the well-being of children converging in the European Union? *The Economic Journal, 109*(459), 692-714.
Millar, J., & Glendinning, C. (2009). Gender and Poverty. *Journal of Social Policy, 18*(3), 363-381. doi:10.1017/S004727940001761X
Milne, P. (2006). *Making the New Indonesia Work for the Poor*: World Bank.
Ministry of Education and Culture Republic of Indonesia. (2019). Data on Elementary and Secondary Education: Number of School Data. Retrieved from http://dapo.dikdasmen.kemdikbud.go.id/sp/2/286100
Miranti, R. (2010). Poverty in Indonesia 1984-2002: The impact of growth and changes in inequality. *Bulletin of Indonesian Economic Studies, 46*(1), 79-97. doi:10.1080/00074911003642252
Montgomery, M. (2009). Urban poverty and health in developing countries. *Population Bulletin, 64*(2).
Morduch, J. (2006). Concepts of Poverty. In *Handbook on poverty statistics: Concepts, methods and policy use* (pp. 23-50): United Nations.
Mukherjee, N. (1999). Consultations with the Poor in Indonesia. *World Bank, Washington, DC. Available online at<http://www.worldbank.org/poverty/voices/reports.htm#national*.
Mukherjee, N. (2006). *Voices of the Poor: Making Services Work for the Poor in Indonesia: a Qualitative Consultation with the Poor at Eight Sites*: World Bank Office Jakarta (East Asia and Pacific Region).
Munier, F., & Ronde, P. (2001). The role of knowledge codification in the emergence of consensus under uncertainty: empirical analysis and policy implications. *Research Policy, 30*(9), 1537-1551.

Murphy, M. K., Black, N. A., Lamping, D. L., McKee, C. M., Sanderson, C. F., Askham, J., & Marteau, T. (1998). Consensus development methods, and their use in clinical guideline development. *Health Technol Assess, 2*.

Murray, W. F., & Jarman, B. O. (1987). Predicting future trends in adult fitness using the Delphi approach. *Research Quarterly for Exercise and Sport, 58*(2), 124-131.

Narayan, D., Patel, R., Schafft, K., Rademacher, A., & Koch-Schulte, S. (1999). Voices of the poor: Can anyone hear us? *World Bank, Washington DC*.

Narayan, D., & Petesch, P. (2002). *Voices of the Poor: From Many Lands* (0821350498). Retrieved from

Nardo, M., Saisana, M., Saltelli, A., & Tarantola, S. (2005). Tools for composite indicators building. European Commission Joint Research Centre. *Institute for the Protection and the Security of the Citizen, Econometrics and Statistical Support to Antifraud Unit, I-21020 Ispra (VA) Italy, Report number: EUR, 21682*.

Natih, P. (2016a). *Bogor City Delphi: Grouping Indicators into Dimensions Exercise*.

Natih, P. (2016b). *Bogor City Delphi: Iterative Questionnaire Results*. Bogor City Delphi Online Questionnaire.

Natih, P. (2020, 13 August 2020). How digital technology can reduce inequality. *The Jakarta Post*. Retrieved from https://www.thejakartapost.com/academia/2020/08/13/how-digital-technology-can-reduce-inequality.html

Niazi, T. H. (2012). Decentralization. In *Diagnosing the Indonesian Economy: Toward Inclusive and Green Growth* (pp. 389-428): Anthem Press.

Nolan, B., & Whelan, C. T. (1996). Resources, deprivation, and poverty. *OUP Catalogue*.

Nolan, B., & Whelan, C. T. (2010). Using Non-Monetary Deprivation Indicators to Analyze Poverty and Social Exclusion: Lessons from Europe? *Journal of Policy Analysis and Management, 29*(2), 305-325. Retrieved from http://www.jstor.org/stable/20685185

Nordstokke, D., & Stelnicki, A. M. (2014). Pairwise Comparisons. In A. C. Michalos (Ed.), *Encyclopedia of Quality of Life and Well-Being Research* (pp. 4575-4576). Dordrecht: Springer Netherlands.

Norton, A., & Stephens, T. (1995). Particlpation in Poverty Assessments.

Nussbaum, M., & Sen, A. (1993). *The quality of life*: Oxford University Press.

Nussbaum, M. C. (2011). *Creating capabilities*: Harvard University Press.

OECD. (2018). Education GPS Indonesia Student Performance (PISA 2015). Retrieved from http://gpseducation.oecd.org/CountryProfile?primaryCountry=IDN&treshold=10&topic=PI

Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: an example, design considerations and applications. *Information & management, 42*(1), 15-29.

Oxfamblogs (Producer). (2010). Guest Blog: World Bank research director critiques the new UN poverty index.

Oxford Poverty and Human Development Initiative (OPHI). (2016a). MPI Study for Indonesia Presented in Jakarta. Retrieved from https://ophi.org.uk/mpi-study-for-indonesia-presented-in-jakarta/

Oxford Poverty and Human Development Initiative (OPHI). (2016b). *Multidimensional Poverty and its Measurement.* Paper presented at the United Nations Economic Commission for Europe Conference of European Statistics, Geneva, Switzerland.

Oxford Poverty and Human Development Initiative (OPHI). (2017). *OPHI Country Briefing June 2017: Indonesia*. Retrieved from www.dataforall.org/dashboard/ophi/index.php/mpi/download_brief_files/IDN

Pelčić, G., Karačić, S., Mikirtichan, G. L., Kubar, O. I., Leavitt, F. J., Cheng-Tek Tai, M., . . . Tomašević, L. (2016). Religious exception for vaccination or religious excuses for avoiding vaccination. *Croatian medical journal, 57*(5), 516-521. doi:10.3325/cmj.2016.57.516

Piachaud, D. (1987). Problems in the definition and measurement of poverty. *Journal of Social Policy, 16*(02), 147-164.
Pietersma, S., de Vries, M., & van den Akker-van, M. E. (2014). Domains of quality of life: results of a three-stage Delphi consensus procedure among patients, family of patients, clinicians, scientists and the general public. *Quality of Life Research, 23*(5), 1543-1556.
Pigou, A. C. (1929). *The economics of welfare*: Macmillan and Company, limited.
Piketty, T., & Atkinson, A. B. (2010). *Top incomes: a global perspective*: Oxford University Press.
PISA 2015 Country Overview of Indonesia. (2015). Retrieved from http://www.compareyourcountry.org/pisa/country/IDN?lg=en
Platt, L. (2010). Ten year transitions in children's experience of living in a workless household: variations by ethnic group. *Population trends, 139*(1), 70-90.
Portela, M. C. A. S., Thanassoulis, E., & Simpson, G. (2004). Negative Data in DEA: A Directional Distance Approach Applied to Bank Branches. *The Journal of the Operational Research Society, 55*(10), 1111-1121. Retrieved from http://ezproxy-prd.bodleian.ox.ac.uk:2084/stable/4101957
Pradhan, M., Suryahadi, A., Sumarto, S., & Pritchett, L. (2000). *Measurements of poverty in Indonesia, 1996, 1999, and beyond*: The World Bank.
Prakarsa. (2015). MPI Indonesia Teknis Perhitungan. Retrieved from http://mpi-indonesia.org/teknis-perhitungan.html. Retrieved 12 December 2016 http://mpi-indonesia.org/teknis-perhitungan.html
Prakarsa. (2016). *Calculation: Multidimensional Poverty in Indonesia 2012-2014*. Retrieved from
Priebe, J. (2010). *Child costs and the causal effect of fertility on female labor supply: an investigation for Indonesia 1993-2008*. Retrieved from Göttingen:
Priebe, J. (2014). Official Poverty Measurement in Indonesia since 1984: A Methodological Review. *Bulletin of Indonesian Economic Studies, 50*(2), 185-205. doi:10.1080/00074918.2014.938406
Pritchett, L., Suryahadi, A., & Sumarto, S. (2000). *Quantifying vulnerability to poverty: A proposed measure, applied to Indonesia*: World Bank Publications.
Public Relations Office Government of Bogor City. (2017). DPRD dan Pemkot Bogor Bahas Perda Penataan Pemukiman Kumuh [Press release]. Retrieved from http://www.jabarprov.go.id/index.php/news/23899/2017/07/06/DPRD-dan-Pemkot-Bogor-Bahas-Perda-Penataan-Pemukiman-Kumuh
Qizilbash, M. (2002). A note on the measurement of poverty and vulnerability in the South African context. *Journal of International Development, 14*(6), 757-772. doi:10.1002/jid.922
Qizilbash, M. (2003). Vague language and precise measurement: the case of poverty. *Journal of economic methodology, 10*(1), 41-58.
Qizilbash, M. (2006). Philosophical accounts of vagueness, fuzzy poverty measures and multidimensionality. In *Fuzzy set approach to multidimensional poverty measurement* (pp. 9-28): Springer.
Raschka, S. (2015). *Python machine learning*: Packt Publishing Ltd.
Rauch, W. (1979). The decision Delphi. *Technological Forecasting and Social Change, 15*(3), 159-169. doi:http://dx.doi.org/10.1016/0040-1625(79)90011-8
Ravallion, M. (1993). *Poverty alleviation through regional targeting: a case study for Indonesia*. Retrieved from Washington D.C.:
Ravallion, M. (1994). *Poverty comparisons* (Vol. 56): Taylor & Francis.
Ravallion, M. (1997). Good and bad growth: The human development reports. *World Development, 25*(5), 631-638. doi:https://doi.org/10.1016/S0305-750X(96)00141-6
Ravallion, M. (2001). Poverty Lines: Economic Foundations of Current Practices. *Unpublished manuscript. The World Bank, Washington, DC*.

Ravallion, M. (2003). The Debate on Globalization, Poverty and Inequality: Why Measurement Matters. *International Affairs, 79*(4), 739-753. doi:10.1111/1468-2346.00334
Ravallion, M. (2010). Poverty lines across the world. *World Bank Policy Research Working Paper Series, Vol*.
Ravallion, M. (2011a). Mashup Indices of Development. *The World Bank Research Observer, 27*(1), 1-32.
Ravallion, M. (2011b). On multidimensional indices of poverty. *The Journal of Economic Inequality, 9*(2), 235-248.
Ravallion, M. (2012). Troubling tradeoffs in the Human Development Index. *Journal of Development Economics, 99*(2), 201-209. doi:https://doi.org/10.1016/j.jdeveco.2012.01.003
Ravallion, M., & Bidani, B. (1994). How robust is a poverty profile? *World Bank Economic Review, 8*(1), 75-102. doi:10.1093/wber/8.1.75
Ravallion, M., Chen, S., & Sangraula, P. (2009). Dollar a day revisited. *The World Bank Economic Review*, lhp007.
Ravallion, M., & Datt, G. (1995). Is Targeting through a Work Requirement Efficient? In D. van de Walle & Nead (Eds.), (pp. 413-444).
Ravallion, M., Datt, G., & Walle, D. (1991). Quantifying absolute poverty in the developing world. *Review of Income and Wealth, 37*(4), 345-361.
Ravallion, M., & Jalan, J. (1998). *Determinants of transient and chronic poverty: evidence from rural China*: The World Bank.
Republic of Indonesia Coordinating Ministry of Administrative and Bureaucratic Reform. (2016). Peneliti Oxford University Lakukan Penelitian Pengentasan Kemiskinan di Kota Bogor. Retrieved from https://www.menpan.go.id/site/berita-terkini/berita-daerah/oxford-university-lakukan-penelitian-pemetaan-kemiskinan-di-kota-bogor
Republic of Indonesia Law Number 3 Year 2014 on Total Coverage of Sanitation for Communities, (2014).
Republic of Indonesia Ministry of Education and Culture. (2003). *Law on the National Education System Number 20 Year 2003*. Jakarta
Law on Compulsory Education 47/2008, (2008).
Republic of Indonesia Ministry of Education and Culture. (2014a). *Law on National Standards for Early Childhood Education 137/2014*. Jakarta Retrieved from https://www.sekolah-id.com/permendikbud-nomor-137-tahun-2014-tentang-standar-nasional-paud/
Republic of Indonesia Ministry of Education and Culture. (2014b). *Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 146 Tahun 2014 Tentang Kurikulum 2013 Pendidikan Anak Usia Dini*. Jakarta
Republic of Indonesia Ministry of Education and Culture. (2015). *Rencana Strategis Kementrian Pendidikan dan Kebudayaan Tahun 2015-2019*. Jakarta
Republic of Indonesia Ministry of Education and Culture. (2016). *Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 19 Tahun 2016 Tentang Program Indonesia Pintar*. Jakarta
Petunjuk Teknis Bantuan Inisiasi Penyelenggaraan Pendidikan Usia Dini 0-3 Tahun Pada Tahun 2018, (2018).
Peraturan Menteri Kesehatan Republik Indonesia 97/2014 tentang Pelayanan Kesehatan Masa Sebelum Hamil, Masa Hamil, Persalinan dan Masa Sesudah Melahirkan, serta Pelayanan Kesehatan Seksual, (2014).
Republic of Indonesia Ministry of Health. (2016). *Immunization Situation in Indonesia*.
Republic of Indonesia Ministry of Health. (2018). *Profil Kesehatan Indonesia Tahun 2017*. Retrieved from Jakarta:
Tujuan Menurunkan Angka Kematian Anak, (2004).
Faktor-Faktor yang Mempengaruhi Kelangsungan Hidup Anak, (2009).

Republic of Indonesia Ministry of National Development Planning. (2019). *Rencana Kerja Pemerintah Tahun 2019: Pemerataan Pembangunan dan Pertumbuhan Berkualitas*. Jakarta
Republic of Indonesia National Constitution. (1945). *Undang-Undang Dasar Republik Indonesia 1945*. Indonesia Retrieved from http://www.bpk.go.id/assets/files/storage/2013/12/file_storage_1386157387.pdf
Republic of Indonesia National Law Number 2 Year 2002 on Policing and Safety (2002).
Republic of Indonesia National Law Number 23 Year 2006 on Citizenship. (2006).
Republic of Indonesia National Law Number 36 Year 2009 on Health. (2009). Jakarta
Rice, D. P. (1967). Estimating the cost of illness. *American journal of public health and the nation's health, 57*(3), 424-440. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/6066903
https://www.ncbi.nlm.nih.gov/pmc/PMC1227175/
Ringen, S. (1988). Direct and Indirect Measures of Poverty. *Journal of Social Policy, 17*(3), 351-365. doi:10.1017/S0047279400016858
Robeyns, I. (2005a). The Capability Approach: a theoretical survey. *Journal of human development, 6*(1), 93-117. doi:10.1080/146498805200034266
Robeyns, I. (2005b). Selecting capabilities for quality of life measurement. *Social Indicators Research, 74*(1), 191-215.
Robeyns, I. (2006). The capability approach in practice. *Journal of Political Philosophy, 14*(3), 351-376.
Rogan, M. (2016). Gender and multidimensional poverty in South Africa: Applying the global multidimensional poverty index (MPI). *Social Indicators Research, 126*(3), 987-1006.
Rossouw, A., Hacker, M., & de Vries, M. J. (2011). Concepts and contexts in engineering and technology education: An international and interdisciplinary Delphi study. *International Journal of Technology and Design Education, 21*(4), 409-424.
Rowe, C., & Broadie, S. (Eds.). (2020). *Aristotle: Nicomachean Ethics*. Oxford: Oxford University Press.
Rowe, G., & Wright, G. (1999). The Delphi technique as a forecasting tool: issues and analysis. *International journal of forecasting, 15*(4), 353-375.
Rowe, G., & Wright, G. (2001). Expert opinions in forecasting: the role of the Delphi technique. In *Principles of forecasting* (pp. 125-144): Springer.
Rowe, G., Wright, G., & McColl, A. (2005). Judgment change during Delphi-like procedures: The role of majority influence, expertise, and confidence. *Technological Forecasting and Social Change, 72*(4), 377-399. doi:https://doi.org/10.1016/j.techfore.2004.03.004
Rowntree, B. S. (1901). *Poverty: a study of town life*: Macmillan.
Rustiadi, E., & Nasution, A. (2017). Can Social Capital Investment Reduce Poverty in Rural Indonesia? *International Journal of Economics and Financial Issues, 7*(2), 109.
Sahn, D. E., & Stifel, D. C. (2000). Poverty Comparisons Over Time and Across Countries in Africa. *World Development, 28*(12), 2123-2155. doi:https://doi.org/10.1016/S0305-750X(00)00075-9
Saisana, M., Saltelli, A., & Tarantola, S. (2005). Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 168*(2), 307-323. doi:10.1111/j.1467-985X.2005.00350.x
Saisana, M., & Tarantola, S. (2002). *State-of-the-art report on current methodologies and practices for composite indicator development*: Citeseer.
Sajogyo. (1977). *Garis Kemiskinan dan Kebutuhan Minimum Pangan* Bogor Agricultural University. Bogor.

Saltelli, A. (2007). Composite Indicators: Between Analysis and Advocacy. *Social Indicators Research, 81*(1), 65-77. Retrieved from http://ezproxy-prd.bodleian.ox.ac.uk:2084/stable/20734414

Sandrey, M. A., & Bulger, S. M. (2008). The Delphi method: an approach for facilitating evidence based practice in athletic training. *Athletic Training Education Journal, 3*(4), 135-142.

Santos, M. E., & Alkire, S. (2011). Training material for producing national human development reports. *MPI: Construction and analysis. Oxford: Oxford Poverty and Human Development Initiative*.

Scheibe, M., Skutsch, M., & Schofer, J. (2002). IV. C. Experiments in Delphi methodology. *In: Linstone H, Turoff M (eds) The Delphi method: techniques and applications*, 257-281.

Schuler, R. S. Nominal Group Technique. *Wiley Encyclopedia of Management*.

Schütze, H., Manning, C. D., & Raghavan, P. (2008). *Introduction to information retrieval* (Vol. 39): Cambridge University Press Cambridge.

Sen, A. (1976a). Poverty: An Ordinal Approach to Measurement. *Econometrica, 44*(2), 219-231. doi:10.2307/1912718

Sen, A. (1976b). Poverty: an ordinal approach to measurement. *Econometrica: Journal of the Econometric Society*, 219-231.

Sen, A. (1979a). Equality of what? The Tanner Lecture on Human Values, 22 May. In: Stanford University.

Sen, A. (1979b). Issues in the Measurement of Poverty. *The Scandinavian Journal of Economics*, 285-307.

Sen, A. (1985). A sociological approach to the measurement of poverty: a reply to Professor Peter Townsend. *Oxford economic papers*, 669-676.

Sen, A. (2001). *Development as freedom*: Oxford University Press.

Sen, A. (2004a). Capabilities, Lists, and Public Reason: Continuing the Conversation. *Feminist economics, 10*(3), 77-80. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=15059158&site=ehost-live

Sen, A. (2004b). *Rationality and freedom*: Harvard University Press.

Sen, A. (2011). *The idea of justice*: Harvard University Press.

Sen, A., & Foster, J. E. (1997). Annexe. In *On Economic Inequality*. Oxford: Clarendon Press.

Shah, H. A., & Kalaian, S. A. (2009). Which Is the Best Parametric Statistical Method For Analyzing Delphi Data? *Journal of Modern Applied Statistical Methods, 8*(1), 20.

Silver, C., Azis, I. J., & Schroeder, L. (2001). Intergovernmental transfers and decentralisation in Indonesia. *Bulletin of Indonesian Economic Studies, 37*(3), 345-362.

Sitlington, H., & Coetzer, A. (2013). Using the Delphi Technique to Identify Components of a Tertiary Strategic HRM Curriculum. *eCULTURE, 5*(1), 1.

Skulmoski, G. J., Hartman, F. T., & Krahn, J. (2007). The Delphi method for graduate research. *Journal of Information Technology Education: Research, 6*, 1-21.

Starkweather, D. B., Gelwicks, L., & Newcomer, R. (1975). Delphi Forecasting of Health Care Organization. *Inquiry, 12*(1), 37-46. Retrieved from www.jstor.org/stable/29770903

Stiglitz, J. E., Sen, A., & Fitoussi, J.-P. (2010a). *Mismeasuring our lives: Why GDP doesn't add up*: The New Press.

Stiglitz, J. E., Sen, A., & Fitoussi, J.-P. (2010b). Report by the commission on the measurement of economic performance and social progress. *Paris: Commission on the Measurement of Economic Performance and Social Progress*.

Strauss, H. J., & Zeigler, L. H. (1975a). The Delphi technique and its uses in social science research. *The Journal of Creative Behavior, 9*(4), 253-259.

Strauss, H. J., & Zeigler, L. H. (1975b). Delphi, political philosophy and the future. *Futures, 7*(3), 184-196.

Streeten, P. (1994). Human development: means and ends. *the american Economic review*, 232-237.
Streeten, P., Burki, S. J., Haq, U., Hicks, N., & Stewart, F. (1981). First things first: meeting basic human needs in the developing countries.
Sugiarto, D. B. A. (2016, 20 July 2016) *Poverty Alleviation Policies in Bogor City/Interviewer: P. G. L. Natih*. Delphi Workshop in Bogor City (Vol 1), Bogor City.
Sullivan, G. M., & Artino, A. R. (2013). Analyzing and Interpreting Data From Likert-Type Scales. *Journal of Graduate Medical Education, 5*(4), 541-542. doi:10.4300/JGME-5-4-18
Sumarto, S., & de Silva, I. (2014). Beyond the Headcount: Examining the Dynamics and Patterns of Multidimensional Poverty in Indonesia.
Sumsion, T. (1998). The Delphi technique: an adaptive research tool. *British Journal of Occupational Therapy, 61*(4), 153-156.
Surbakti, P. (1995). Indonesia's National Socio-Economic Survey. *Central Bureau of Statistics, Jakarta-Indonesia*.
Suryadarma, D., & Jones, G. W. (2013). *Education in Indonesia*: Institute of Southeast Asian Studies.
Suryadarma, D., Suryahadi, A., & Sumarto, S. (2007a). Measuring Unemployment in Developing Countries: The Case of Indonesia. *LABOUR: Review of Labour Economics & Industrial Relations, 21*(3), 541-562. doi:10.1111/j.1467-9914.2007.00374.x
Suryadarma, D., Suryahadi, A., & Sumarto, S. (2007b). *Reducing unemployment in Indonesia: Results from a growth-employment elasticity model*: SMERU Research Institute.
Suryahadi, A., & Sumarto, S. (2010). Poverty and Vulnerability in Indonesia before and after the Economic Crisis. *Poverty and Social Protection in Indonesia*, 36.
Suryahadi, A., Sumarto, S., & Pritchett, L. (2003). Evolution of poverty during the crisis in Indonesia. *Asian Economic Journal, 17*(3), 221-241. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-0142151533&partnerID=40&md5=9ec455b953f09ee894bddc15186a67c1
Suryahadi, A., Suryadarma, D., & Sumarto, S. (2006). *Economic growth and poverty reduction in Indonesia: The effects of location and sectoral components*. Retrieved from
Syebubakar, A. (2014). Income poverty is only part of the story. *The Jakarta Post*. Retrieved from https://www.thejakartapost.com/news/2014/10/14/income-poverty-only-part-story.html
Symonds, P. M. (1924). On the Loss of Reliability in Ratings Due to Coarseness of the Scale. *Journal of Experimental Psychology, 7*(6), 456.
Tarricone, R. (2006). Cost-of-illness analysis: What room in health economics? *Health Policy, 77*(1), 51-63. doi:https://doi.org/10.1016/j.healthpol.2005.07.016
Tehusijarana, K. M. (2019). Not even mediocre? Indonesian students score low in math, reading, science: PISA report. *The Jakarta Post*. Retrieved from https://www.thejakartapost.com/news/2019/12/04/not-even-mediocre-indonesian-students-score-low-in-math-reading-science-pisa-report.html
Thanassoulis, E. (2001). *Introduction to the theory and application of data envelopment analysis*: Springer.
The Australia Indonesia Partnership for Economic Governance (AIPEG). (2017). *Women's Economic Participation in Indonesia: A study of gender inequality in employment, entrepreneurship, and key*
enablers for change. Retrieved from
The Lancet. (2020). Redefining vulnerability in the era of COVID-19. *The Lancet, 395*(10230), 1089. doi:10.1016/S0140-6736(20)30757-1
Thorbecke, E. (2007). Multidimensional poverty: Conceptual and measurement. In N. Kakwani & J. Silber (Eds.), *The Many Dimensions of Poverty*: Palgrave Macmillan.

Townsend, P. (1979a). *Poverty in the United Kingdom : a survey of household resources and standards of living*. Harmondsworth: Penguin Books.
Townsend, P. (1979b). *Poverty in the United Kingdom: a survey of household resources and standards of living*: Univ of California Press.
Townsend, P. (1985). A Sociological Approach to the Measurement of Poverty--A Rejoinder to Professor Amartya Sen. *Oxford economic papers*, 659-668.
Tran, V. Q., Alkire, S., & Klasen, S. (2015). Static and dynamic disparities between monetary and multidimensional poverty measurement: Evidence from Vietnam. *OPHI Working Paper 97*.
Trinity College. (2018). Graduate Guest Lecture with the Mayor of Bogor City. Retrieved from https://www.trinity.ox.ac.uk/guest-lecture/
Triyono, B., Raharjo, H., Faiq;, Pratiwi, M. I., Ayu, N. M., Rahmawaty, A. N., & Nuryawani, T. P. (2014). *Kumpulan Ringkasan Kajian dan Evaluasi Sektoral 2008-2013*. Jakarta: Badan Perencanaan Pembangunan Nasional (BAPPENAS)
Triyono, M. B., & Mateeke Moses, K. (2019). Technical and Vocational Education and Training and Training in Indonesia. In B. Bai & Paryono (Eds.), *Vocational Education and Training in ASEAN Member States: Current Status and Future Development* (pp. 45-79). Singapore: Springer Singapore.
Tsui, K. Y. (2002). Multidimensional poverty indices. *Social Choice and Welfare, 19*(1), 69-93.
Turoff, M. (1970). The design of a policy Delphi. *Technological Forecasting and Social Change, 2*(2), 149-171. doi:http://dx.doi.org/10.1016/0040-1625(70)90161-7
Turoff, M. (1975). The Policy Delphi. In H. A. Linstone & M. Turoff (Eds.), *The Delphi method: Techniques and applications*.
Ulschak, F. L. (1983). *Human resource development: The theory and practice of need assessment*. Reston, VA: Reston Publishing Company.
UN General Assembly. (1948). Universal declaration of human rights. *UN General Assembly*.
UNESCO. (2017). Official Entrance Age by Level of Education. Retrieved from http://data.uis.unesco.org/index.aspx?queryid=216. from UNESCO Institute for Statistics http://data.uis.unesco.org/index.aspx?queryid=216
UNICEF. (2007). *A human rights-based approach to education for all: A framework for the realization of children's right to education and rights within education*: United Nations Publications.
UNICEF. (2012). Immunization: Keeping Children Alive and Healthy. Retrieved from https://www.unicef.org/immunization/files/Immunization_brochure.pdf
UNICEF. (2013). *Every Child's Birth Right: Inequities and Trends in Birth Registration*. Retrieved from New York:
United Nations. (2014). Crime is both cause, consequence of poverty, third committee told as it begins discussion of crime prevention, international drug control [Press release]
United Nations. (2016). *Sustainable Development Goals Report 2016*: UN.
United Nations (UN). (2015). The Human Right to Water and Sanitation Media Brief [Press release]
United Nations Children's Fund (UNICEF). (2020). *The State of Children in Indonesia: Trends, Opportunities and Challenges for Realizing Children's Rights*. Retrieved from
United Nations Department of Public Information. (2009). *Millennium Development Goals Report 2009*: United Nations Publications.
United Nations Development Programme (UNDP). (2015). MPI Statistical Programmes. Retrieved from http://hdr.undp.org/en/content/mpi-statistical-programmes
United Nations Population Fund (UNFPA). (2010). *The Rights to Contraceptive Information and Services for Women and Adolescents*. Retrieved from New York:
Ura, K., Alkire, S., Zangmo, T., & Wangdi, K. (2012). *An extensive analysis of GNH index*.

Usman, S. (2002). *Regional Autonomy in Indonesia: Field Experiences and Emerging Challenges*. Paper presented at the 7th PRSCO Summer Institute/ The 4th IRSA International Conference: "Decentralization, Natural Resources and Regional Development in the Pacific Rim", Bali, Indonesia.

Van de Ven, A. H., & Delbecq, A. L. (1974). The effectiveness of nominal, Delphi, and interacting group decision making processes. *Academy of management Journal, 17*(4), 605-621.

von der Gracht, H. A. (2012). Consensus measurement in Delphi studies: Review and implications for future quality assurance. *Technological Forecasting and Social Change, 79*(8), 1525-1536. doi:http://dx.doi.org/10.1016/j.techfore.2012.04.013

Waldron, J. S. (1971). An investigation into the relationships among conceptual level, time delay of information feedback, and performance in the Delphi process.

Walker, A., & Selfe, J. (1996). The Delphi method: a useful tool for the allied health researcher. *British Journal of Therapy and Rehabilitation, 3*(12), 677-681.

Wang, C. C., Wang, Y., Zhang, K., Fang, J., Liu, W., Luo, S., . . . Li, V. C. (2003). Reproductive health indicators for China's rural areas. *Social Science & Medicine, 57*(2), 217-225. doi:http://dx.doi.org/10.1016/S0277-9536(02)00341-6

Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association, 58*(301), 236-244. doi:10.2307/2282967

Weaver, W. T. (1971). The Delphi Forecasting Method. *The Phi Delta Kappan, 52*(5), 267-271. Retrieved from http://ezproxy-prd.bodleian.ox.ac.uk:2084/stable/20372868

Whelan, C. T., & Maître, B. (2007). Measuring material deprivation with EU-SILC: Lessons from the Irish survey. *European Societies, 9*(2), 147-173.

Whelan, C. T., Nolan, B., & Maître, B. (2014). Multidimensional poverty measurement in Europe: An application of the adjusted headcount approach. *Journal of European Social Policy, 24*(2), 183-197.

Whelan, C. T., Watson, D., & Maître, B. (2019). From Income Poverty to Multidimensional Quality of Life. *The Economic and Social Review, 50*(4), 683-705.

Whitman, N. I. (1990). The Delphi technique as an alternative for committee meetings. *Journal of Nursing Education, 29*(8), 377-379.

Williams, P. L., & Webb, C. (1994). The Delphi technique: Methodological discussion. *J Adv Nurs, 19*. doi:10.1111/j.1365-2648.1994.tb01066.x

Witkin, B. R., & Altschuld, J. W. (1995). *Planning and conducting needs assessments: A practical guide*: Sage.

World Bank. (1990). *World Development Report: Poverty,* . New York: Oxford University Press.

World Bank. (2005). *Introduction to Poverty Analysis*: World Bank Institute.

World Bank. (2015). INDO-DAPOER Data Bank for Indonesia. Retrieved from http://databank.worldbank.org/data/reports.aspx?source=indo~dapoer-(indonesia-database-for-policy-and-economic-research)#

World Bank. (2016). *Indonesia - National Slum Upgrading Project (Project Appraisal Document No. 1477)*. Retrieved from Washington DC: http://documents.worldbank.org/curated/en/226091468417040569/Indonesia-National-Slum-Upgrading-Project

World Health Organization (WHO). (2004). *The Sanitation Challenge: Turning Commitment into Reality*. Retrieved from Switzerland:

World Health Organization (WHO). (2009). *WHO guide to identifying the economic consequences of disease and injury* (9241598298). Retrieved from

World Health Organization (WHO). (2013). *Self care for health: a handbook for community health workers & volunteers*. Retrieved from India:

World Health Organization (WHO). (2018a). *Guidelines on sanitation and health*. Retrieved from Switzerland:

World Health Organization (WHO). (2018b). Immunization. Retrieved from http://www.who.int/topics/immunization/en/

World Health Organization (WHO) and the United Nations Children's Fund (UNICEF). (2017). *Progress on drinking water, sanitation and hygiene: 2017 update and SDG baselines*. Retrieved from Switzerland:

Yalonetzky, G. (2011). A Note on the Standard Errors of the Members of the Alkire Foster Family and its Components. *OPHI Research in Progress, 25*.

Yang, Y. N. (2003). Testing the Stability of Experts' Opinions between Successive Rounds of Delphi Studies.

Yi, J., Heckman, J. J., Zhang, J., & Conti, G. (2015). Early Health Shocks, Intra-household Resource Allocation and Child Outcomes. *The Economic Journal, 125*(588), F347-F371.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*(3), 338-353. doi:https://doi.org/10.1016/S0019-9958(65)90241-X