Working Paper No. 871

Simulations of Employment for Individuals in LIMTCP Consumption-poor Households in Tanzania and Ghana, 2012

by

Thomas Masterson, **Kijong Kim**, and **Fernando Rios-Avila**
Levy Economics Institute of Bard College

August 2016

Levy Economics Institute
P.O. Box 5000
Annandale-on-Hudson, NY 12504-5000
http://www.levyinstitute.org

ISSN 1547-366X

ABSTRACT

New methodology for producing employment microsimulations is introduced, with a focus on farms and household nonfarm enterprises. Previous simulations have not dealt with the issue of reduced production in farm and nonfarm household enterprises when household members are placed in paid employment. In this paper, we present a method for addressing the trade-off between paid employment and the farm and nonfarm business activities individuals may already be engaged in. The implementation of the simulations for Ghana and Tanzania is described and the quality of the simulation results is assessed.

Keywords: LIMTCP; Microsimulation; Ghana; Tanzania; Employment; Unpaid Family Labor; Household Production; Time Use

JEL Classifications: C14, C40, D31, J22

INTRODUCTION

This paper documents the creation of employment simulations for Tanzania and Ghana to test the impact of employment gains on the time and income poverty of individuals and households using our estimates of the Levy Institute Measure of Time and Consumption Poverty (LIMTCP). Unlike previous simulations in countries like Turkey or Mexico (Masterson 2012; Masterson 2013), in this case we only assign paid employment to people who are already employed on household farms or in household enterprises if the earnings at least replaced a significant portion of their estimated contribution to farm output and/or household business income. As a result, the assignment of jobs was done for a smaller portion of eligible adults in poor households than in previous simulations. The simulation will be more plausible, since we are not imposing large losses on any households.

The purpose of the exercise is to estimate the real impact on time and consumption poverty of some policy aimed at alleviating poverty through the promotion of paid employment. Any such shift into paid employment entails not only changes in household earnings from paid work and the distribution of time allocated to necessary household production in affected households, but also a shift away from time spent on productive activities already being carried out by members of consumption-poor households on the family farm or in a family business enterprise. In previous simulations we rejected job assignments if the resulting changes in individuals' earnings were negative (if the individual was already doing paid work, but we attempted to assign full-time employment, for example), since we were attempting to estimate the effect of voluntary, not mandatory, paid employment. No individual or household was made worse off in terms of income or consumption in those simulations.

In this case we also consider the individuals' contributions to farm income and nonfarm business income when assessing whether to reject the assignment for a given individual in the simulation. This requires a further step of estimating these individuals' contributions to farm and nonfarm income. To make the assessment we also need to draw a line in the sand: set a threshold below which assignment of paid work would be rejected for an individual. In previous simulations we implicitly assumed that the threshold for the net benefit of paid employment was zero in purely

monetary terms. If the change in earnings from the job assignment was negative, we did not keep the assignment. In this case, we assume that there is a non-monetary benefit to paid employment not captured by the change in earnings. Therefore we set the threshold to a bit below zero. We discuss the details in the section on methodology, below.

As always with these types of simulations, it is not possible to assess how well the assignment is done. Since we are creating a counterfactual distribution of earnings and time allocation, we have nothing against which to compare the results, other than the baseline actual scenario. Given that fact, we do check that the results are not implausible given the characteristics of the recipients and donors and the actual distribution of time and income. These checks are presented in the "results and quality" section below. We conclude with an overall assessment of the exercise.

DATA AND METHODOLOGY

The base data sets for the two countries are the LIMTCP estimates created for this project. These are synthetic data files, created with a statistical match of the each country's household survey (with which official poverty statistics are calculated), as well as a time use survey for that country. For Ghana, the sixth Ghana Living Standards Survey (GLSS6), conducted in 2012, is matched with the 2009 Ghana Time Use Survey (GTUS). For Tanzania, the 2012 Household Budget Survey (HBS) is matched with the 2006 Integrated Labour Force Survey Time Use add-on module (TUS).[1] With the matched files we calculate the LIMTCP for Ghana and Tanzania. The base file for the simulation is the matched file plus the LIMTCP estimates.

The simulations, as always, involved several steps. However, we incorporated a new step into the simulations for this project: estimating the contribution of each individual to household farm income and nonfarm business income. We now specify in detail the steps we took to produce the simulation estimates.

[1] See Rios Avila (2016).

The first step is to identify donor and recipient pools for job assignments. We first determine which of the individuals in the base data set is eligible for the analysis. By eligible we mean between the ages of 18 and 70 and not in school, retired, or disabled. In the Ghanaian simulation, this step reduces the number of records to 36,146 (representing 13,624,024 people) from the total of 71,717. In the Tanzanian simulation, 21,991 of 46,535 records (representing 18,933,118 people) were eligible. From these records we identify donors and recipients. The recipients are those who may be assigned a paid job in the simulation. These are individuals in LIMTCP poor households who are either: not employed; are working for pay for less than 10 hours per week; or working in an actual primary activity other than as a paid employee or apprentice. The latter categories included "non-agricultural contributing family worker," "agricultural self-employed without employees," or "agricultural contributing family worker" in the case of Ghana[2] and "working on the household farm" or "helping without pay in household business" in the case of Tanzania.[3] The donors are those who are currently working for pay for 10 hours per week or more as their primary activity.

ESTIMATING A PRODUCTION FUNCTION

We need to account for the reduction in output due to each individual recipient leaving the family farm or nonfarm business to take up paid employment. In order to approximate each member's contributions to household farms and enterprises, we estimate a *log-linear* production function defined as:

$$\ln Y = \alpha + \boldsymbol{\beta} \ln \boldsymbol{L_F} + \gamma_1 \ln L_H + \gamma_2 \ln H + \gamma_3 \ln K + \gamma_4 \ln X + \boldsymbol{\varphi Z} + \mu$$

Where Y is the value of output, $\boldsymbol{lnL_F}$ is a vector of the log of the amount of family labor by age categories[4] and sex; L_H is the amount of hired labor; H is the amount of land operated (in the case of farm businesses); K is the amount of capital employed; X is the amount of other inputs

[2] From section 4, part A, question 20 of the GLSS 6: "What was the status of (NAME) in this job?"
[3] From section 12, question 10a of the HBS 2012: "Which of these activities is (NAMES) primary activity?"
[4] The six categories are: less than 18 years old, 18 to 24 years old, 25 to 44 years old, 45 to 64 years old, and 65 or older.

into production; and $\boldsymbol{Z}$ is a vector of household characteristics, including dummies for agro-climatic zone (in the case of farms), region, rural/urban status, age, sex, and education level of the household head.

The measure of output in the case of farms includes all agricultural products produced, whether sold or consumed, valued at reported prices. For households that did not report prices for given items, a local average price for the item was used. If no local price was available, we used a regional or national average price. A similar procedure is used for aggregating the value of land, capital equipment, livestock, and other inputs into production.

ESTIMATING INDIVIDUAL CONTRIBUTIONS TO HOUSEHOLD FARM/BUSINESS

We then estimate the contribution to production of each individual. First, we predict the level of output for each farm/business using the results of the regression. Next we calculate the level of operating expenses per weekly hour of family labor employed. Then, for each individual in the household that works on the farm or in the business we subtract their weekly hours worked from the household total for their age-sex category and we subtract the amount of inputs (operating expenses) for their hours of work. Then we predict the output for that household *at the individual level* using the same regression results with adjusted household totals. This produces an estimate of the gross contribution of each individual family worker to gross output. We scale the sum of these individual contributions to equal the actual gross output for the household and then subtract (for each individual) the cost of the operating expenses that would not be used due to their not working on the farm/business.[5] The result is an estimate for each individual of their net contribution to the family farm or nonfarm business enterprise.

[5] We assume here that the relationship between operating costs (inputs) and family labor inputs is linear.

ASSIGNING INDUSTRY AND OCCUPATION

The next step in assigning jobs to recipients is to determine the likeliest industry and occupation for each of the potential job recipients. This is done using a multinomial logit procedure. Industry and occupation are regressed on age, age squared, sex, rural/urban status, education, and geographic region in the donor pool. The likelihood for each industry and occupation is then predicted in the recipient pool using the results of the multinomial logit. Then each recipient is assigned the industry and occupation corresponding to the largest predicted likelihoods.

IMPUTED WAGES AND HOURS

The imputations for the earnings and usual weekly hours of paid work are performed using a three-stage Heckit procedure (Berndt 1996: 627) separately for each combination of four age categories[6] and sex. The first stage is a probit estimation of labor force participation:

$$lf_i = \alpha_1 + \beta X + \varepsilon_i$$

The vector of explanatory variables, *X*, comprises the number of children under the age of five and the number of children ages six to seventeen in the household, and the individual's education, as well as the individual's spouse's age, education, and labor force status. The regression is run on the universe of all eligible adults. The Mills ratio is calculated for all individuals using the results of the first stage regression:

$$\lambda = f\left(\frac{-\hat{lf}}{\sigma_{\hat{lf}}}\right) \Bigg/ \left(1 - F\left(\frac{-\hat{lf}}{\sigma_{\hat{lf}}}\right)\right)$$

[6] Less than 25 years old, 25 to 34 years old, 35 to 54 years old, and 55 and older.

Where f is the normal density function, F is the normal distribution function, $\hat{lf}$ is the estimated probability of labor force participation, and $\sigma_{\hat{lf}}$ is the standard deviation of $\hat{lf}$.

The second stage is an ordinary least squares (OLS) estimate of the log of hourly wage:

$$\ln w_i = \alpha_2 + \gamma_2 Z + \theta_2 \lambda + \mu_i$$

This regression is run only on those that are actually employed for pay. The vector of explanatory variables, *Z*, in this stage includes the individual's education, age, industry, occupation, geographic region, rural/urban location, spouse's labor force status, and finally, λ, the Mills ratio calculated in the first stage. Inclusion of the Mills ratio corrects for the selection bias induced by limiting the regression to those in paid employment. The imputed log of wage is predicted for donors and recipients from the results of the regression, with industry and occupation replaced for the latter by the industries and occupations assigned in the previous step.

The third stage is a regression of usual hours of paid work per week:

$$h_i = \alpha_3 + \gamma_3 Z + \omega \ln \hat{w}_i + \theta_3 \lambda + \eta_i$$

The regression is once again run only on those in paid employment. The vector of explanatory variables, *Z*, in this stage is the same as the previous stage, with the addition of the number of children under five years of age and the number of children ages six to seventeen in the household. Finally, the imputed wage predicted in the second stage and the Mills ratio calculated in the first stage are included. Imputed hours per week are predicted for donors and recipients using the results of the regression, replacing the industry and occupation of the latter with their assigned values. The results of the last two stages give us the remaining variables with which we perform the hot-decking procedure to assign actual earnings, hours, industry, and occupation to recipients.

ASSIGNING EARNINGS AND HOURS

We can now assign earnings, usual hours of work, industry, and occupation to those individuals in the recipient pool. The assignment method is statistical matching with hot-decking (Andridge and Little 2010). The matches are performed within cells formed from combinations of age, sex, and educational attainment. The variables used to assess nearness of match are family type, spouse's labor force status and educational attainment, assigned industry and occupation, the number of children under five years of age and the number of children ages six to seventeen in the household, and the two imputed variables (log of wage and hours worked). We use affinity score matching, which allows us to weight the matches of each of the matching variables by importance. Industry and occupation are the most heavily weighted variables, followed by imputed hours and wage. After these, we weight family type and spouse's full-time/part-time status, followed by marital status and spouse's education and labor force status, and finally the variables detailing the number of children in the household. Matches are drawn randomly from all those donor records with the highest affinity score for an individual recipient. Industry, occupation, earnings, and hours from both the donor's primary and secondary activity are transferred to the recipient.

COMPARING SIMULATED EARNINGS TO ACTUAL EARNINGS/CONTRIBUTION

Once the hot-decking is finished, we compare the earnings each recipient gets with the value of lost production, calculated as described above. We cancel any assignments with a large enough negative impact, and for the rest adjust income from household farm/business. We define the cutoff for a "large enough" negative impact using the ratio of the simulated earnings to the recipient's estimated net contribution to family farm/business output plus reported individual earnings. For those individuals for whom this ratio is less than 75%, we reverse the results of the simulation. The rest of the recipients remain in the "adjusted" recipient pool.

REASSIGNING HOUSEHOLD PRODUCTION SHARES

Finally, we need to reallocate the shares of required household production in order to recompute individuals' time deficits/surpluses as a result of the simulation. As many individuals' paid/unpaid work hours may have changed as a result of the simulation, we need to adjust the shares of household production for all the adult members of all the households with simulation job recipients. We use a second round of hot-decking to assign new weekly hours of household production, new hours caring for young children (since we will be reassigning child care hours contracted in the next stage), and new commuting hours to each of the adults, based on updated labor force participation variables for the recipients of jobs in the first stage. The method is the same as the first stage, with the exception of the matching variables used and their relative weighting in the procedure. In this stage, the variables used to assess nearness of match are family type, spouse's labor force status, number of adults, number of children, and the number of children under five years of age and those ages six to seventeen in the household, simulated net household income, the income share of each individual,[7] simulated usual weekly hours of employment, and household total simulated hours of employment. All income and labor force variables are updated to reflect the new job assignments received in the previous stage. In this round of hot-decking, the number of children and number of adults in the household are weighted most heavily of all the variables. Next most-heavily weighted are family type and income share. Finally, the variables detailing the number of young children in the household, followed by net household income, hours of employment and household hours of employment, and finally spouse's labor force status receive the lowest weights. For each match, the weekly hours of household production are transferred. We now have the income and time use variables necessary to recalculate time and income poverty for recipient households.

[7] Income share is included to reflect changes in bargaining power within the household and its impact on the distribution of household production work.

RESULTS AND QUALITY

In order to assess the quality of the simulation we do a number of comparisons between the donor and recipient pools. It should be noted that since we are creating a counterfactual distribution of earnings, employment hours, and household production hours, there exists no standard against which to measure the quality of the simulation. Nevertheless, we check that the resulting distributions are not too different from the donor pool. They cannot be the same, because (for obvious reasons) the donor pool is very different than the pool of recipients. To emphasize the point, we begin with a comparison of the recipient and donor pools for each country, by several categories used in the simulation itself. Table 1, below, provides the breakdown in percentages as well as the overall total numbers of weighted individuals in each country. The distributions of donors and recipients are quite similar in the two countries. In both countries the recipient pool is majority female while the donor pool is majority male. This simply reflects the fact that paid employment is dominated by men in each country. In both countries, donors are concentrated in the 25–44 year age range, while recipients are more likely to be younger. In both countries, the recipient pool has less education than the donor pool, and is three to four times as likely to live in rural areas as in urban areas. The pools are clearly not representative of the same population within each country. The implication of this fact is that although the donor pools are sizable in each country, the matches will be drawn from a small subset of the available donors in many instances.

Table 1. Donor and Recipient Pools for Jobs Simulations, Ghana and Tanzania

	Ghana		Tanzania	
	Donors	**Recipients**	**Donors**	**Recipients**
Sex				
Male	70.0%	46.7%	66.8%	44.5%
Female	30.0%	53.3%	33.2%	55.5%
Age Category				
Less than 25	12.8%	29.2%	18.4%	26.1%
25 to 34	38.4%	22.0%	34.8%	25.5%
35 to 44	24.6%	21.0%	24.0%	23.0%
45 to 54	16.5%	14.8%	15.1%	13.2%
55 to 64	6.7%	8.6%	6.8%	8.4%
65 or older	1.0%	4.5%	0.9%	3.7%
Educational Attainment				
Never attended	4.7%	43.4%	4.6%	27.0%
Primary	6.2%	19.9%	5.5%	15.9%
Middle	35.2%	26.7%	48.6%	50.3%
Secondary or above	54.0%	10.0%	41.3%	6.9%
Rural/Urban Status				
Rural	18.3%	81.3%	27.8%	85.1%
Urban	81.7%	18.7%	32.2%	12.8%
Dar Es Salaam			40.0%	2.1%
Total	1,825,717	2,634,239	2,003,097	5,513,792

Source: Author's calculations using synthetic data files described in Rios Avila (forthcoming).

We next estimate individuals' contributions to farm and nonfarm business output. The results of the regressions we run are shown in tables 2 and 3, below.[8] The variables for family labor are specified as, for example, *lnflm_lt18* is the natural log of the weekly hours of family labor contributed by males less than 18 years old to the family farm. The agroecological region variables are groupings of the regional variables.[9] They perform well, given the fact that it is

[8] Zero values for inputs were set to the weighted mean of the natural log of the respective inputs and a dummy variable for zero values for each input was included in the regression. The results for the dummy variables for the various inputs are omitted from the tables presented here but can be furnished upon request.

[9] For Ghana, Upper East, Upper West, and Northern regions are in the Savannah agroecological region; Brong Ahafo is in the Transitional agroecological region; Ashanti, Eastern, and Volta are in the Deciduous Forest agroecological region; Western is in the Rainforest agroecological region; and Central and Greater Accra are in the Coastal Savannah agroecological region. For Tanzania, the groupings were as follows: Pwani, Dar es Salaam, and Mtwara in the Coastal Plains agroecological region; Dodoma, Kilimanjaro, Tanga, Morogoro, Ruvuma, and Manyara in the Eastern Plateaux and Mountain agroecological region; Iringa in the High Plains and Plateaux agroecological region; Rrusha and Mara in the Volcanoes and Rift agroecological region; Mbeya, Singida, Tabora, Rukwa, Kigoma, Shinyanga, and Mwanza in the Central Plateaux agroecological region; Lindi in the Sediments agroecological region; and Kagera in the Western Highlands agroecological region.

cross-sectional data. The results of these regressions are used to predict output both at the household level and at the individual level. In the latter case, we predict the output after subtracting that individual's labor and a proportional amount of other inputs (other than land) used. The difference between the household prediction and each individual's predicted output is then scaled to add up to the total output, yielding estimates of each individual's contributions to the family farm and nonfarm business output.

Table 2. Farm and Nonfarm Production Regression Results for Ghana

Farm	Coefficient	Standard Error		Nonfarm	Coefficient	Standard Error
lnH	0.140	0.013		lnnflm_lt18	-0.004	0.082
lnflm_lt18	0.028	0.027		lnnflm_1825	0.141	0.126
lnflm_1825	0.028	0.036		lnnflm_2545	0.124	0.240
lnflm_2545	0.048	0.027		lnnflm_4564	0.379	0.394
lnflm_4564	-0.002	0.033		lnnflm_ge65	0.028	0.524
lnflm_ge65	-0.006	0.061		lnnflf_lt18	-0.139	0.065
lnflf_lt18	0.031	0.029		lnnflf_1825	0.229	0.089
lnflf_1825	0.040	0.038		lnnflf_2545	0.117	0.147
lnflf_2545	0.012	0.025		lnnflf_4564	-0.535	0.286
lnflf_4564	-0.007	0.034		lnnflf_ge65	0.030	0.492
lnflf_ge65	-0.037	0.057		lnnfL	0.419	0.016
lnL	0.536	0.009		lnnfK	0.060	0.010
lnK	0.052	0.008		lnnfX	0.328	0.007
lnX	0.142	0.010		Female	-0.197	0.033
Agroecological Region				**Region**		
Transitional	-0.375	0.061		Central	-0.747	0.072
Deciduous Forest	-0.387	0.060		Greater Accra	-0.739	0.059
Rainforest	-0.206	0.066		Volta	-0.419	0.065
Coastal Savannah	-0.538	0.080		Eastern	-0.641	0.064
Region				Ashanti	-0.008	0.054
Central	0.370	0.071		Brong Ahafo	-0.382	0.069
Volta	0.257	0.039		Northern	-0.292	0.075
Eastern	0.250	0.040		Upper East	-0.502	0.108
Northern	-0.248	0.060		Upper West	-0.806	0.114
Upper East	-0.197	0.067		Urban	0.311	0.033
Age	-0.002	0.001		**Education**		
Education				Primary not complete	0.003	0.052
Primary not complete	0.003	0.033		Primary complete	-0.003	0.044
Primary complete	-0.057	0.029		Secondary or above	0.107	0.051
Secondary or above	-0.142	0.039		Age	-0.004	0.001
Female	-0.117	0.033		Constant	3.422	3.004
Urban	-0.056	0.026				
Constant	3.205	0.363				
Adjusted R^2	0.551			Adjusted R^2	0.440	

Table 3. Farm and Nonfarm Production Regression Results for Tanzania

Farm	Coefficient	Standard Error		Nonfarm	Coefficient	Standard Error
lnH	0.289	0.018		lnnflm_lt18	0.004	0.073
lnflm_lt18	0.008	0.027		lnnflm_1825	0.255	0.085
lnflm_1825	-0.038	0.038		lnnflm_2545	0.112	0.097
lnflm_2545	0.092	0.032		lnnflm_4564	-0.174	0.136
lnflm_4564	0.134	0.046		lnnflm_ge65	3.446	0.567
lnflm_ge65	0.072	0.074		lnnflf_lt18	-0.117	0.066
lnflf_lt18	0.011	0.028		lnnflf_1825	0.102	0.081
lnflf_1825	0.085	0.040		lnnflf_2545	-0.002	0.070
lnflf_2545	-0.083	0.033		lnnflf_4564	-0.302	0.143
lnflf_4564	0.063	0.045		lnnflf_ge65	-0.052	0.254
lnflf_ge65	-0.132	0.067		lnnfL	0.142	0.086
lnL	0.170	0.016		lnnfX	0.533	0.013
lnK	0.150	0.008		Female	-0.175	0.051
lnX	0.193	0.011		**Region**		
Agroecological Region				Arusha	0.135	0.134
Eastern Plateaux and Mountain	-0.115	0.094		Kilimanjaro	-0.517	0.165
High Plains and Plateaux	-0.157	0.094		Tanga	0.313	0.120
Volcanoes and Rift	-0.005	0.099		Morogoro	-0.211	0.125
Central Plateaux	0.105	0.089		Pwani	0.013	0.139
Sediments	0.542	0.105		Dar es Salaam	0.235	0.109
Western Highlands	0.584	0.092		Lindi	0.222	0.145
Region				Mtwara	0.203	0.154
Arusha	0.210	0.100		Ruvuma	0.244	0.144
Kilimanjaro	0.163	0.098		Iringa	-0.036	0.131
Tanga	-0.193	0.083		Mbeya	-0.071	0.120
Morogoro	0.474	0.084		Singida	0.020	0.149
Pwani	-0.210	0.116		Tabora	0.177	0.156
Dar es Salaam	0.682	0.165		Rukwa	0.216	0.134
Ruvuma	0.194	0.082		Kigoma	-0.065	0.125
Mbeya	-0.193	0.074		Shinyanga	0.192	0.134
Singida	-0.354	0.091		Kagera	-0.273	0.125
Tabora	-0.020	0.087		Mwanza	0.101	0.117
Rukwa	0.236	0.091		Mara	0.323	0.125
Kigoma	-0.305	0.080		Manyara	0.093	0.134
Shinyanga	0.037	0.073		Rural	-0.144	0.052
Manyara	0.290	0.107		**Education**		
Age	-0.001	0.002		Primary not complete	-0.081	0.075

Education				Primary complete	-0.035	0.064
Primary not complete	0.005	0.044		Secondary or above	-0.016	0.082
Primary complete	-0.026	0.039		Age	-0.001	0.002
Secondary or above	-0.096	0.068		Constant	-1.346	1.737
Female	-0.088	0.045				
Rural	0.149	0.045				
Land quality index	-0.159	0.025				
Share of soil type 1	-1.777	10.985				
Share of soil type 2	-1.572	10.985				
Share of soil type 3	-1.603	10.985				
Share of soil type 4	-1.532	10.985				
Share of land irrigated	0.185	0.064				
Constant	8.469	10.998				
Adjusted R^2	0.473			Adjusted R^2	0.467	

The earnings from the preliminary assignment are now compared to individual recipients' estimated contributions to see if they would take the job if it were offered. Those individuals whose earnings were less than 75% of their estimated contribution were left out of the simulation. Rates of attrition by sex and by participation in farm and nonfarm family enterprises for each country are reported in table 4, below. The overall rate of attrition is similar for both countries, but there are some large differences by sex and by activity. In Ghana, women are 25% more likely to have been dropped from the simulation than in Tanzania, while the rates for men are the same. The rates for individuals that work neither on a family farm nor in a family business are unsurprisingly low and equal in the two countries. Family farm workers are more likely to drop out of the simulation in Ghana than in Tanzania. Family farm workers were 74% of the original recipient pool in Ghana while 88% of those in the pool in Tanzania were family farm workers, but in both countries the bulk of those dropping out were family farm workers (91% in Ghana and 94% in Tanzania). This is an indicator that for many poor people in both Ghana and Tanzania, family farm work is a better option than paid employment, given what is currently available.

Table 4. Percent of Recipient Pool Dropped by Sex and Activity, Ghana and Tanzania

	Ghana	Tanzania
Male	24%	24%
Female	30%	24%
Not family worker	2%	2%
Family farm	33%	26%
Nonfarm family	87%	53%
Both farm and nonfarm	84%	33%
Total	27%	24%

The actual assignments are done within cells constructed from age, sex, and educational categories, as described above. Thus we show breakdowns of the donor and adjusted recipient pools by sex and age and by sex and education for Ghana and Tanzania in figures 1 through 4, below. In both countries, the recipient pool is younger, less well-educated, and more likely to be female than the donor pool. This result is not surprising of course, given the nature of this exercise.

Tables 5 and 6 contain the results of the job assignment in Ghana and Tanzania. Not all individuals receive their likeliest industry and occupation in the matching. This happens when there are no records in the donor cell for that combination of industry and occupation. However, in Ghana 90% of recipients got a job in their likeliest industry and 95% in their likeliest occupation. In Tanzania, the respective proportions were 87% and 85%. So for the most part, job recipients received the job we estimate they are likeliest to get.

Figure 1. Donor and Adjusted Recipient Pools by Sex and Age, Ghana

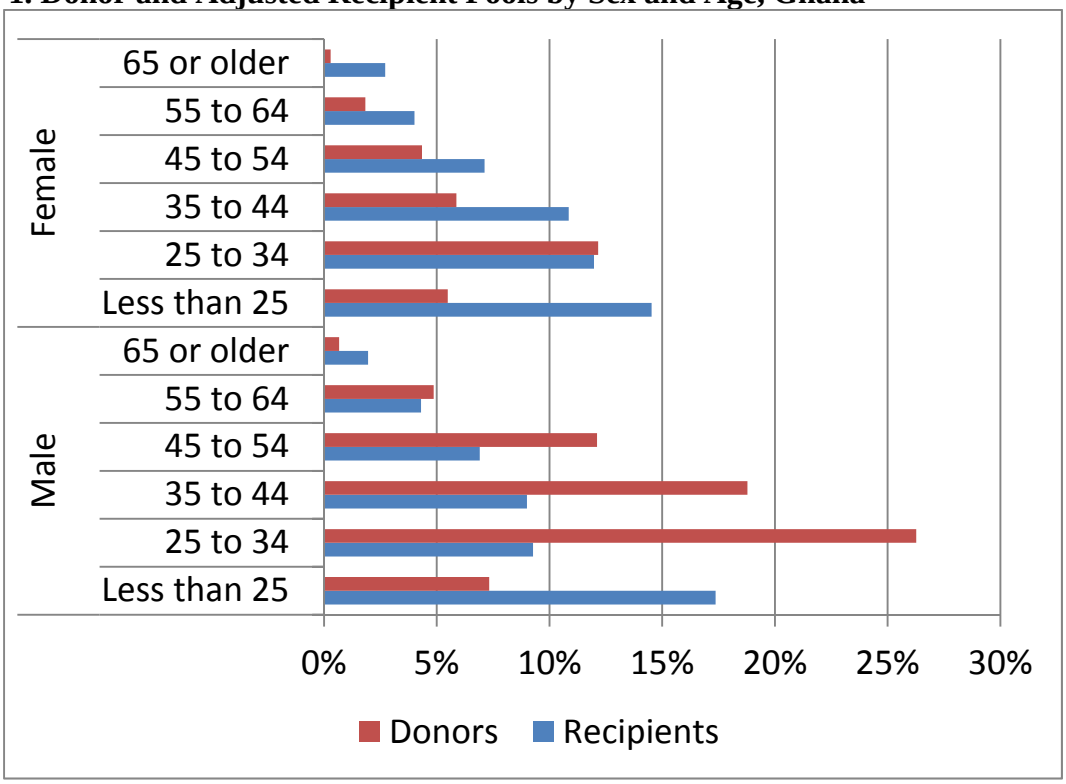


Figure 2. Donor and Adjusted Recipient Pools by Sex and Education, Ghana

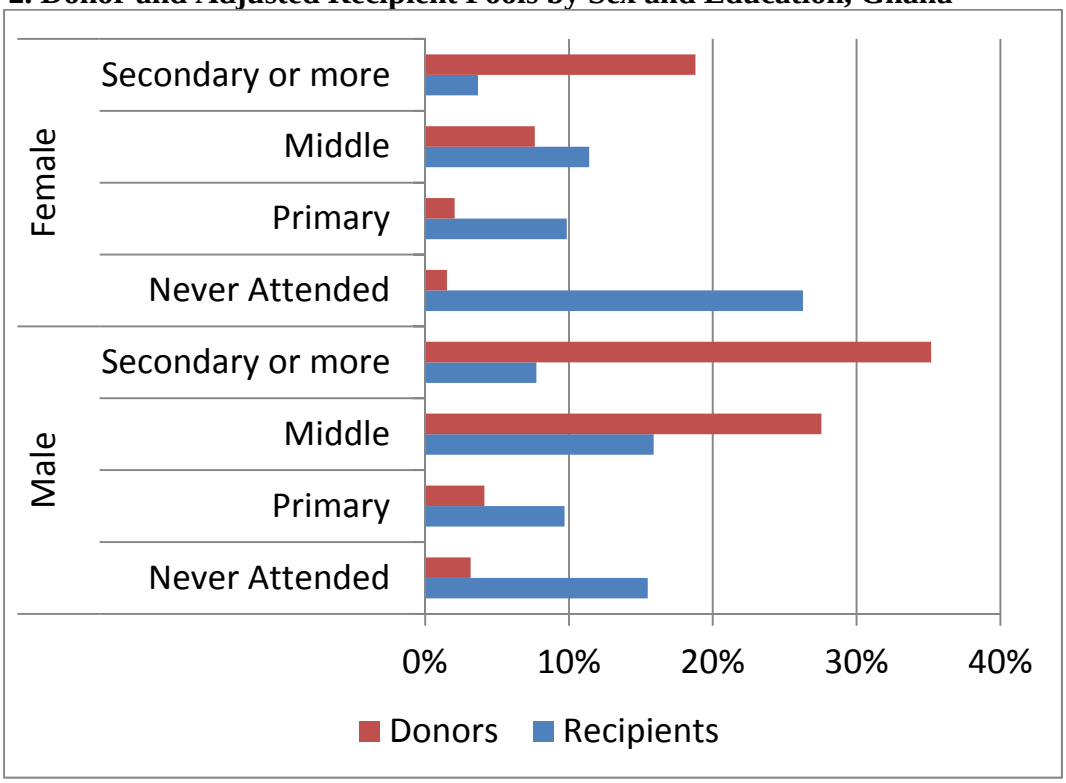


Figure 3. Donor and Adjusted Recipient Pools by Sex and Age, Tanzania

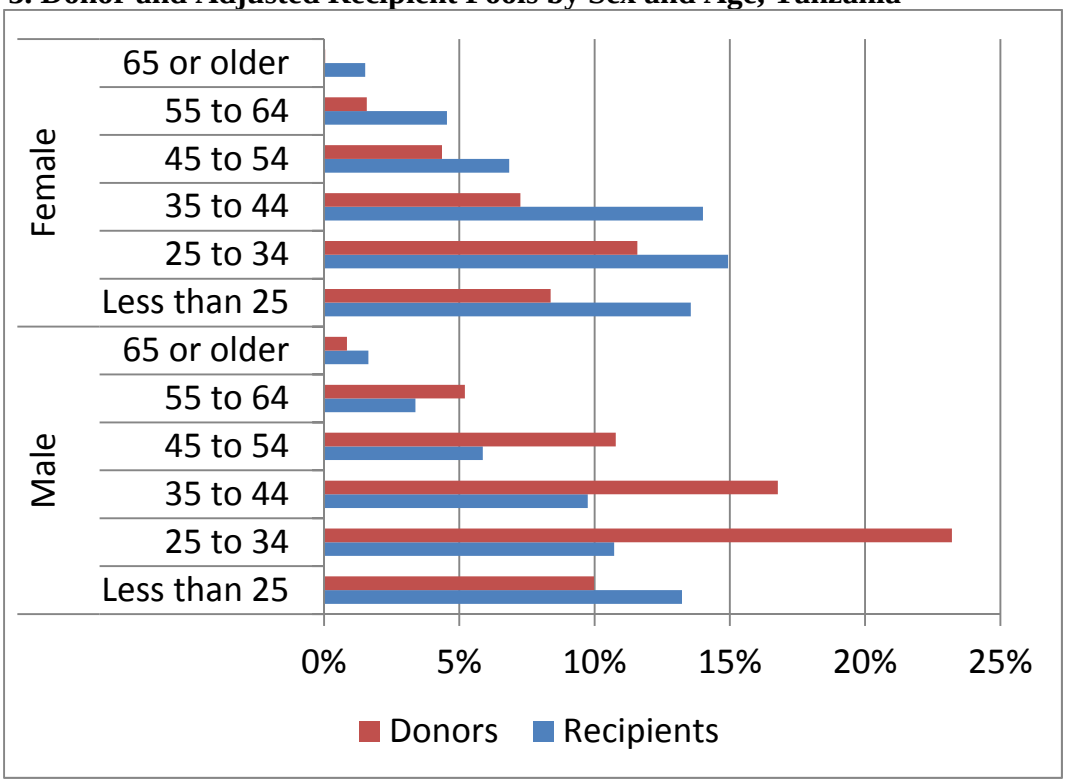


Figure 4. Donor and Adjusted Recipients by Sex and Education, Tanzania

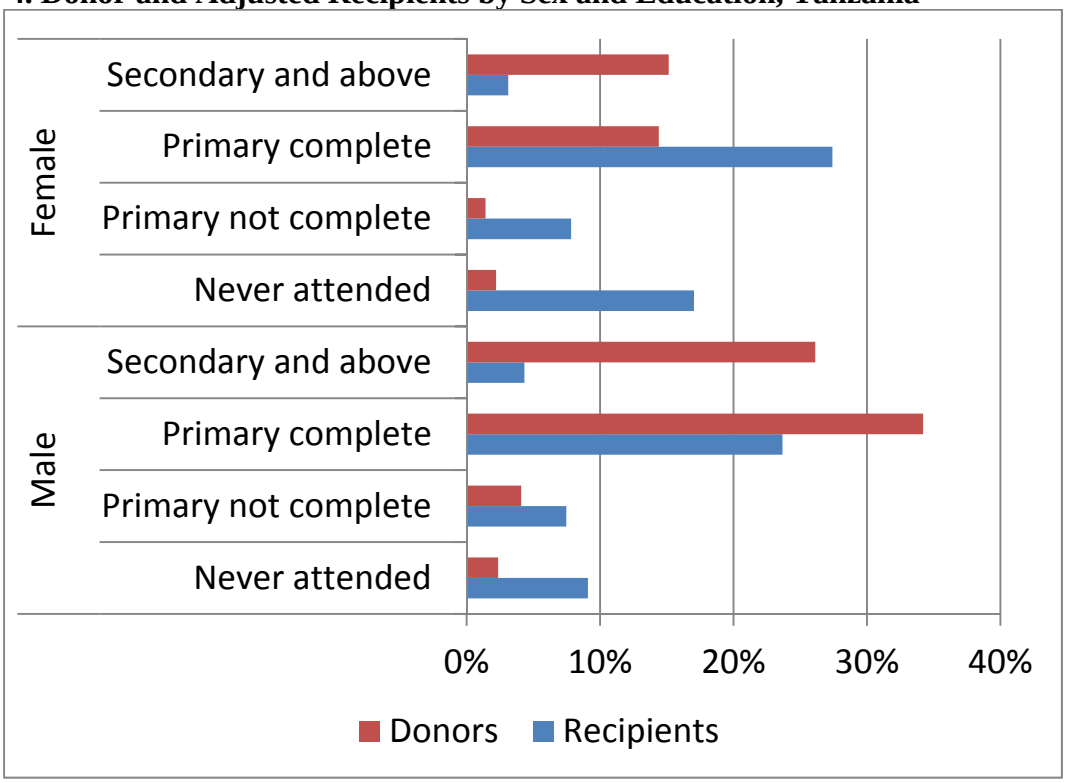


Table 5. Job Assignment Results for Ghana

Assigned Industry	Likeliest Industry					
	Agriculture, forestry, and fishing	Mining, manufacturing, and utilities	Accommodation and food services	Finance, insurance, and real estate	Education, human health, and social work	**Total**
Agriculture, forestry, and fishing	**47536**	0	3562	0	1870	**52968**
Mining, manufacturing, and utilities	1216	**65518**	1518	0	20109	**88361**
Accommodation and food services	9527	12578	**1197020**	0	66840	**1285965**
Finance, insurance, and real estate	1115	8185	0	**18648**	18000	**45948**
Professional and administrative	294	0	0	0	0	**294**
Education, human health, and social work	10595	2803	15249	0	**395043**	**423690**
Arts, entertainment, recreation, and other	397	7748	0	0	18357	**26502**
Total	**70680**	**96832**	**1217349**	**18648**	**520219**	**1923728**

Assigned Occupation	Likeliest Occupation					
	Technicians and associate professionals	Skilled agricultural, forestry, and fish	Craft and related trades workers	Plant and machine operators, and assemb	Elementary occupations	**Total**
Technicians and associate professionals	**201841**	1717	0	0	597	**204155**
Clerical support workers	74	0	0	0	0	**74**
Service and sales workers	0	0	0	0	485	**485**
Skilled agricultural, forestry, and fish	8723	**924508**	4253	726	17904	**956114**
Craft and related trades workers	0	5944	**8192**	424	1356	**15916**
Plant and machine operators, and assemb	552	26672	0	**18407**	15824	**61455**
Elementary occupations	1904	0	2728	0	**680897**	**685529**
Total	**213094**	**958841**	**15173**	**19557**	**717063**	**1923728**

Table 6. Job Assignment Results for Tanzania

Assigned Industry	Likeliest Industry										
	Agriculture, forestry, and fishing	Mining, manufacturing, and utilities	Construction	Whole-sale and retail; repair; transport	Accommodat-ion and Food services	Professional and administrative	Public Adminis-tration	Education, human health and social work	Arts, entertainment, recreation, and other	Activities of households as employers	Total
Agriculture, forestry, and fishing	**2118535**	6549	5024	1570	0	0	6834	41561	1674	474	2182221
Mining, manufacturing, and Utilities	18699	**61530**	0	0	415	0	0	22037	676	0	103357
Construction	0	0	**132439**	0	0	0	326	0	0	0	132765
Wholesale and retail; repair; transport	73876	7694	0	**269436**	1194	0	474	20319	0	6585	379578
Accommodation and food services	14281	0	0	3967	**14529**	0	0	17482	275	0	50534
Finance, insurance, and real estate	1706	0	0	0	0	0	0	1270	0	0	2976
Professional and administrative	16178	0	0	0	0	**9585**	0	1964	0	0	27727
Public administration	12946	322	0	0	0	355	**42793**	0	0	0	56416
Education, human health, and social work	76120	893	0	1963	0	0	6404	**711188**	0	0	796568
Arts, entertainment, recreation, and other	37357	1746	3518	547	0	0	0	12283	**3031**	0	58482
Activities of households as employers	78822	966	275	8324	456	0	6288	30667	407	**270730**	396935
Total	2448520	79700	141256	285807	16594	9940	63119	858771	6063	277789	4187559

Assigned Occupation	Likeliest Occupation								
	Managers	Professionals	Technicians and associated professionals	Service and sales workers	Skilled agricultural forestry and fishery workers	Craft and related trades workers	Plant and machine operators, and assemblers	Elementary occupations	Total
Managers	**18243**	0	346	0	0	0	0	0	18589
Professionals	0	**3867**	1340	21230	0	0	0	0	26437
Technicians and associate professionals	0	0	**460289**	15087	93021	1561	0	53987	623945
Clerical support workers	0	0	272	0	0	0	3354	0	3626
Service and sales workers	0	1157	13282	**1192198**	14689	11862	957	5334	1239479
Skilled agricultural, forestry, and fishery workers	2028	0	14719	45243	**736253**	0	0	110436	908679
Craft and related trades workers	326	0	0	26645	3954	**157345**	346	0	188616
Plant and machine operators, and assemblers	0	0	2189	2182	0	0	**94110**	2380	100861
Elementary occupations	6564	0	74589	66451	24598	3025	0	**902100**	1077327
Total	27161	5024	567026	1369036	872515	173793	98767	1074237	4187559

In table 7, we see the most relevant results of the jobs assignment: the earnings and hours worked. We compare the mean and median for each to the donor pool for reference. Since the recipient pool is composed of different people, we see that the earnings are different. Given the labor market conditions and the nature of the recipient pool (see above) compared to the donor pool, it is not surprising that the simulated mean and median earnings are lower than the actual earnings in both countries. It is also unsurprising that the hours worked are quite similar, as hours vary much less by industry and occupation than earnings. These differences are reflections both of the characteristics of the recipients as well as of the differential rate of return on those characteristics (see, for example, Elu and Loubert [2013]). In figures 5 and 6, below, we show the same ratios for more detailed cells (by sex and educational attainment). The ratios closest to unity correspond to the cells with the largest numbers of records in the sample. Generally speaking, individuals with greater educational attainment earn less in the simulation than their counterparts already in paid employment.

Table 7. Earnings and Hours Assignment Results

		Donor		Recipient		Ratio	
Ghana		**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**
First activity	Earnings	7563.2	4200	3465.7	2080	45.8%	49.5%
	Usual hours	51.6	48	54.0	50	104.8%	104.2%
Second activity	Earnings	263.7	0	65.9	0	25.0%	
	Usual hours	1.7	0	1.7	0	103.9%	
Tanzania							
First activity	Earnings	4,199,696	2,040,000	1,516,568	960,000	36.1%	47.1%
	Usual hours	57	56	53	56	92.6%	100.0%
Second activity	Earnings	3,024					
	Usual hours	5		6	-	130.7%	

Figure 5. Ratio of Recipients' to Donors' Mean and Median Earnings and Hours, Ghana

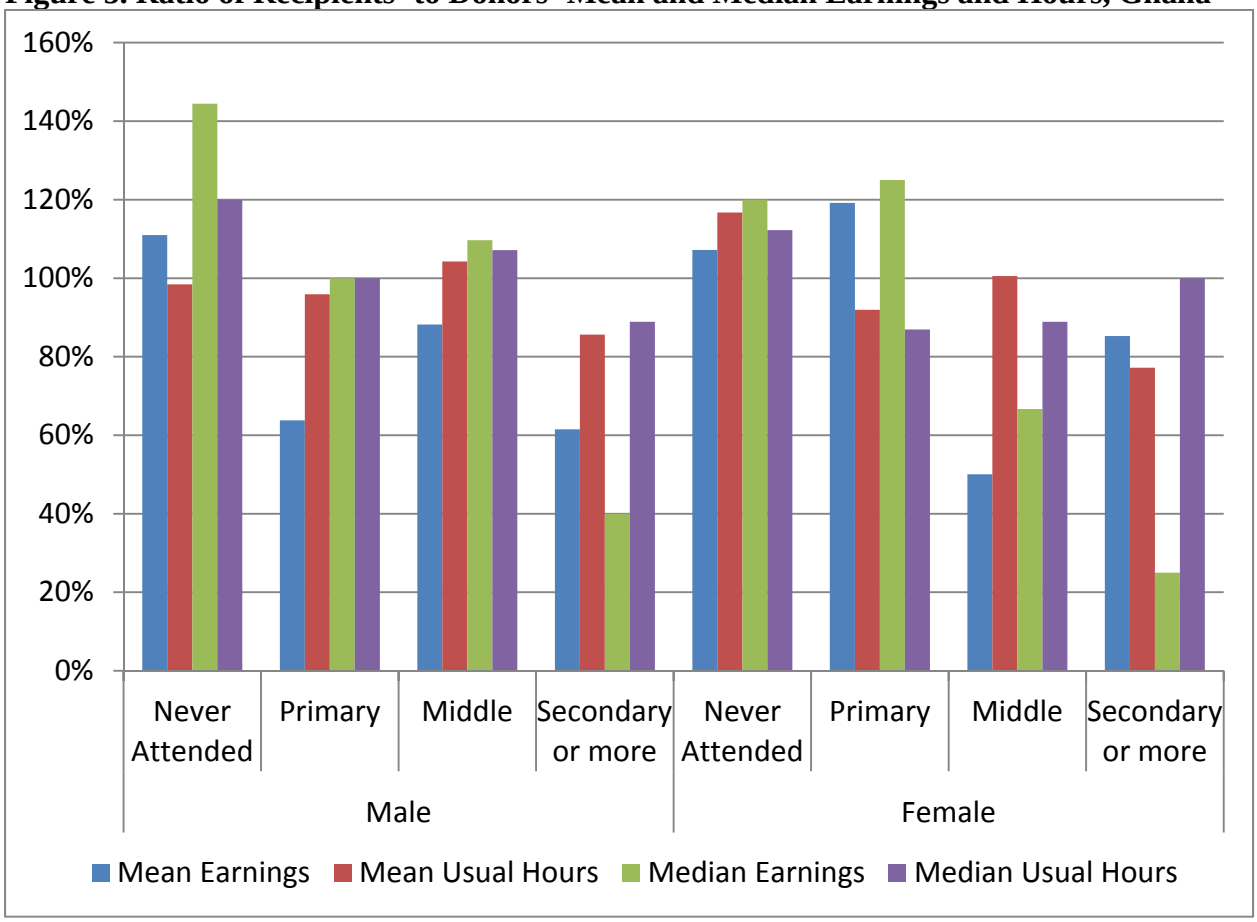
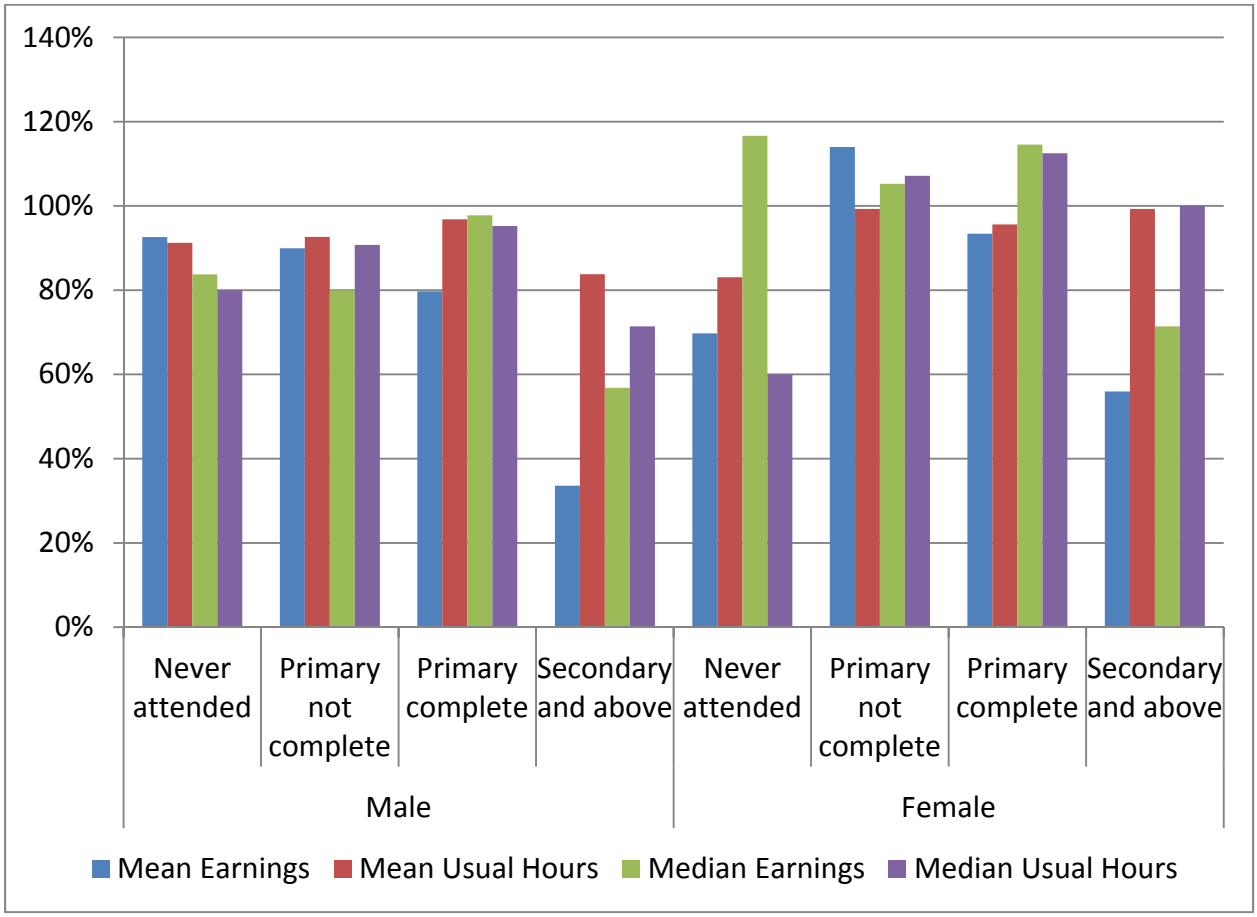


Figure 6. Ratios of Recipients' to Donors' Mean and Median Earnings and Hours, Tanzania



Figures 7 and 8 provide breakdowns of the recipient and donor pools for the time use assignment by sex and age for Ghana and Tanzania, respectively. As expected, these pools are more closely matched between recipients and donors. In Ghana, the recipient pool tends to be somewhat younger than the donor pool for both sexes. This trend is evident in Tanzania as well, but not as strongly.

Figure 7. Recipient and Donor Pools for Time Use Assignments, by Sex and Age, Ghana

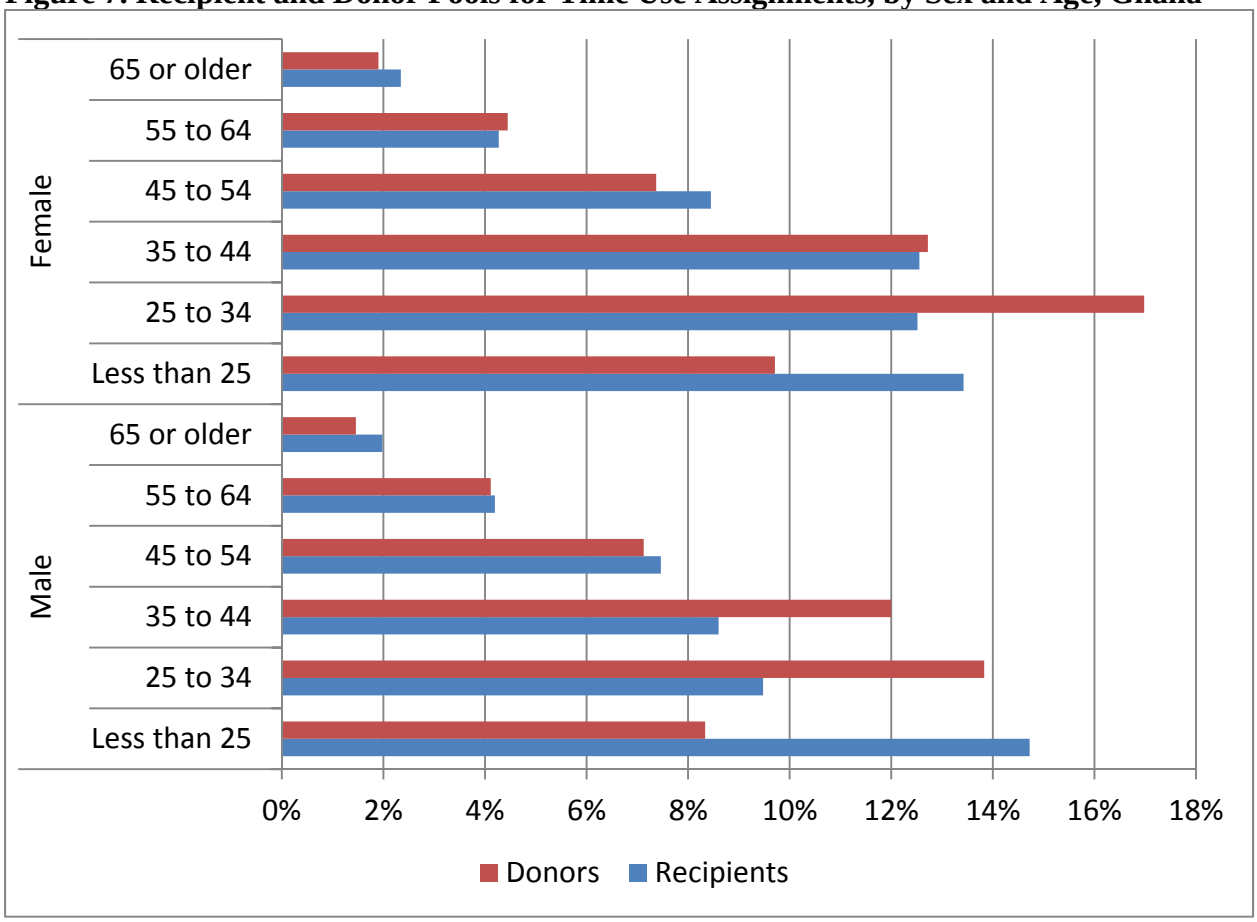
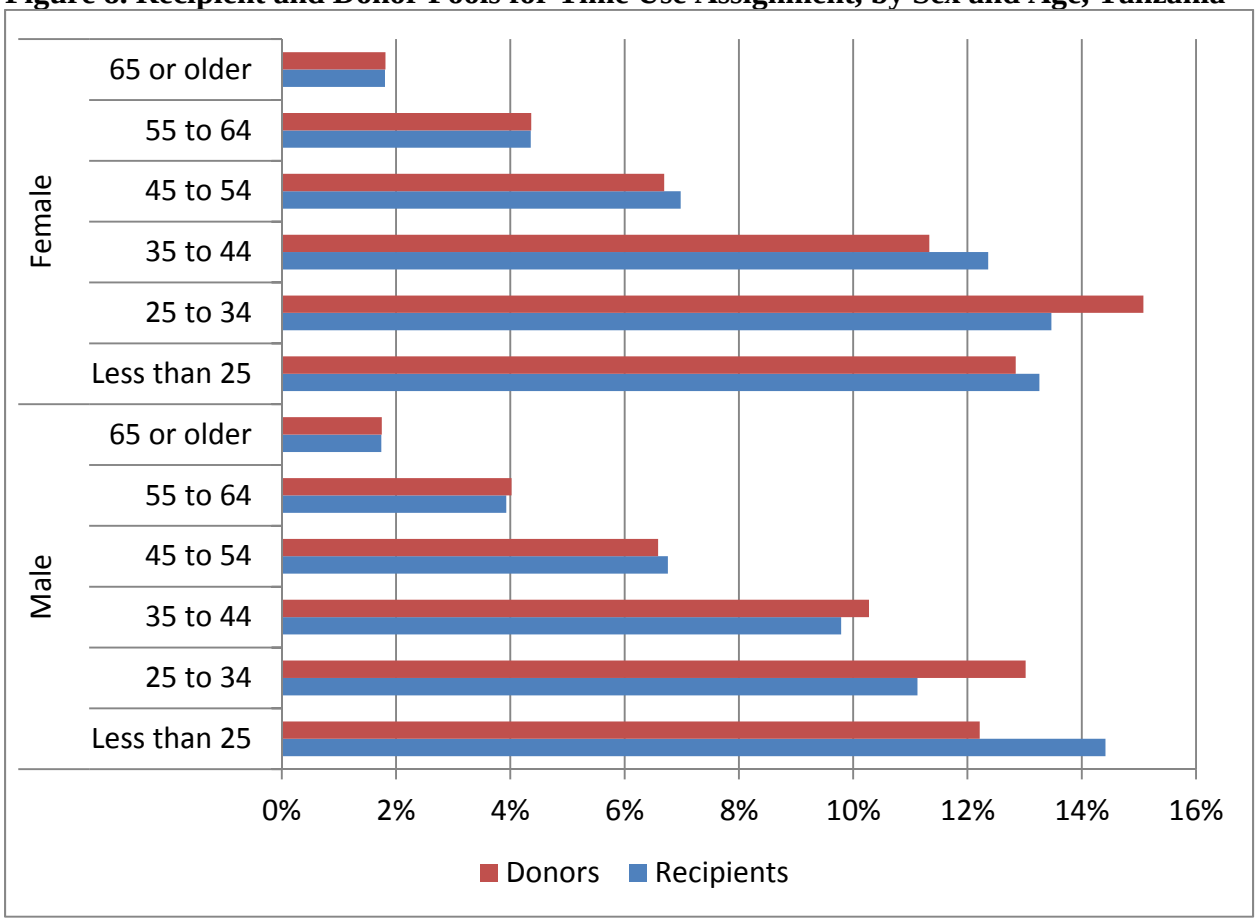


Figure 8. Recipient and Donor Pools for Time Use Assignment, by Sex and Age, Tanzania



Finally, the results of the time use reassignment are presented in figures 9 and 10 for Ghana and Tanzania, respectively. The assigned hours are very similar to the donors' hours of household production. This is especially true of women in Tanzania.

Figure 9. Ratios of Recipients' to Donors' Weekly Hours of Household Production, Ghana

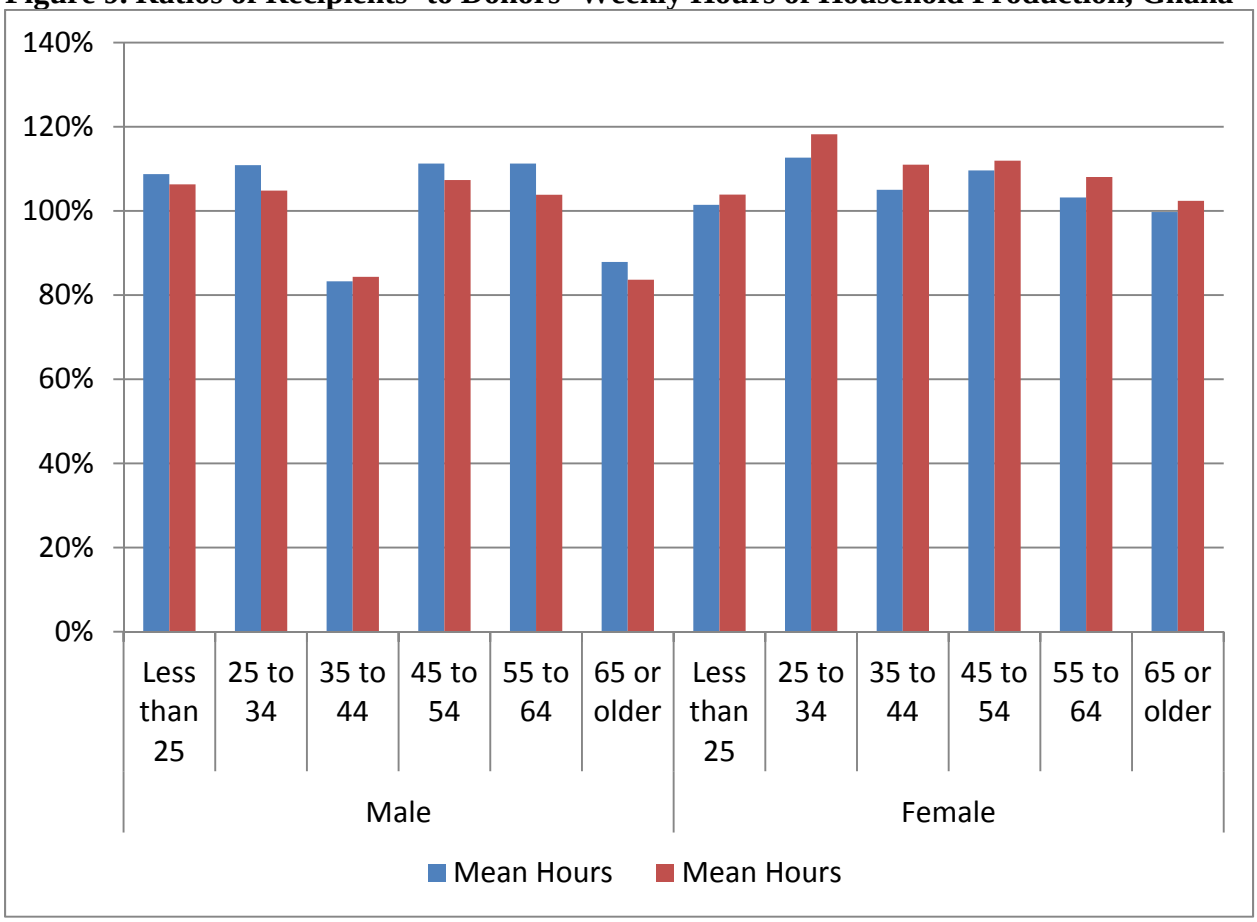
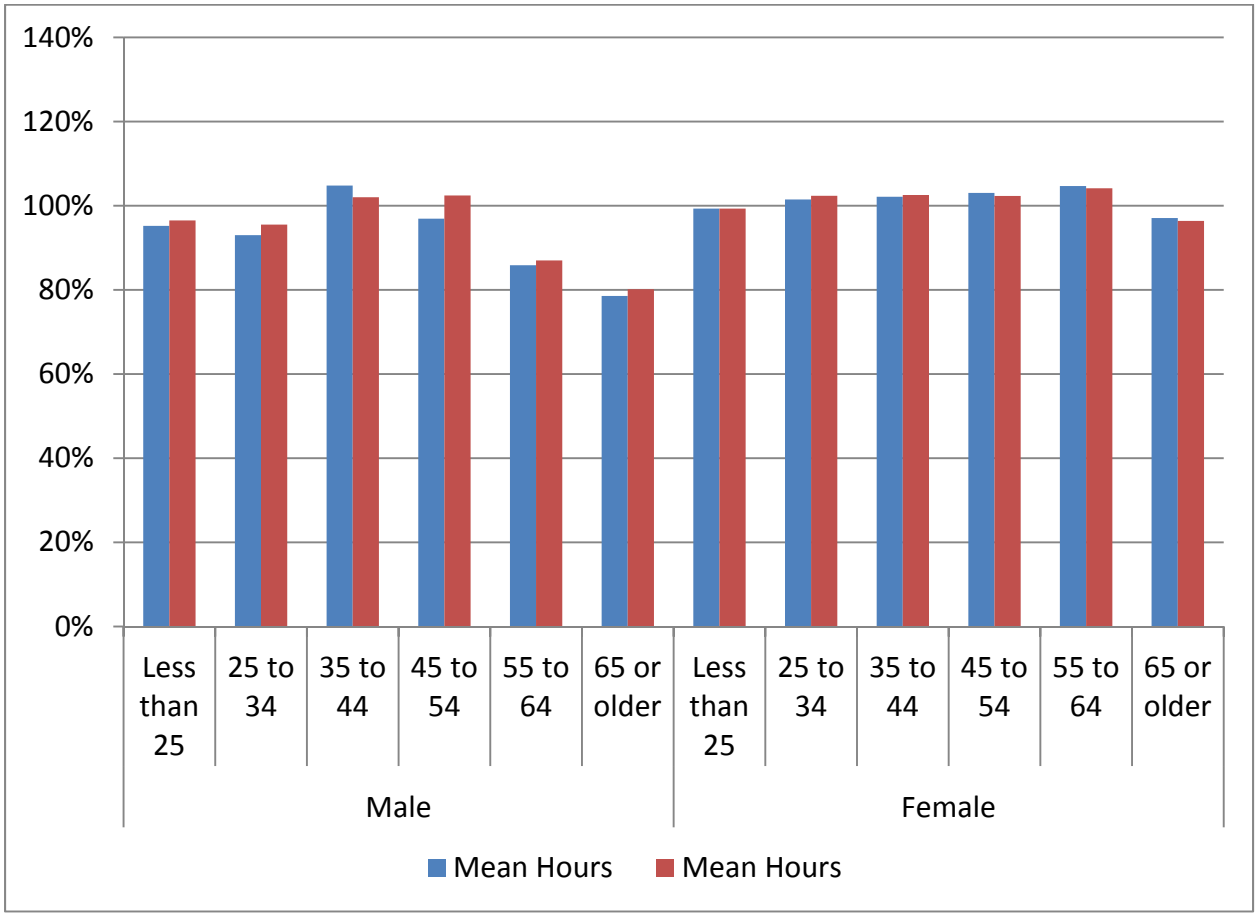


Figure 10. Ratios of Recipients' to Donors' Weekly Hours of Household Production, Tanzania



CONCLUSION

Overall, the results of the employment simulation are plausible in terms of earnings and hours worked and the reassignment of hours of household production. Most of the individuals received jobs in the simulation in their likeliest industry and occupation. Individuals in nonfarm household enterprises were more likely to be dropped from the simulation due to the fact that their simulated earnings were well below their contributions to nonfarm business output. In summary, we are confident that the results of the simulation are a plausible representation of the impact on recipient (i.e., consumption-poor) households of those not in paid employment receiving the paid work they are most likely to receive given actual labor market conditions in Ghana and Tanzania.

REFERENCES

Andridge, R.R., and R.J.A. Little. 2010. "A Review of Hot Deck Imputation for Survey Non-Response." *International Statistical Review* 78(1): 40–64. doi:10.1111/j.1751-5823.2010.00103.x.

Berndt, E.R. 1996. *The Practice of Econometrics: Classic and Contemporary*. Reading, MA: Addison Wesley.

Elu, J.U., and L. Loubert. 2013. "Earnings Inequality and the Intersectionality of Gender and Ethnicity in Sub-Saharan Africa: The Case of Tanzanian Manufacturing." *The American Economic Review* 103(3): 289–92.

Masterson, T. 2012. "Simulations of Full-Time Employment and Household Work in the Levy Institute Measure of Time and Income Poverty (LIMTIP) for Argentina, Chile, and Mexico." Levy Economics Institute Working Paper 727. Annandale-on-Hudson, NY: Levy Economics Institute of Bard College.

———. 2013. "Quality of Statistical Match and Simulations Used in the Estimation of the Levy Institute Measure of Time and Consumption Poverty (LIMTCP) for Turkey in 2006." Levy Economics Institute Working Paper 769. Annandale-on-Hudson, NY: Levy Economics Institute of Bard College.

Rios Avila, F. forthcoming. "Quality of Match for Statistical Matches Used in the Development of the Levy Institute Measure of Time and Consumption Poverty (LIMTCP) for Ghana and Tanzania." Levy Economics Institute Working Paper. Annandale-on-Hudson, NY: Levy Economics Institute of Bard College.