2005

Sources of gender difference in rural to urban migration in Kenya: does human capital matter?

Richard U. Agesa
Marshall University, agesa@marshall.edu

Jacqueline Agesa
Marshall University, AGESAJ@MARSHALL.EDU

Sources of gender difference in rural to urban migration in Kenya: does human capital matter?

Richard U. Agesa* and Jacqueline Agesa

Division of Finance and Economics, Marshall University, One John Marshall Drive, Huntington, WV 25755, USA

Using data from Kenya this article estimates the urban to rural gender gap in the rate of migration and then decomposes the gap into the explained portion and the portion due to gender differences in coefficients. The former is further decomposed to unveil the relative influence of each explanatory variable on the explained portion of the gender gap in the rate of migration. A non-trivial finding suggests that human capital variables may exert the strongest influence on gender differences in migration, partially explaining the higher incidence of male migration.

I. Introduction

Migration literature offers two rather contrasting explanations to explain gender differences in rural to urban migration in Africa. One strand suggests that institutional factors (e.g. customs and traditions) may be responsible for the relatively higher incidence of male migration (Brockerhoff and Eu, 1993; Lucas 1997). Another strand contends that economic factors (especially the larger wage gain for males) may cause gender differences in migration (Armitage and Sabot, 1991; Agesa and Agesa, 1999).

This article augments the latter literature by estimating and decomposing the urban to rural gender gap in the rate of migration into the portions due to (1) gender differences in human capital characteristics and (2) the influence of human capital characteristics. The former is further decomposed in order to identify variables that may have the greatest influence on gender differences in migration. Such an exercise is important because it would account for the incompatibility in attributes for migrant workers in urban areas and for non-migrant workers in rural areas, thus providing the sources for the larger wage gain for males as a result of migration. No other study has, to the best of the authors' knowledge, attempted such an exercise.

II. Data

This study uses data from two sources: the 1986 Urban Labor Force Survey (ULFS) and the 1988/89 Rural Labor Force Survey (RLFS). The ULFS consists of 5749 migrants (3288

*Corresponding author. E-mail: agesa@marshall.edu

male and 2461 female). The RLFS consists of 1200 male workers and 3518 female workers who have made the decision not to migrate to the urban area and considered the rural home as their permanent residence. Earnings of individuals in the ULFS and the RLFS were converted to weekly wages, and the national Consumer Price Index (CPI) taken from the *Economic Survey* (1994b), was used to deflate the weekly earnings of workers to 1986 Kenya Shillings.

III. Empirical Model

First, a reduced form probit of the migration decision, adjusted for non-randomness of the data, was estimated as follows: [1]

$$\text{Prob}(M_{igig})$$

$$=\phi(\alpha+\beta Z+\gamma(\hat{W}_{ig,u}-\ln\hat{W}_{ig,r})+\varepsilon_i) \quad (1)$$

Where Prob(Migig) is a binary variable that takes the value one of the individual is a migrant and zero otherwise. The matrix Z consists of variables thought to determine migration. Plausible candidates include age, size of rural holding, number of children at the time of migration, distance from the urban top the rural areas, the urban unemployment rate, marital status and human capital dummy variables. The vector β is the corresponding vector of estimated coefficients (partial derivatives are evaluated at the mean values of regressors to estimate marginal effects).

The earnings equations ($\ln Wi$) are specified as follows:[2]

$$\ln W_{ig,u}=\delta+\beta_u X_u+c_u\lambda_{iu}+\mu_u \quad (2)$$

$$\ln W_{ig,r}=\gamma+\beta_r X_r+c_r\lambda_{ir}+v_u \quad (3)$$

Second, the rural to urban gender gap in the rate of migration was estimated. This gap was then decomposed into the portion due to gender differences in characteristics, and gender differences in the influence of human capital characteristics.

[1]The dependent variable is binary and takes the value 1 if the individual is a migrant in the urban area and zero if the individual is a non-migrant in the rural area. The independent variables consist of variables thought to proxy the determinants of rural to urban migration without affecting wages: size of the rural land holding, number of children at the time of migration, and distance to the urban area. Strictly speaking, however, identification of the instrumental variables would have been more satisfactorily resolved if one had a larger number and variety of variables that would shift the probability of migration without affecting wages—indeed this is a perennial problem in the literature; however, the very limited nature of the data precluded such attempts by the authors.

[2]Where ln $W_{ig,u}$ (the subscript $_{ig,u}$ = the ith migrant in the urban area for each gender) is the log weekly wage for migrant workers in urban areas, for each gender. Similarly, ln$W_{ig,r}$ (the subscript $_{ig,r}$ = the ith non-migrant in the rural area for each gender) is the log weekly wage for non-migrant workers in rural areas, for each gender. The matrices X_u and X_r consist of individual characteristics for migrant workers in urban areas and non-migrant workers in rural areas for each gender respectively. These variables include age, square of age, marital status and categorical variables for education. For the education variables, the combination of individuals with no education and those with primary education constitute the base group. The vectors β_u and β_r are the regression coefficients for migrant workers in urban areas and for non-migrant workers in rural areas, respectively.

An unbiased estimator of individual participation in the migratory labour stream for each gender is:

$$\text{Prob}(X_{ig}\beta_g) = \frac{1}{n_g}\Sigma\phi(X_{ig}\beta_g) \quad (4)$$

where n_g is the sample size for each gender. Thereafter, the rural to urban gender gap in the rate of migration is estimated as follows:

$$MF\text{gap} = \text{Prob}(X_M\beta_M) - \text{Prob}(X_f\beta_f) \quad (5)$$

where *MF*gap is the rural to urban gender gap in the rate of migration (*M* indexes male and *F* indexes female).

A mathematical equivalent of Equation 5 that allows for the decomposition of the rural to urban gender gap in the rate of migration is expressed as follows:

$$MF\text{gap} - [\text{Prob}(X_m - \beta_m) - \text{Prob}(X_f\beta_m)]$$

$$+ [\text{Prob}(X_f\beta_m) - \text{Prob}(X_f\beta_f)] \quad (6)$$

The first two terms on the right-hand side of the equal sign in Equation 6 represent the portion of the gender gap in the rate migration that can be explained by gender differences in human capital characteristics. The last two terms capture the portion of the gap that can be explained by gender differences in the influence of human capital characteristics on the migration decision.

The fraction of the portion due to gender differences in characteristics (that is the former portion of Equation 6) that can be attributed to the *j*th explanatory variable is defined as follows:

$$\text{explained}_j = \text{Prob}(X_M\beta_M) - \text{Prob}(X_f\beta_M)]$$

$$\text{x} \left[\frac{(X_{jm} - X_{if})\beta_{jm}}{(X_m - X_f)\beta m}\right] \quad (7)$$

Equation 7 unveils the relative influence of the *j*th explanatory variable on the gender gap in the rate of migration. The unexplained portion of the gender gap in the rate of migration has two potential sources: gender differences in unobserved characteristics between rural workers and urban workers that may partially explain the higher incidence of male migration, and rural to urban differences in urban differences in returns to observable attributes – or the gain in

Table 1. Definition of variables

Included in the migration status equation	
S-land	Size of rural land holding (in acres).
Children	Number of children at the time of migration.
Distance	The distance between the rural and urban area (in miles).
Unemprate	The urban unemployment rate (%).
Wagediff	The difference between the wage for migrants in urban areas and non-migrants in rural areas, for each gender.
In the earnings equations	
Age	Measured in years.
O level	Ordinary level education; percentage with four years of high school education.
A level	Advanced level education; percentage with six years of high school education (two additional years of high school, post O level).
Tertiary	Percentage with teacher; technical; polytechnic, and institute of technology education. University Percentage with three or four years of university education.
Married	Percentage married.

Table 2. Estimates of probit model and earnings equations for males (corrected for sample selection bias)

	Probit model		Rural earnings (non-migrants)		Urban earnings (migrants)	
Variable	Mean (1)	Coefficient (2)	Mean (3)	Coefficient (4)	Mean (5)	Coefficient (6)
Constant	1	0.0171 (0.749)	1	−2.1711 *** (−2.664)	1	4.2392*** (8.911)
Age	30	0.0190*** (2.411)	29	0.1160*** (2.385)	33	0.0579*** (5.127)
S-land	0.3541	−0.7245*** (−2.861)	–	–	–	–
Children	3	−0.0619*** (−3.420)	–	–	–	–
Distance	125	0.0017 (0.749)	–	–	–	–
Wagediff	0.667	2.8742*** (3.257)	–	–	–	–
Unemprate			3%	0.4291 (0.127)	12.77	0.5921 (0.397)
Sqaure of age	–	–	841	−0.0012*** (−5.242)	1089	−0.0005*** (−4.522)
O level	35%	1.4693*** (2.713)	30%	1.7147*** (4.081)	38%	1.7914*** (2.878)
A level	2.2%	1.4777*** (2.497)	1.7%	1.6160 (1.563)	2.5%	1.8599*** (3.244)
Tertiary	1.5%	2.1938*** (4.925)	0.3%	1.9341*** (2.933)	1.7%	2.2667*** (7.444)
University	1.2%	2.2378*** (6.219)	0.1%	2.2619*** (2.295)	1.4%	3.5648*** (5.778)
Married	69%	0.9274*** (3.892)	60.2%	0.1141*** (2.158	73.5%	1.6854*** (3.481)
Lambda	–	–	0.298	−0.1436*** (−2.553)	0.8477	0.2395*** (3.691)
F value	144		154.77		177.65	
Adjusted R^2			0.29774		0.31447	
Total number	3288		1200		2360	

Notes: *** indicates coefficeint is significant at the 1% level.
Note that the Z statistics in the probit model and the T statistics in the earnings equations are in parentheses.

Table 3. Estimates of probit model and earnings equations for females (corrected for sample selection bias)

	Probit model		Rural earnings (non-migrants)		Urban earnings (migrants)	
Variable	Mean (1)	Coefficient (2)	Mean (3)	Coefficient (4)	Mean (5)	Coefficient (6)
Constant	1	0.0046 (0.437)	1	−0.0251*** (−6.877)	1	1.4926*** (2.681)
Age	29	0.0480*** (4.983)	28	0.2587*** (3.772)	29	0.217*** (6.206)
S-land	0.175	−0.1340*** (−3.785)	–	–	–	–
Children	3	−0.0570*** (−2.042)	–	–	–	–
Distance	127	0.0001 (0.574)	–	–	–	–
Wagediff	0.329	0.2313 (1.385)	–	–	–	–
Unemprate			3%	0.5668 (0.1287)	12.89	0.239 (0.187)
Sqaure of age	–	–	784	−0.0001*** (−3.849)	841	−0.0011*** (−4.729)
O level	30%	1.7427* (1.699)	24%	0.0793*** (2.459)	32.8%	1.5881*** (3.678)
A level	1.9%	1.6263** (2.072)	1.1%	0.7394* (1.669)	2.1%	1.4430*** (2.945)
Tertiary	0.9%	1.9388*** (2.592)	0.1%	1.8370*** (2.949)	1.4%	2.1939*** (7.198)
University	0.05%	1.9964*** (3.692)	0.02%	1.2695*** (2.595)	0.9%	3.1127*** (5.331)
Married	68%	0.4927* (1.647)	66.9%	0.117*** (3.584)	75.7%	1.0285*** (4.359)
Lambda	–	–	0.2318	0.1849 (0.023)	0.3578	0.1268* (1.756)
F value	125.88		164.7		188.4	
Adjusted R^2		0.28556		0.27145		
Total number	2461		3518		1567	

Notes: *** indicates that coefficient is significant at the 1% level; ** indicates that coefficient is significant at the 5% level; * indicates that coefficient is significant at the 10% level.
Note that the *Z* statistics in the probit model and the *T* statistics in the earnings equations are in parentheses.

Returns to observable attributes as a result of migration.[3] (Actually, unobserved and omitted variables are a part of the residual and are captured in the unexplained portion. If significant variables are omitted from the regression, or variables are measured with error, then the estimated coefficients will be biased. However, if the model is properly specified, then the unexplained portion is an unbiased estimator of the higher incidence of male migration, due to their larger gain in returns to observable attributes.)

Results and Conclusion

The definitions of the variables used are provided in Table 1. Table 2 summarizes the results of the structural probit model for males (columns 1 and 2) and the separate earnings functions for male non-migrant workers in rural areas (columns 3 and 4) and for male migrant workers in urban

[3] And as Jones (1983) demonstrates, the latter portion (that is the portion due to coefficients) cannot be decomposed further. Therefore, the validity of the interpretation of the portion due to coefficients largely depends on the adequacy in the specification of the status equations.

areas (columns 5 and 6). Table 3 displays similar results, but for females. Table 4 provides a breakdown of the sources of gender differences in migration.

Perhaps the most important finding (from Table 4) is the influence of the human capital variables on the explained portion of the gender gap in the rate of migration. Human capital variables

Table 4. Decomposition of the gender gap in the rate of rural to urban migration

Variable	Proportion
Age	0.02
Square of age	0.03
S-land	-0.11
Children	0.21
Distance	0.04
Wagediff	0.75
Unemprate	0.08
Human capital*	1.15
Total explained	2.17
Total unexplained	3.02
Total predicted change, *MF*gap	5.19

Notes: The influence of each explanatory variable on the explained portion of the gender gap in the rate of migration.

* This variable is the summation of O level, A level, tertiary and university education, that is all the human capital variables.

constitute the largest share that is 75% (1.15/2.17), of the explained portion of the gender gap in the rate of migration. This finding is consistent with a priori expectations: a 1.15 percentage point in male advantage in migration could be explained by gender differences in human capital attributes. Human capital attributes therefore may exert a considerable influence on gender differences in migration. As a remedy, society should place considerable emphasis of the education of females. Another remarkable finding is the influence of the urban to rural wage difference on the explained portion of the gender gap in the rate of migration: the rural to urban wage difference (wagediff) constitutes the second largest share, 35% (0.75/2.17), of the explained portion of the gender gap in the rate of migration. This finding suggests that the higher incidence of male migration may be positively influenced by their higher anticipated earnings in urban areas as migrants versus their anticipated earnings in the rural area as non-migrants.

References

Agesa, J. and Agesa, R. (1999) Gender differences in the incidence of rural to urban migration: evidence from Kenya, *Journal of Development Studies*, 35, 36–58.

Armitage, J. and Sabot, R. H. (1991) *Discrimination in East Africa's Urban Labor Markets, Unfair Advantage: Labor Market Discrimination in Developing Countries*, World Bank Regional and Sectoral Studies, Washington, DC.

Brockerhoff, M. and Eu, H. (1993) Demographic and socioeconomic determinants of female rural to urban migration in Sub-Saharan Africa, *International Migration Review*, 27, 557–77.

Jones, F. (1983) On decomposing the wage gap: a critical comment on Blinder's method, *Journal of Human Resources*, 18, 126–30.

Lucas, R. E. B. (1997) Internal migration in developing countries, in *Handbook of Population and Family Economics*, Vol. 1B, Elsevier Science, Oxford, pp. 721–98.