Operationalizing Amartya Sen's Capability Approach through Capability-Equivalent Income

By

Yuan Cheng[1], John K. Dagsvik[2], Xuehui Han[3]

Abstract:
Economic conditions that have a strong impact on individual well-being and freedom are sometimes not reflected in income. Amartya Sen proposed the Capability Approach to capture the welfare impact of such non-income economic condition. However, empirical applications of the capability approach are commonly expressed in reduced form which is not suitable for conducting welfare analysis wherein the distribution of preferences and choice constraints across households are separately identified and accounted for. Using the random utility theory, we incorporate capability into the job market choice set, as well as to the utility through disposable income and leisure. In this approach, we can transform capability to the traditional concept of income, which yields the equivalent utility or welfare as if there were no capability constraints. We deliberately chose Cambodia for the empirical analysis, because the civil war and genocide that occurred in the 1970s in the country can be considered a "natural shock". These events have resulted in a high proportion of war-related disabilities, which can be considered as a source of capability constraint. By incorporating disability-associated constraints on the job choices and the disability-induced utility difference with respect to income and time spent on leisure, we show that capability disadvantages can be converted to a form of income through the compensation variation, which we denote as capability-equivalent income. We simulate a transfer program where disabled individuals are compensated with capability-equivalent income through an income tax to the top 20% earners. We find that inequality would be mitigated, and the overall welfare of the whole society would improve by 4%.

Keywords: Capability approach, Discrete choice, Welfare function, Inequality
JEL classification: C25, C35, D31, D63

---

[1] Population Research Institute, Fudan University, 220 Handan Road, Shanghai, PRC, 200433. Email: chengyuan@fudan.edu.cn
[2] Research Department, Statistics Norway, Oslo, Norway. Email: john.dagsvik@ssb.no
[3] Economic Research and Regional Cooperation Department, Asian Development Bank, Metro Manila, 1550, Philippines. Email: xuehuihan@adb.org

Acknowledgement: The views expressed in this paper are those of the authors and do not necessarily reflect the views and policies of the Asian Development Bank, its Board of Governors, or the governments they represent. The authors are grateful for the comments from the participants of an ADB seminar and ASSA 2018 annual conference and for the excellent research assistance by Ma. Christina F. Epetia and Orlee P. Velarde. The authors are responsible for any remaining errors in the paper.

1. Introduction

Sen has underlined in several of his publications that economic inequality is not necessarily the same as income inequality. He said in the opening statement of his article entitled "From Income Inequality to Economic Inequality" published in 1997:

Focus must be shifted from income inequality to economic inequality because of the presence of causal influences on individual well-being and freedom that are economic in nature but cannot be expounded by simple statistics of incomes and commodity holdings. Attention must be given to heterogeneous magnitudes.

Economic conditions that have a strong impact on individual well-being and freedom are sometimes not reflected in income. Sen proposed the Capability Approach to capture the welfare impact of such non-income economic conditions. The capability approach has been developed and discussed in a series of papers and books like Sen (1979, 1985a, 1985b, 1992, 1997), Drèze and Sen (2002), Robeyns (2003, 2006), and Robeyns and Kuklys (2005).

Many studies based on Sen's capability approach present typical reduced-form analyses, such as Anand, Hunter, and Smith (2005); Anand, Santos, and Smith (2008); Kuklys (2005); and Anand et al. (2011), to name a few. Using reduced-form analyses implies that they were not able to establish welfare measures that are based on an explicit representation of the distribution of household preferences and capability sets. Even Sen himself has not discussed the challenge of developing an empirical strategy that is based on a structural quantitative model. He also seems to be sceptical of using conventional economic choice theory to develop a methodological framework for generating welfare measures in the context of his capability approach. Sen stated on page 5 of his 1997 paper:

...for many purposes, the appropriate space is neither that of utilities (as claimed by welfarists) nor that of primary goods (as demanded by Rawls). If the object is to concentrate on the individual's real opportunity to pursue her objectives, then account would have to be taken not only of the primary goods the person holds but also of the relevant personal characteristics that govern the conversion of primary goods into the person's ability to promote her ends. For example, a person who is disabled may have a larger basket of primary goods and yet have less chance to lead a normal life (or to pursue her objectives) than an able-bodied person with a smaller basket of primary goods....

This may be because the typical welfare implications that stem from conventional economic choice theory is *utilitarian*. Conventional textbook theory usually deals with unconstrained choice (subject to the budget constraint), where preferences are deterministic and exhibit perfectly transitive choice behaviour. Evidently, conventional unconstrained choice theory is hardly relevant for the purposes of the capability theory. In principle, constrained choice theory can offer an alternative methodology, because it allows the researcher to include restrictions like the ones represented by the capability set. However, in contrast to the very simple cases, constrained (or deterministic) choice theory is extremely complicated and impractical (Kuklys, 2005).

To fill the gap, Dagsvik (2013) developed a structural empirical framework to construct welfare measures based on the capability approach. Dagsvik (2013) argued that in Sen's framework, capabilities (or the capability set) represent the opportunities to achieve functionings. He used the labor market as an example, saying that while working is a functioning, the opportunity to work is an element of a person's capability set. In Sen's original terminology, the concept of capability (or capability set) is equivalent to the set of functionings available to the individual (choice set). Robeyns (2003) and Robeyns and Kuklys (2005) pointed out that the concept of capability has been used in different ways in the theoretical literature. However, according to Dagsvik (2013)'s further reading of Sen's work, the capabilities as freedoms refer to the presence of valuable options or alternatives in the sense of opportunities that exist not only formally or legally but also effectively, because they are available to the individual. In turn, the capabilities will be measured and reflected as the availability of job choices in our analysis. Building on the past literature, we extend the framework to accommodate the capability of individuals to enjoy consumption and leisure time. Individuals with disability not only face job opportunity constraints but also rely more on consumption and demand for more leisure time; that is, consumption and leisure time are more valuable to individuals with disability.

By applying the methodology of probabilistic choice theory (McFadden, 1973) and Dagsvik (2013), we generate practical welfare measures that explicitly accommodate observed and latent choice constraints, like reduced job opportunities caused by disability. This approach deviates from the utilitarian method, because the degree of freedom of the households, which is represented by household-specific choice constraints defining the capability sets, is accounted for. We compute welfare measures which consider the fact that disabled individuals have fewer opportunities in the labor market and rely more on consumption and leisure compared to people who do not have disability. Disability is only one of the many potential capability constraints and far from representative of the scope of Sen's Capability Approach. However, it is still useful in illustrating how to measure capability and translate it to the traditional, easy-to-understand concept of income and welfare.

We use a sample of individuals obtained from the Cambodia Socio-Economic Survey (CSES) 2009. Cambodia experienced a civil war and genocide in the 1970s, which caused a surge in the proportion of people with disability. A major implication of the surge is that disability in Cambodia can more likely be considered exogenous compared to the common cases of disability. The detailed causes of disability are disclosed in the CSES and war is explicitly listed as one of the options. Utilizing the available data in the Cambodian survey, we develop a measurable labor supply model which explicitly incorporates the job choice constraints due to disability. Furthermore, we translate the job choice constraints (capability limit) to a money metric measure, which we denote as the capability-equivalent income. Thus, the advantage of this method is that it helps visualize the abstract concept of welfare loss associated with capability limit in monetary terms. We simulate the impact of compensating for the lost welfare due to capability limit through the capability-equivalent income which, in turn, is fully financed by income tax. In particular, we find that inequality in both income and welfare would be mitigated significantly if people with disability were compensated with the amount of capability-equivalent income that corresponds to the constraints they face by taxing the top 20% eaners. The overall welfare level would increase by 4% with no impact on the balance of the government budget. To our best knowledge, we are among the pioneers who have applied a structural quantitative model to assess the welfare impact in Sen's capability framework. As demonstrated in our income distribution simulation, the straightforward money metric measure produced with this method can be easily communicated to stakeholders and the net welfare gain from the proposed transfer program might be appealing to policymakers.

The paper is organized as follows. In the next section, we briefly introduce the background on Cambodia's civil war and genocide in 1970s. Section 3 presents the model for labor supply behaviour. Section 4 discusses the data and shows the empirical estimation results. Section 5 explains how the capability-equivalent income is calculated. Section 6 presents the simulation of a transfer program by compensating individuals with disability and taxing the 20% earners to finance the compensation scheme and evaluate the impact of the program on income and welfare. Section 7 concludes the paper.

2. Cambodia Civil War/Genocide and Disability Exogeneity[4]

After being freed from the French empire in 1953, Cambodia started as a new and independent country. However, Cambodia immediately experienced economic difficulties, especially the lack of food for peasants, and became involved in the neighbouring Viet Nam's civil war, triggering heated conflicts between pro- and anti- communist forces. A civil war then broke out in Cambodia at the end of 1960s and it lasted until 1975. Right after the Civil War, the so-called Pol Pot's Regime (1975 -1979) ruled the country, during which a genocide was carried out and caused massive deaths and injuries. Heuveline (2015) estimated that the genocide resulted in a median value of 1.9 million excess deaths.

The unique history of Cambodia makes it an ideal case to study disability as an outcome of external shocks rather than reckless individual behaviour which is more common. The injuries and disabilities caused by the war, such as injuries by bombing, generated a disproportionally high share of people with disability in Cambodia, providing an exogenous case for analysing the capability-induced loss of opportunity and distorted dependence on non-labor income and demand for leisure.

Cambodia has available data on the specific details of individual disability. In the Cambodia Socio-Economic Survey (CSES) 2009, there are questions on the type of difficulties or disabilities and the corresponding causes of such disabilities, like "mental trauma due to war" and "war injuries".[5] As a follow-up question, the respondents were asked whether their disabilities have prevented them from participating in employment and income-generating activities. Among the 633 individuals who indicated that their disabilities have prevented them from participating in employment and income-generating activities, 272 individuals have pure exogenous causes for their disabilities, including congenital, difficulty

[4] I have talked with one friend who has personally experienced the Cambodian Civil War about the exogeneity of disability in Cambodia. In addition to war, she has mentioned that there are other conditions causing the exogeneity of disability in Cambodia. For example, she mentioned one case like one fly fled into someone's left eye, which should be easy to cure. However, because of lack of medical staffs and skills, the right eye of the person was affected soon and the person became blind. Other cases like people who injured their leg by small accidence can easily develop into serious condition and have to loss the whole leg.

[5] Section 14 on page 48 of the CSES 2009 questionnaire. Refer to Appendix 2 for the scanned copy of this section.

in delivery, mental trauma due to war, war injuries, and old age, which accounts for around 43% of the respondents. We will discuss the detailed summary statistics in Section 5.

3. Model of labor market behaviour

3.1. *The Utility Model*

We shall now present the model of individual behaviour in the labor market. We refer to Dagsvik and Strøm (2006) for a more rigorous discussion and Dagsvik et al. (2013) for a survey of this type of models. This model significantly departs from the traditional approaches in the literature, since job choice becomes the fundamental decision variable and it extends the conventional discrete choice approach to labor supply models as proposed by van Soest (1995).

In our model, the agent derives utility from household consumption C (here it is set to equal household disposable income), leisure h (represented by working hours), and non-pecuniary latent attributes of jobs. The utility of agent i is assumed to have the form:

$$U^{D_i}(C_i, h_i, j, z) = v^{D_i}(C_i, h_i)\varepsilon_i(j, z), \tag{3.1}$$

where (C_i, h_i) denote disposable income and monthly hours of work of agent i, respectively, and the term $v(\cdot)$ is a suitable positive deterministic function. The term $\{\varepsilon_i(j, z)\}$ is positive sector j- and job z-specific random taste shifters for agent i, which account for unobserved individual characteristics and unobserved job-specific attributes. The taste shifters in $\{\varepsilon_i(j, z)\}$ are assumed to be i.i.d. across jobs z and agents i, with the extreme value distribution of c.d.f. $exp(-1/x), x > 0$. This distribution function follows the Independence of Irrelevant Attributes assumption (IIA). The reason why the index z enters the utility function is that job-specific attributes beyond the attributes of sector j, such as wage and hours of work, may affect the utility of the agents.

We first divide jobs into sectors because there are large differences in the nature of jobs across sectors, making the comparison of jobs with similar wage and working hours across sectors quite different from the comparison of jobs with similar wage and working hours within the same sector. Let $j = 1, 2, 3$ correspond to the index of the sectors of

Agriculture, Manufacturing, and Other industries, respectively, and let $j = 0$ represent the alternative of not working. Let $z = 1, 2, \ldots, Z_j$ be the index of jobs within each sector.

Suppose disposable income C_i has two sources: wage income and non-labor income. Let w_j be the wage the agent faces in sector j. For a given job z, hours of work is expressed by $h_i = H_i(z)$. This means that the agent can only change hours of work by changing jobs; that is, jobs within each sector can be differentiated by their associated hours of working. The real working hour distribution can be represented by an appropriate indexation. For given hours of work h_i and wage w_j, the disposable household income is defined as $C_i = h_i w_j + I_i$, where I_i denotes the non-labor income of agent i.

Let $B_j(h_i)$, which is equivalent to $B_j(z)$ in our context, be the set of jobs in sector j with hours of work h_i that are available to the agent i. Let $\theta(j)$ be the total number of available jobs in sector j and $g_j(h_i)$ be the mean fraction of jobs with hours of work h_i within sector j. Thus, $\theta(j)g_j(h_i)$ refers to the mean number of jobs in $B_j(h_i)$. We call $\theta(j)g_j(h_i)$ the opportunity measure of the agent, which is the choice set of jobs realized by agent i with working hours h_i. Recall that in our context, working hours h_i also represents the types of jobs. Individual characteristics, such as disability, affect the utility function through the available job choices in $B_j(h_i)$ wherein the number of job choices in $B_j(h_i)$ when agent i has disability is smaller than the number of job choices when agent i has no disability.

From the assumptions defined above, it follows immediately from the theory of discrete choice that the probability $\varphi(h_i, j)$ of choosing hours of work h_i in sector j when working in either one of the three sectors such that $h_i > 0$ can be expressed as:

$$(3.2) \qquad \varphi(h_i, j) = P\left(\max_{z \in B_j(h)} U\big(h_i w_j + I_i, h_i, j, z\big) = \right.$$

$$\left. max\left(U(I_i, 0), \max_r \max_{x \in G_i} \max_{z \in B_r(x)} U(x w_r + I_i, x, r, z)\right)\right) = \frac{\theta(j) g_j(h_i) v^{G_i}(h_i w(j) + I_i, h_i)}{v(I_i, 0) + \sum_r \theta(r) \sum_{x \in D_i} g_r(x) v^{G_i}(x w(r) + I_i, x)}.$$

where $\mathrm{w_j} = \{\mathrm{w}(1), w(2), w(3)\}$, $j = 1, 2, 3$ and G_i is the job subset faced by agent i, which, in turn, is determined by agent i's capability. Equation (3.2) implies that choosing the job with hours of work h_i in sector j brings the highest utility to the agent among all the job choices determined by the capability of agent i in each sector and the alternative choice of not working. The triple max in Equation (3.2) illustrates the utility maximizing process as agent i first chooses the best job x in each sector r conditional on the choice set that is constrained by his or her capability, then compares the best job across all sectors, and finally chooses the job that brings the highest utiltiy. In other words, the probability given in Equation (3.2) is the probability that the most preferred job of agent i is a job in sector j with job-specific hours of work equal to h_i among all the jobs in the individual's latent choice set of available jobs.

In the traditional case, as suggested by in Dagsvik and Strøm (2006) and Dagsvik et al. (2013), the probability $\varphi(i, 0,0)$ of not working with h = 0 can be presented as:

$$\varphi(i, 0,0) = \frac{v^{G_i}(I_i,0)}{v(I_i,0)+\sum_r \theta(r) \sum_{x\in D_i} g_r(x) v^{G_i}(xw(r)+I_i,x)}. \tag{3.3}$$

In our analysis, we extend the probability to further allow the utility of consumption and leisure to depend on the individual characteristics of individuals with and without disability. The reason we only introduce disability and not the other individual characteristics to affect the utility with respect to consumption and leisure is to emphasize the impact of the capability factor. Additionally, introducing more individual characteristics might generate identification difficulties. When individual i has no disability, the probability of choosing not to work will take the form shown in Equation (3.3). When an individual has disability, we allow the utility of not working to be different from Equation (3.3) by adding the term $\alpha_0 D_i$ such that:

$$\varphi(i, 0,0) = \frac{v^{G_i}(I_i,0)}{(v(I_i,0)+\alpha_0 D_i)+\sum_r \theta(r) \sum_{x\in G_i} g_r(x) v^{G_i}(xw(r)+I_i,x)}. \tag{3.4}$$

where $D_i = 1$, if individual i has disability and equal to zero if otherwise. The coefficient α_0 represents the difference in the utility of not working between an individual with disability and without disability.

3.2. Empirical specification of the model

In this section, we further define the deterministic part of the utility function as a log-linear function in the form of:

$$(3.5) \qquad logv^{G_i}(C_i, h_i) = \alpha_1 C_i + \alpha_2 h_i;$$

that is, the logarithmic transformation of $v(C_i, h_i)$ is linear in both parameters and variables of disposable income C_i and working hours h_i, which can be easily derived by using the Taylor extension. Examples can be found in Han (2010, p. 360, Equations (3) – (4)). The deterministic part of the utility function given in Equation (3.5) fulfils the requirement of positive deterministic function for $v(C_i, h_i)$. We expect higher disposable income to bring higher utility (a positive estimate of α_1) and more hours of work to reduce utility (a negative estimate of α_2).

We expect consumption and leisure to affect the utility of an individual with disability differently. We revise the utility function in Equation (3.5) to allow individuals with and without disabilities to have different utility with respect to consumption and working hours to:

$$(3.6) \quad logv(C_i, h_i) = (\alpha_1 + \alpha_{1D} * D_i)C_i + (\alpha_2 + \alpha_{2D} * D_i)h_i \,,$$

where coefficients α_{1D} and α_{2D} reflect the difference in utility between individuals with and without disability with respect to consumption and working hours, respectively. We expect individuals with disabilities to demand more consumption and less working hours.

Similar to the deterministic part of the utility function $v(C_i, h_i)$, we express in log-linear form the opportunity sets of jobs faced by agent i if he chooses to work. The opportunity sets can be measured and specified as:

$$(3.7) \qquad log\left(\theta(j)g_j(h_i)\right) = log\theta(j) + logg_j(h_i) = \gamma_{0j} + \gamma_j X_i,$$

where $log\left(\theta(j)g_j(h_i)\right)$ can be decomposed into two parts: $log\theta(j)$ and $logg_j(h_i)$. The first part, $log\theta(j)$, represents the total number of jobs in sector j and does not depend on agent i. Therefore, it can be expressed as a constant term that only varies across sectors and is denoted as γ_{0j}.

In contrast, the second part, $log g_j(h_i)$, varies across sectors and agents. We assume that the interacting terms are independent and can be formulated as $\gamma_j X_i$, where the parameter vector γ_j denotes the variation across sectors and characteristic vector X_i reflects the variation across agents. Conditional on data availability in the household survey, we assume that the characteristic vector X_i consists of the following variables: gender, dummy for speaking more than one language, years of schooling, dummy for living in the current location since birth, dummy for being located in either an urban or rural area, years of work experience, years of work experience squared, and dummy for having disability. All the variables included in X_i are assumed to be affecting the opportunity sets of the agent. Years of work experience is calculated as the difference between an individual's age and years of schooling minus six years of pre-school age. We can include both education and work experience in the model without causing perfect multicollinearity, because we do not include the age of agents in vector X_i.

The job opportunity measure $\left(\gamma_{0j} + \gamma_j X_i\right)$ in Equation (3.7) can be explicitly linked to the number of jobs faced by agent i in sector j by taking the exponential of both sides of Equation (3.7):

$$(3.8) \qquad \widetilde{m}_{ij} = \theta(j) g_j(h_i) = exp\left(\gamma_{0j} + \gamma_j X_i\right),$$

where $\widetilde{m}_{ij}$ can be understood as the number of latent jobs accessible to agent i in sector j.

With the number of latent jobs $\widetilde{m}_{ij}$, Equation (3.8) introduces the impact of capability constraints faced by individuals who have disability through X_i. The probability $\{\varphi(h_i, j)\ \}$ of agent i with hours of work h_i in sector j as shown in Equation (3.2) is affected by disability through two channels: (1) directly through the disposable income and leisure time entering the utility function and (2) indirectly through the size of job choice sets $\widetilde{m}_{ij}$ faced by the agent which is determined by his or her capability (characteristics).

In our modelling framework, disability affects both the demand for and supply of labor: the job opportunities provided by the job market conditional on individual characteristics through $\widetilde{m}_{ij}$ and the direct utility generated by supplying labor through $log v(C_i, h_i)$.

Recall that in Section 3.1, we defined the hourly wage rate for sector j as $w(j)$. To estimate $w(j)$, we attempt first to follow the classic approach in estimating wage equations; that is, estimating wage as a function of individual characteristics. Within each sector, we include individual characteristics vector $\tilde{X}_i$ with gender, language, length of schooling, immigrant indicator, urban/rural indicator, work experience, work experience squared, and disability as explanatory variables for wage:

(3.9) $$w_i(j)=\beta_{0j} + \beta_j \tilde{X}_i.$$

As shown in Table A1 in Appendix 1, only few of the explaining variables included in the vector have significant impact on wage in all three sectors. Only the indicator for living in the current location since birth is significant for the agriculture sector, length of schooling and indicator for living in the current location since birth are significant for manufacturing sector, and length of schooling and indicator for urban/rural are significant for other industries. Moreover, the overall explanatory power of the model is very low with the R-squared being not significantly different from zero: 0.003 for agriculture, 0.008 for manufacturing, and 0.01 for other industries. However, when we pool the three sectors together and include sector dummies, the coefficients of these dummies are significant and the R-squared increased dramatically to 0.42. On the other hand, R-squared remains low at 0.01 if sectors are not differentiated by dummies. We understand that the result implies that in Cambodia the wage level does not depend much on individual characteristics but depend more on which sector the individual choose to work in. The variation of wages is reflected more across sectors rather than within sectors.

The estimation results of the wage equations imply that the sector average wage rate will be at least as good as the wage estimates based on wage equation. Therefore, for the wages, we take the sector average wage and assume that every individual face the same sector average wage rate when they make sector choices; that is, $\mathrm{w_j} = \{\overline{w}_1, \overline{w}_2, \overline{w}_3\},\ j = 1, 2, 3$. The differences in their job choice in each sector can only be reflected in the hours of working h_i.

The final step is to estimate the parameters in the probability function. Based on Equations (3.2) and (3.4), we use the Maximum Likelihood method to estimate the parameters. The log-likelihood function takes the form of:

(3.10) $$\log L = \sum_i \sum_{j=1}^{3} Y_{ij} log\big(\varphi_i(h_i, j)\big) + \sum_i Y_{i0} log\big(\varphi_i(i, 0,0)\big),$$

where $Y_{ij} = 1$ if individual i works in sector j and zero if otherwise, $j = 1,2,3$; $Y_{i0} = 1$ if individual i is not working and zero if otherwise; and h_i is the observed hours of work of individual i. Thus, the first term in Equation (3.10) is the log-likelihood of observing agent i working in sector j and the second term is the log-likelihood of observing agent i not working. We then search the values of the parameters which bring the maximum value of $\log L$ by using the quasi-Newton optimization algorithm.

4. Data and Empirical Results

The data used in the study is extracted from the CSES 2009. The survey sampled 11,971 households and 57,105 individuals. Two sets of questionnaires are included in the survey: Household Questionnaire and Diary Questionnaire. As previously mentioned, we chose the CSES 2009, because the civil war and genocide that occurred in Cambodia make the country a distinct case where disability can be attributed to natural shocks Accordingly, the dataset contains detailed information on individuals' disabilities—for example, the type of disability and the corresponding reasons for such disability—which makes it uniquely suitable for the analytical purposes of this paper. The questionnaire pages concerning disability-related questions are included in Appendix 2. We emphasize, however, that we are not exploring any country-specific research implications for the Cambodia case in this study.

The analysis focuses on people who had evaluated their job market choices and provided complete data on the variables we need in the study. We exclude the following: individuals who are younger than 15 years old or older than 60 years old; individuals who reported that they are students, seeking jobs, or retired; and individuals who are employed but did not report weekly working hours.

Jobs are categorized into sectors that correspond to the nature of their economic activities. In the survey, economic activity conducted by individuals is coded using the International Standard Industrial Classification (ISIC). We group all activities into three broad sectors; namely Agriculture, Forestry, and Fishing (section A in ISIC); Mining and Manufacturing (sections B and C in ISIC); and the remaining sectors (sections D-U in ISIC). For convenience, we will subsequently refer to these sectors as agriculture, manufacturing, and other industries, respectively.

The information on individual characteristics, such as age, gender, education, industry where the person is working, weekly working hours, and monthly salary, are obtained from the household questionnaire. To get the non-labor income of an individual, we add the monthly salaries of all other family members reported in the household questionnaire and the household income from other sources reported in the diary sheet. Individuals who neither have salary nor non-labor income information are excluded. To curb the effect of positive skewness in income, we exclude the 1% of individuals with the highest total income. The remaining sample after cleaning consists of 14,103 individuals with 6,767 males and 7,336 females.

Among the 6,767 males, 453 individuals reported that they have disability in three degrees of severity—mild, moderate, and severe—which accounts for around 6.7% of the males.[6] The detailed information allows us to examine whether the degree of severity affects job opportunities. Among the 453 individuals with disabilities, there are 171, 193, and 89 males with mild, moderate, and severe disability, respectively, which are around 2.5%, 2.9%, and 1.3% of the male sample, respectively. As shown in the questionnaire in Appendix 2, the disability and the degree of severity are based on self-assessment. Taking advantage of the information revealed by the question on whether the disability prevents the respondent from participating in and accessing employment and income generating activities, there are 380 males who indicated that they face such difficulties, which is smaller than the total number of males reporting that they have disability. In the empirical estimations, we will examine whether it is the reported disability or the self-reported harm of job opportunities that have the actual impact.

Among the 7,336 females, 529 individuals reported that they have some degree of disability, which accounts for around 7.2% of the females. There are 214, 213, and 102 females with mild, moderate, and severe disability, respectively, which are around 2.9%, 2.9%, and 1.4% of the female sample, respectively. There are 350 females who reported that their disability has hindered their participation in employment and income generating activities.

We break down the descriptive statistics of the sample by grouping individuals according to gender, severity of disability, and sector. In Table 1, we present four health statuses (mild/moderate/severe disability, and without disability), three sectors for working individuals, and an additional category of not working. The upper panel is for males, while the lower panel is for females.

[6] In the questionnaire, up to three disabilities can be included. We identified the highest level of severity of disability in the reported disability as the degee of severity in the final data.

Table 1. Distribution of sample by disability severity and gender across sectors

Health status	Agriculture	Manufacturing	Other industries	Not working
Males				
Without disability	48.3%	8.5%	23.8%	19.3%
Disability of all degrees	49.2%	10.8%	21.0%	19.0%
Mild disability	49.7%	15.2%	19.9%	15.2%
Moderate disability	49.7%	8.8%	21.2%	20.2%
Severe disability	47.2%	6.7%	22.5%	23.6%
Females				
Without disability	47.7%	9.3%	24.4%	18.5%
Disability of all degrees	44.8%	10.2%	23.6%	21.4%
Mild disability	48.6%	8.4%	25.7%	17.3%
Moderate disability	41.8%	10.8%	21.6%	25.8%
Severe disability	43.1%	12.7%	23.5%	20.6%

As shown in Table 1, among those who are not working, the proportion of individuals who have disability of all degrees is close to the proportion of individuals without disability, but the gap is slightly larger for females compared to that of males. The small difference between the share of females and males who are not working is consistent with the published labor participation rates. For individuals without disability, the shares of working in the three sectors are almost similar between females and males. However, there is a smaller share of females with disability of any degree working in the agriculture sector, while there is a larger share of females with disability of any degree working in other industries than that of males.

We also examine the wage rates. Table 2 presents the summary statistics of wage rates across sectors for the entire sample. Both the mean and median wage rates of the other industries are the highest, while those of the agriculture sector are the lowest. By looking at the maximum and minimum wage rates in each sector, we find that the spread of wage rate in the manufacturing sector is the smallest, while the spread in agriculture is the largest. Following the disaggregation of the sample in Table 1, we further differentiate the wage rates by severity of disability and by gender in Table 3. In all three sectors, the average wage rates of workers with severe disability are lower than the wage rates of individuals with mild and moderate disability and without disability. However, the wage rates of individuals with moderate disability do not differ much from the wage rates of individuals without disability, especially for the females. As shown in Table 3, the average wage rates of males with mild disability in manufacturing and other industries even exceed the wage of their counterparts who do not have disability. However, since there are only limited number of observations in the group of males with mild disability in both manufacturing (16 observations) and other industries (14 observations), a few observations which reported high wages might induce a higher average wage rate, which induces the higher average wage of males with mild disability than their peers without disability.

After comparing the wage rates of females and males, we did not find much discrepancy between the wage rate of females and males without disability. Furthermore, among individuals with severe disability, the wage rates of females are higher than those of males. The wage rate comparison between females and males suggests that gender inequality is not much of a concern in Cambodia, at least in terms of wage rates. It can also help explain the lack of explaining power of gender in the wage equation, which will be discussed later.

Table 2. Summary statistics of wage rate per hour across sectors (unit: Riel)

	Agriculture	Manufacturing	Other Industries
Mean	714	1,296	1,488
Median	889	1,480	2,103
Minimum	21	79	29
Maximum	32,031	13,333	25,938

Note: 1USD is roughly equivalent to 4,000 Riel.

Table 3. Summary statistics of average wage rate per hour by gender and disability severity across sectors (unit: Riel)[7]

	Agri. (No.of Obs)	Manuf. (No.of Obs)	Other Ind. (No.of Obs)
Males with mild disability	855 (16)	1,835 (16)	3,362 (14)
Females with mild disability	826 (21)	1,427 (10)	2,557 (29)
Males with moderate disability	1,007 (19)	1,429 (10)	1,665 (13)
Females with moderate disability	1,187 (9)	1,382 (11)	2,040 (17)
Males with severe disability	569 (6)	763 (3)	1,601 (9)
Females with severe disability	767 (5)	959 (7)	2,002 (12)
Males without disability	900 (443)	1,463 (288)	2,041 (656)
Females without disability	876 (449)	1,499 (351)	2,135 (692)

[7] All individuals disclosed their working hours. However, only a fraction of them reported their monthly salary.

We show the summary statistics for non-labor income in Table 4 and total income in Table 5 by degree of disability severity. There is no significant difference in non-labor income and total income between individuals with moderate disability and individuals without disability. However, both the non-labor income and total income of individuals with severe disability are lower than that of their counterparts with moderate or without disability. There are two possible explanations for this. First, individuals with severe disability might need more attention and care from other family members, who otherwise can participate in income-generating activities. Second, severe disability could weaken the competitiveness in marriage matching markets, which, in turn, may be reflected in the lower household income.

Table 4. Summary statistics of monthly non-labor income by degree of disability (unit: Riel)

	Moderate disability	Severe disability	Without disability
Mean	1,437,000	1,148,000	1,413,000
Median	483,000	469,500	454,400
Minimum	8,000	8,000	0
Maximum	21,060,000	16,626,000	24,140,000

Table 5. Summary statistics of monthly total income by degree of disability (unit: Riel)

	Moderate disability	Severe disability	Without disability
Mean	1,603,000	1,306,000	1,599,000
Median	684,000	616,800	644,200
Minimum	15,400	8,000	8,000
Maximum	21,060,000	16,626,000	24,140,000

We maximised the log-likelihood function in Equation (3.10) and the parameters in the probability Equations in (3.6) and (3.7) are estimated and presented in Table 6. The first three columns of Table 6 show the estimates when a dummy for individuals who reported they have disability is included, while the last three columns report the estimates when a dummy for individuals who indicated that their disability has hindered their participation in employment and income generating activities.

The first three columns in Table 6 show the parameter estimates γ_j for the job choice sets given in Equation (3.7) for the three sectors (γ_1 = agriculture, γ_2 = manufacturing, and γ_3 = other industries). The coefficient estimates for gender in all three sectors are positive but insignificant, suggesting that, holding all other things constant, in Cambodia females have the same or more opportunities or job choice sets than males do, especially in the manufacturing and other industry sectors. This might be surprising at first, but after benchmarking with the summary statistics on wage rate and labor market participation rate of males and females, this finding accurately reflects the state of gender equality in Cambodia.

The coefficient on ability to speak more than one language is positive but insignificant across all three sectors. The coefficient estimate of length of schooling is significant for the job opportunities in the agriculture sector and insignificant for the other two sectors. The length of schooling has a negative effect for the agriculture and manufacturing sectors. This implies that a person with higher education tends to have fewer job opportunities in the agriculture and manufacturing sectors, possibly indicating that the agriculture and manufacturing sectors use education as a proxy to some unobserved quality of labor, such as endurance to do heavy labor jobs. In contrast, the length of schooling has a positive effect for the industry sector. It is likely that the other industry sector has a lower demand for manual labor and puts higher value on technical and soft skills like communication skills, thus treating a higher level of education as an advantage for the job.

Being a newcomer in a locality, which is indicated by not living in his current location since birth, reduces job opportunities in all three sectors, implying that the job market of Cambodia still gives importance to the local ties of employees. The signs of the coefficient estimates for the urban/rural regions reiterate the perception that local ties are highly valued. Since the agriculture sector is usually located in rural areas, living in urban areas reduces the job opportunities in the agriculture sector. On the other hand, the other industry sector is more likely to be established in urban areas, so living in urban areas expands job opportunities. However, only the coefficient for the agriculture sector is significant.

Table 6. Coefficient estimates of the deterministic utility function and opportunity sets

	Agri. (γ_1)	Manuf. (γ_2)	Other Ind. (γ_3)	Agri. (γ_1)	Manuf. (γ_2)	Other Ind. (γ_3)
Intercept	1.185**	-0.661**	0.218	1.182**	-0.658**	0.208
	(0.127)	(0.181)	(0.141)	(0.126)	(0.18)	(0.141)
Gender (0=male; 1=female)	0.014	0.109	0.056	0.015	0.113	0.056
	(0.046)	(0.068)	(0.052)	(0.046)	(0.068)	(0.052)
Language (=1, if speaking more than 1 language)	0.041	0.21	0.008	0.038	0.218	0.004
	(0.126)	(0.177)	(0.144)	(0.126)	(0.176)	(0.144)
Length of schooling	-0.156**	-0.139	0.065	-0.161**	-0.141	0.061
	(0.069)	(0.104)	(0.077)	(0.069)	(0.104)	(0.077)
Living in the current location since birth (0=yes; 1=no)	-0.096**	-0.111	-0.013	-0.095**	-0.111	-0.012
	(0.047)	(0.07)	(0.053)	(0.047)	(0.07)	(0.053)
Urban/Rural (Urban=1; Rural=0)	-0.252**	-0.078	0.075	-0.262**	-0.087	0.064
	(0.054)	(0.081)	(0.06)	(0.054)	(0.081)	(0.06)
Experience	-0.129*	-0.178*	-0.173**	-0.132*	-0.183*	-0.173**
	(0.071)	(0.104)	(0.079)	(0.071)	(0.104)	(0.079)
Experience squared	0.03**	0.04*	0.037**	0.031**	0.041*	0.038**
	(0.014)	(0.021)	(0.016)	(0.014)	(0.021)	(0.016)
Disability (=1, if disability is indicated in the survey)	-1.746**	-1.581**	-1.817**	-	-	-
	(0.426)	(0.444)	(0.439)	-	-	-
Disability (=1, if disability is claimed affecting employment)	-	-	-	-1.272*	-1.284	-1.299*
	-	-	-	(0.762)	(0.783)	(0.779)
Disposible income (1,000,000 KHR)		1.642**			1.702**	
		(0.275)			(0.271)	
Disposible income $*D_i$		1.458**			0.767	
		(0.513)			(0.821)	
Working hours (5,040 hours)		-0.663			-0.541	
		(1.548)			(1.516)	
Working hours $*D_i$		-2.076			0.833	
		(5.79)			(8.264)	
$\alpha_0 D_i$		-0.881**			-0.743**	
		(0.061)			(0.228)	
Likelihood at maximum		-17,241			-17,247	
Likelihood at benchmark		- 37,958			-37,958	
McFadden's ρ^2		**0.55**			**0.55**	
Total no. of observations (no. of disabled)		14,103(982)			14,103(506)	

The coefficient estimates of experience (joint with experience squared) for all three sectors are convex and significant which means that individuals with longer experience get fewer job opportunities up to a certain level of experience. After experience reaches a certain threshold, it can bring more job opportunities. The interesting convexity of experience suggests a distinct industrial feature of Cambodia. In each sector, people with less experience are more likely to have less resistance to the new ways of doing work and could be trained more easily to fit the need of the firm they are working at. However, if the experience of individuals already reaches a certain number of years, they could apply the skills they learned from their own experience and could more likely contribute significantly to operations and even improve the operational process itself. Although this is an interesting perspective that is worthy of deeper research, this topic deviates from the scope of our study and will not be further discussed here.

The negative estimates for disability show that disability reduces job opportunities in all three sectors. Disability reduces job opportunities in agriculture and other industries more than it does in manufacturing, which is consistent with our expectation that in the manufacturing sector people with disability can do heavy jobs by, like by operating machines. However, the severity of disability does not affect opportunities differently.

We rescale the two variables in the deterministic utility function, which are disposable income and hours of work, to make their magnitude and the magnitude of the other variables included in the job choice easily comparable. Disposable income is rescaled to the unit of 1 million Riels. Hours of work is rescaled to the unit of one calendar-month by dividing it by 5,040 (24 hours per day times 7 days times 30 days). As expected, the utility of individuals increases with higher disposable income and fewer hours of work or equivalently more leisure. Consistent with our expectations, for individuals with disability, utility increases more sharply with higher disposable income and fewer hours of work, which means individuals with disabilities depend more on income and demand more leisure time. Although the signs are in the right direction, only the coefficient estimates of disposable income are significant.

The coefficient estimate of $\alpha_0 D_i$ is negative and significant, which indicates that compared with individuals without disability, individuals with disability have lower utility for not working. That is, individuals with disability are more willing to have a job, holding all other things constant.

To measure the goodness-of-fit of the model, we introduce McFadden's ρ^2, which is defined as:

$$\text{McFadden's } \rho^2 = 1 - \frac{L_{max}}{L_{reference}},$$

where L_{max} is the log-likelihood value at maximization and $L_{reference}$ is the log-likelihood value at the reference case. In our analysis, the reference case is to allow individuals to randomly choose one of the four alternative choices (not working and working in either of the three sectors). As shown in Table 6, the McFadden's ρ^2 is 0.55, indicating a very good fit of the model. In the context of discrete choice, the value of McFadden's ρ^2 at 0.1 is usually considered an acceptable goodness of fit.

As previously mentioned, the estimates in the last three columns of Table 6 only differ from the estimates in the first three columns in the definition of dummy for disability. The results in the last three columns seem to be consistent with those of the first three columns. However, in the alternative scenario where the disability dummy refers to individuals who reported that their disability has hindered their participation in employment, the disability dummy for the manufacturing sector is not significant and the likelihood at maximum is smaller than the first scenario. Therefore, we choose the first scenario—where the disability dummy is equal to one if the person has a disability—as the baseline case to perform the subsequent analysis.

We also tried differentiating the severity of disability into mild, moderate, and severe in the estimations. We obtained the same coefficient estimates across severity and there is no change of likelihood at optimal; that is, the severity of disability does not affect opportunities differently. One possible explanation for the same estimates for different degrees of severity is that the disadvantage of severely disabled people has already been incorporated in the lower average wage (see Table 4) in the deterministic utility function rather than in the choice opportunities. By assuming that any job in sector j offers the sector average wage rate, people with disability who are located in the lower tail of the wage rate distribution will not be reflected. Based on this result, we did not pursue the specification that differentiates disability severity.

We also show that, in comparison to the linear regression model, our model captures the significant impact of disability on income through reduced job opportunity. The linear regression model which used monthly salary as the left-hand side variable showed that disability has no impact on salary (refer to Table A2).

5. Welfare of disabled persons measured by capability-equivalent income

We shall now discuss how we can use the parameter estimates obtained in the previous section to compute the welfare measures. In particular, we focus on a measure which we call *Capability-Equivalent Income*. First, we define the expected utility of agent i given his non-labor income I_i, sector average wage rate as $\overline{w}_j$, and the vector of individual characteristics X_i as:

$$(5.1) \quad V(\theta, g_i, w, I_i, X_i) = v(I_i, 0, X_i) + \sum_{j=1}^{3} \theta(j) v(\overline{w}_j h_i + I_i, h_i) g_j(h_i),$$

Based on the specifications in Equations (3.4) to (3.6) and the parameter estimates in Table 6, we can also express the expected utility in (5.1) as:

$$(5.2) \quad V(\theta, g_i, w, I_i, X_i) = \exp\big((\hat{\alpha}_1 + \hat{\alpha}_{1D} * D_i) I_i + \hat{\alpha}_0 * D_i\big)$$

$$+ \sum_{j=1}^{3} exp\big(\hat{\gamma}_{0j} + \hat{\gamma}_j X_i\big) \exp\Big((\hat{\alpha}_1 + \hat{\alpha}_{1D} * D_i)\big(\overline{w}_j h_i + I_i\big) + (\hat{\alpha}_2 + \hat{\alpha}_{2D} * D_i) h_i\Big).$$

$V(\theta, g, w, I_i, X_i)$ is interpreted as the highest utility an individual can expect given his non-labor income, the sector wage rate he receives, his job opportunity sets, and personal characteristics. Specific to individuals with disabilities, there are two ways in which disability can affect expected utility: one is the indirect impact through the available job opportunities and the other is the direct impact through non-labor income and working hours. Let X_i^0 denote the vector of reference characteristics wherein the components of X_i^0 are equal to the components of X_i except for the disability dummies which are set to zero in X_i^0. Setting the disability dummies to zero aims to artificially remove the disabilities of individuals while holding all other variables constant. Correspondingly, we define the expected utility as V_i for an individual with disability and its equivalent expected utility as V_i^0 with X_i replaced by X_i^0, D_i artificially assigned to equal to zero, and the non-labor income I_i replaced by a new non-labor income $\tilde{I}_i$. We denote the new non-labor income, $\tilde{I}(V_i, V_i^0)$, as capability-equivalent income for an individual with characteristics X_i. Given these specifications, we set:

(5.3) $V(\theta, g_i, w, I_i, X_i) = V(\theta, g_i^0, w, \tilde{I}_i, X_i^0)$.

Therefore, the difference between non-labor income I_i when the individual is disabled and the simulated non-labor income $\tilde{I}_i$ when the person is assumed to be without disability is the money metric value of job opportunity loss caused by the disability and the utility changes with respect to non-labor income and the working hours due to disability. The difference can be mathematically expressed as:

(5.4) $CV(X_i, X_i^0) = I_i - \tilde{I}_i$,

that is, $CV(X_i, X_i^0)$ is the amount needed to compensate individuals for their reduced opportunity sets and change in utility with respect to non-labor income and working hours due to disability. In other words, $CV(X_i, X_i^0)$ is a money metric value that represents the loss of welfare for the disabled. Thus, $I_i + CV(X_i, X_i^0)$ is the value of non-labor income necessary for an individual with disability to achieve the same mean utility as an individual who is not disabled, assuming that disabled individuals have the same opportunity sets as individuals who are not disabled.

We applied the formulas in Equations (5.2) and (5.4) to estimate capability-equivalent income for all the disabled individuals in the sample. Table 7 summarises the actual observed disposable income and simulated capability-equivalent income for three scenarios, including compensating for both reduced job opportunities and changed utility with respect to non-labor income and leisure due to disability, partially compensating for reduced job opportunities, and partially compensating for changed utility. For each scenario, we reported both the required compensation and the new total non-labor income by adding the compensation to the original non-labor income.

The partial compensation for reduced job opportunities is positive in all scenarios because disability universally reduced the size of the job set across all three sectors. Therefore, individuals all require a positive compensation to offset such loss.

However, the partial compensation for the changed utility due to disability is positive for some individuals and negative for the others depending on their non-labor income and hours of working. Higher non-labor income would bring higher utility to individuals with disability, while this would lower the utility of their counterparts who are free of disability. This will generate a negative “compensation” when we artificially remove the disability. For the same length of hours of working, artificially transforming individuals from having disability to being free of disability would reduce the negative impact on compensation. Thus, to achieve higher utility, individuals should be compensated for their disability, which would offset their loss from reduced employment opportunities. The net impact between the negative "compensation" from non-labor income and positive compensation from the sacrificed leisure could be either positive or negative. As shown in Table 7, the median value for the partial compensation to utility is negative at -124,566 riels, which implies that the compensation for more than half of the individuals with disability is mostly due to the income impact rather than the leisure impact. This means that the net impact is more likely to be negative. The negative impact also explains why some individuals obtain negative overall compensation.

With the above result, it is more suitable to specify partial compensation for reduced employment opportunities as the capability-equivalent income. This is because the lost job opportunities due to disability can be compensated by income. However, achieving higher utility by getting the same amount of consumption due to disability is very hard to be artificially "assumed away". The psychological dependence on income is very hard to remove unless a person's disability can be physically cured. Therefore, in the following section, we design a transfer program which only applies the partial compensation scheme for reduced employment opportunities.

Table 7. Observed income versus capability-equivalent income for disabled individuals (unit: Riel)

Observed Income and *Capability-Equivalent Income*	Min.	Median	Mean	Max.
The observed I_i for the disabled	8,000	563,700	1,488,881	21,061,400
1. Compensate for both reduced employment opportunities and utility				
The compensation $CV(X_i, X_i^0)$	-6,227,761	3,932,020	3,458,611	4,384,015
The sum of $I_i + CV(X_i, X_i^0)$	387,711	4,487,104	4,947,492	14,833,639
2. Partially compensate for sole reduced employment opportunties				
Partial compensation $CV(X_i, X_i^0)$	3,318,764	3,989,520	3,941,213	4,315,273
The sum of $I_i + partial\ CV(X_i, X_i^0)$	3,882,606	4,572,607	5,430,093	24,693,626
3. Partially compensate for sole utility				
Partial compensation $CV(X_i, X_i^0)$	-9,888,420	-124,566	-554,084	88,043
The sum of $I_i + partial\ CV(X_i, X_i^0)$	-3,164,063	417,116	934,796	11,172,980

6. Welfare assessment for a transfer program by subsiding the disabled through tax

As shown in the previous section, the compensation required by individuals with disability to compensate for the loss in job opportunities and affected utility with respect to consumption and leisure can be measured in money metrics, which we denote the capability-equivalent income. In this section, we will discuss the scenario where we design a transfer program to impose an income tax to the individuals with high income and without disability as a means to subsidize the disabled individuals and see how such transfer program would affect the welfare of and corresponding degree of inequality in the whole society.

In designing this type of program, first we calculate the total amount of monetary compensation. Second, we assume balanced financing of the compensation by imposing an income tax to the top 20% earners without disability; that is, we calculate a tax rate which is sufficient to cover the required subsidy and multiply it to the income of the top 20% earners.[8]

To compare the impact of this transfer program, we first calculate the overall welfare W^B as the baseline welfare by summing up V_i of all individuals as implied in the data. We then calculate the post-transfer welfare W^T by subtracting the income tax from individuals who belong to the top 20% income earners without disability and adding the required capability-equivalent income calculated in section 5 to individuals with disability. That is,

(6.1) $$W^B = \sum_{i \in I_D} V(\theta, g_i, w, I_i, X_i) + \sum_{i \in I_{ND}} V(\theta, g_i, w, I_i, X_i),$$

(6.2) $$W^T = \sum_{i \in I_D} V(\theta, g_i, w, (I_i + CV_i), X_i) + \sum_{i \in I_{ND}} V(\theta, g_i, w, I_i(1 - t), X_i) + \sum_{i \in I_{ND'}} V(\theta, g_i, w, I_i, X_i).$$

where I_D refers to the group of individuals with disability, I_{ND} refers to the group of top 20% eaners without disability, and $I_{ND'}$ refers to the group of individuals without disability but with income not reaching the top 20% threshold. The term CV_i is the needed compensation for individual i with disability and t is the income tax rate. In our analysis, we only compensate for reduced job opportunities due to disability.

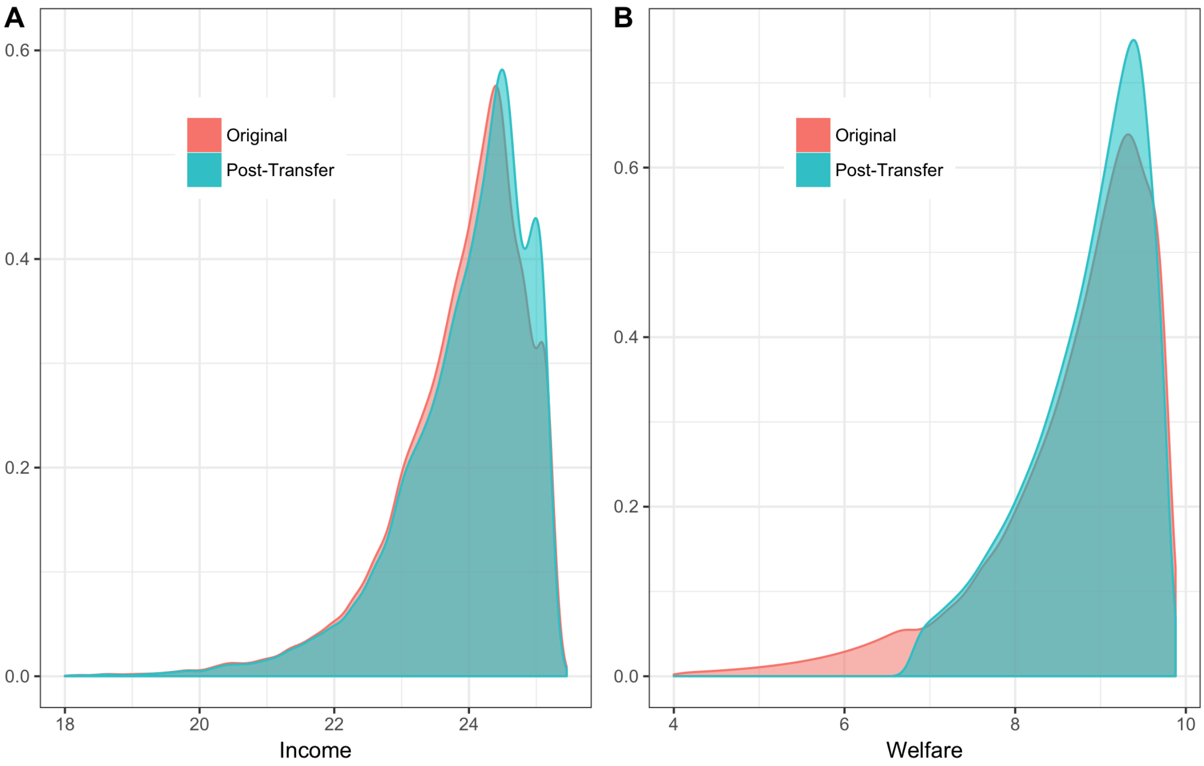
Using the results presented in the fourth row of Table 7, we get the required compensation $\sum_{i \in I_D} CV_i$ for the 982 individuals with disability and divided it by the sum of

[8] We chose the top 20% earners to illustrate the impact. Our framework can be easily extended to do optimal tax scheme simulation where the percentage of income earners who are subjected to tax and the tax rate are chosen simultaneously.

the income of the top 20% earners without disability to get the tax rate. With this method, the tax rate would generate a zero-net impact on the government's balance sheet; that is, the compensation given to individuals with disability would be fully financed through tax. We then evaluate the impact of such transfer program on the distribution of income and welfare.

In Figure 1, we first plot the distribution of total income in the pre- and post- transfer program cases and then plot the distribution of welfare in the pre- and post- transfer program cases. We see that the expected utility $V(i)$ provides a perfect measure for welfare. Although the level of expected utility itself has no real meaning, its distribution and the change when compared with the original level can help us reveal the impact on welfare.

Figure 1. Distribution of Income and Welfare Pre- and Post- Transfer Program



The Gini coefficient of total income declined from 0.66 in the pre-transfer program case to 0.58 in the post-transfer program case. The Gini coefficient of welfare also dropped from 0.14 to 0.07. The absolute level of the welfare of the whole society increased from 25280 to 26237, a 3.8% net increase.

We are fully aware that such a transfer program is a simplified case. Our intention is to demonstrate the impact of this type of transfer scheme on reducing inequality and improving welfare. Our framework can be easily extended to accommodate more sophisticated designs of subsidy-tax transfer program. However, we leave it for further research, since this is beyond the scope of this study.

7. Conclusion

This paper develops an empirical model to construct welfare measures in the context of the capability approach. This approach is based on the concept of probabilistic choice which is rooted in the random utility theory, and it has the advantage of supporting practical tools for structural analyses of qualitative choice behaviour. A key contribution of this analysis is the construction of a measure of capability-equivalent income that allows us to empirically examine welfare effects.

We develop a labor supply model that accommodates the major institutional features of the labor market; namely, that the market is organized into several sectors and that labor supply can be viewed as a choice among jobs, wherein the observed hours of work pertains to job-specific fixed hours chosen by the individual from his latent set of available jobs. The latent set of available jobs, in turn, is subject to constraints associated with the individuals' capability, such as disability. We chose to use the Cambodia Socio-Economic Survey 2009 data that discloses detailed information on an individual's disability. Based on the utility function, we deduce the corresponding compensation variation equation and compute the capability-equivalent income for disabled persons. Capability-equivalent income is measured in terms of non-labor income, which can be understood as a subsidy to individuals with disability to compensate for the utility loss due to fewer opportunities in the job market.

With the estimated parameters, we simulate the amount of non-labor income of individuals after being compensated for their capability disadvantage by taxing the top earners. We show that after the transfer, the inequality in both income and welfare decreases and the overall welfare increases. Our analysis illustrates how to estimate the money metric measure for capability inequality, which is of high interest for policymakers who want to quantitatively assess the welfare loss caused by differences in capability. One drawback of our modelling framework is that it is a static one-period model. There are no lifetime considerations about job choice. An adoption of dynamic choice model where individuals maximize lifetime utility, as suggested by Blundell et al. (2016), would allow the simulations of alternative policies, like expanding the choice set or giving subsidies, more realistic. However, such dynamic choice model requires using panel data of high quality with individual choices being observed in a considerable length of time. Moreover, to evaluate capability-related policies accurately, a well-functioning labor market is a necessary condition. As a country which was recently upgraded to the lower middle-income status,[9] Cambodia is still far from having a mature labor market, which is more appropriate when building a dynamic modelling framework.

The labor supply modelling framework can be extended for future research to accommodate the complexity of real labor market behaviours. The optimization algorithm also remains a challenge when more variables are included in the model.

[9] Based on the World Bank Group classification, Cambodia became a lower middle-income country in 2016.

References

Anand, P., G. Hunter and R. Smith (2005): Capabilities and Wellbeing: Evidence Based on the Sen–Nussbaum Approach to Welfare. *Social Indicators Research*, **79**, 9–55.

Anand, P., C. Santos and R. Smith (2008): The Measurement of Capabilities. In K. Basu and R. Kanbur (eds), *Arguments for a Better World: Essays in Honour of Amartya Sen.* Oxford University Press, Oxford.

Anand, P., K. Jaya and N. B. Tran (2011): Measuring Welfare: Latent Variable Models for Happiness and Capabilities in the Presence of Unobserved Heterogeneity. *Journal of Public Economics*, **95**, 205-215.

Blundell, R., Costa Dias, M., Meghir, C. and Shaw, J. (2016), Female Labor Supply, Human Capital, and Welfare Reform. Econometrica, 84: 1705–1753.

Dagsvik, J. K. (2012): Justification of Functional Form Assumptions in Structural Models: Application and Testing of Qualitative Measurement Axioms. *Theory and Decision,* DOI 10.1007/s11238-9308-5.

Dagsvik, J. K. (2013): Making Sen's Capability Approach Operational: A Random Scale Framework. *Theory and Decision*, **74**, 75-105.

Dagsvik, J. K. and S. Strøm (2006): Sectoral Labor Supply, Choice Restrictions and Functional Form. *Journal of Applied Econometrics*, **21**, 803–826.

Dagsvik, J. K. and Z. Jia (2006): Labor Supply as a Choice among Latent Job Opportunities. A Practical Empirical Approach. Discussion Papers, no. 481, Statistics Norway.

Dagsvik, J. K. and S. Røine Hoff (2011): Justification of Functional Form Assumptions in Structural Models: Application and Testing of Qualitative Measurement Axioms. *Theory and Decision,* **70**, 215-254.

Dagsvik, J. K., Z. Jia, T. Kornstad and T. O. Thoresen (2012): Theoretical and Practial Arguments for Modeling Labor Supply as a Choice among Latent Jobs. *Journal of Economic Surveys*, DOI: 10.1111/joes.12003.

Drèze, J., and A. Sen (2002): *India: Development and Participation*. Oxford University Press, Oxford.

Han,X. (2010): Housing demand in Shanghai: A discrete choice approach. *China Economic Review* **21**: 355-376

Heuveline, P. (2015) The boundaries of genocide: Quantifying the uncertainty of the death toll during the Poll Pot regime in Cambodia (1975-79). Population Studies (69): 201-218.

Kuklys, W. (2005): *Amartya Sen's Capability Approach. Theoretical Insights and Empirical Applications.* Springer-Verlag, Berlin.

McFadden, D. (1973): Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (ed.), *Frontiers in Econometrics*. Academic Press, New York.

Robeyns, I. (2003): The Capability Approach: An Interdisciplinary Introduction. Mimeo, Department of Political Science and Amsterdam School of Social Science Research, University of Amsterdam.

Robeyns, I. (2006): The Capability Approach in Practice. *Journal of Political Philosophy*, **14**, 351–376.

Robeyns, I., and W. Kuklys (2005): Sen's Capability Approach to Welfare Economics. In W. Kuklys (ed.), *Amartya Sen's Capability Approach: Theoretical Insight and Empirical Applications.* Springer-Verlag, Berlin.

Sen, A. K. (1979): Personal Utilities and Public Judgments: or What's Wrong with Welfare Economics. *Economic Journal*, **89,** 537–558.

Sen, A. K. (1985a): *Commodities and Capabilities*. North-Holland, Amsterdam.

Sen, A. K. (1985b): Social Choice Theory. In A. Arrow and M. D. Intriligator (eds), *Handbook of Mathematical Economics*, Vol. III, North-Holland, Amsterdam.

Sen, A. K. (1992): *Inequality Re-examined*. Clarendon Press, Oxford.

Sen, A. K. (1997): From Income Inequality of Economic Inequality. *Southern Economic Journal*, **64**, 384–396.

Sen, A. K. (1998): *Welfare Economics and the Quality of Life.* Chung-Hua series of lectures by invited eminent economists, no 24. Institute of Economics, Academia Sinica. Nankan, Taipei, Taiwan.

Van Soest (1995): Structural Models of Family Labor Supply: A Discrete Choice Approach. *Journal of Human Resources*, **30**, 63-88.

Appendix 1: Summary of Empirical Results

Table A1. Estimation results of wage equations using OLS

	Agri.	Manuf.	OtherInd.	Pooled
Intercept	640**	1171**	1912**	
	'(150)	'(175)	'(260)	
Dummy for Agriculture				668**
				'(143)
Dummy for Manufacture				1252**
				'(149)
Dummy for Other Industry				1885**
				'(141)
Gender (0=male; 1=female)	-42	11	-33	-26
	'(72)	'(84)	'(124)	'(65)
Language (=1, if speaking more than 1 language)	-61	-270	350	30
	'(201)	'(223)	'(418)	'(194)
Length of schooling	86	349**	307**	257**
	'(108)	'(125)	'(176)	'(95)
Living in the current location since birth (0=yes; 1=no)	130**	143**	82	106**
	'(74)	'(86)	'(126)	'(66)
Urban/Rural (Urban=1; Rural=0)	64	100	522**	303**
	'(91)	'(96)	'(138)	'(75)
Experience	29	7	-115	-56
	'(108)	'(129)	'(185)	'(98)
Experience squared	10	0	11	11
	'(21)	'(25)	'(37)	'(19)
Disability (=1, if disability is indicated in the survey)	12	-69	218	66
	'(134)	'(160)	'(257)	'(128)
Ajusted R-sqaured	0.003	0.008	0.010	0.42
No. of Obs.	1143	824	1740	3707
	(7749)	(1495)	(3954)	(14103)

Table A2. Impact of Disability on Observed Salary Income (unit: Riel)

	Agri.	Manuf.	OtherInd.	Pooled
Intercept	141052**	222189**	370386**	
	(24715)	(31881)	(50932)	
Dummy for Agriculture				136378**
				(27296)
Dummy for Manufacture				253862**
				(28420)
Dummy for Other Industry				361476**
				(26892)
Gender (0=male; 1=female)	-10454	15644	-9247	-3440
	(11865)	(15406)	(24193)	(12394)
Language (=1, if speaking more than 1 language)	-18837	-65405**	57446	-4346
	(33140)	(40135)	(81798)	(36896)
Length of schooling	2465	59146**	50061	40007**
	(17903)	(22902)	(34422)	(18135)
Living in the current location since birth (0=yes; 1=no)	22381**	20933	34982	27129**
	(12186)	(15724)	(24593)	(12670)
Urban/Rural (Urban=1; Rural=0)	21345	12190	97641**	57696**
	(14953)	(17601)	(26985)	(14409)
Experience	10226	24294	-21514	-3111
	(17870)	(23576)	(36228)	(18655)
Experience squared	477	-4224	1705	350
	(3487)	(4645)	(7224)	(3684)
Disability (=1, if disability is indicated in the survey)	17239	-31109	46430	14850
	(22176)	(29350)	(50306)	(24404)
Ajusted R-sqaured	0.006	0.008	0.009	0.43
No. of Obs.	1146	827	1741	3714
	(7749)	(1457)	(3954)	(14103)

Appendix 2: Section 14 of CSES 2009 Questionnaire

14. DISABILITY | ***Respondent: Head of household or the spouse of the head of household*** | ***WEEK 3***

Please provide information on all members who usually reside in this household.

ID NUMBER	Does ..[NAME].. have any of the following? **Enter the 3 most important** 01 = Difficulty seeing 02 = Difficulty hearing 03 = Difficulty speaking 04 = Difficulty moving 05 = Difficulties in feeling or sensing 06 = Psychological or behaviourial difficulties 07 = Learning difficulties 08 = Fits 09 = Other (specify) 98 = Don't knoW **Enter '0' if none, (>> NEXT PERSON)** **The 3 most important**			Is the difficulty ... 1 = Mild 2 = Moderate 3 = Severe **Enter one code for each of the difficulties reported in Col 2a-2c**			What was the cause? 01 = Mine/UXO 02 = Traffic Accident 03 = Work Accident 04 = Disease(s) 05 = Congenital 06 = Fever 07 = Difficulty Delivery 08 = Chemical Accident 09 = Rape 10 = Violent Attack 11 = Domestic Violent 12 = Suicide Attempt 13 = Mental Trauma due to war and other traumatic events 14 = War Injuries 15 = Malnutrition 16 = Burns 17 = Torture 18 = Old Age 19 = Other (specify) 98 = Don't know **Enter one code (the most important) for each of the difficulties reported in Col 2a-2c**		
(1)	(2a)	(2b)	(2c)	(3a)	(3b)	(3c)	(4a)	(4b)	(4c)
01									
02									
03									
04									
05									
06									
07									
08									
09									
10									
11									
12									
13									
14									
15									

Cambodia Socio-Economic Survey 2009 (CSES) | Page 48 of 63

14. DISABILITY (CONTINUED) | ***Respondent: Head of household or the spouse of the head of household*** | ***WEEK 3***

Please provide information on all members who usually reside in this household.

ID NUMBER	Does the difficulty/difficulties prevent ..[NAME].. from participation or access to any of the following? 1 = Education (ask if aged 3 or over) 2 = Housing 3 = Land ownership (ask if aged over 18) 4 = Employment and income generation 5 = Health services 6 = Transport **Enter the three most important** **Leave blank if "None"**		
(1)	(6a)	(6b)	(6c)
01			
02			
03			
04			
05			
06			
07			
08			
09			
10			
11			
12			
13			
14			
15			

END OF WEEK 3