

Hours by Type of Work and Schooling in Ethiopia: Nonparametric and Parametric Estimates

Evangelos M. Falaris

University of Delaware

I estimate regressions of the relationship between time devoted to alternative types of work activities and worker characteristics using data from Ethiopia. The regressions are useful in investigating how changes in individuals' schooling are likely to affect their type of work activity which has implications for poverty reduction and their economic well-being. I estimate nonparametric regression as well Tobit models. The less restrictive nonparametric models are superior and reveal patterns in the relation of schooling and time devoted to three different work activities that are not apparent with Tobit models. The nonparametric and Tobit models result in significantly different predictions in the majority of the cases considered.

Keywords: hours by type of work, nonparametric regression, labor supply, Ethiopia

Received November 18, 2021

Accepted January 31, 2022

Corresponding Author

Evangelos M. Falaris

Department of Economics, University of Delaware

E-mail: falaris@udel.edu

JEL Codes: J22, O15, C14, C24

INTRODUCTION

Over the course of economic development, work time devoted to agricultural activities decreases and work time devoted to non-agricultural activities including self-employment and wage employment increases. What is the relationship between time devoted to different types of work and worker characteristics such as years of schooling completed? Answering this question is useful in predicting labor supply by type of work, and thus changes in economic structure. This is important given changes in worker schooling that are likely to take place as the economy develops or in response to policy changes. In many developing countries a shift of workers from agriculture to non-agricultural self-employment and from both of these activities to wage employment indicates a shift to better paying and more stable work (Fields, 2011). Exploring how an increase in schooling may affect the supply of different types of work has implications for poverty reduction and increased worker well-being.

In this paper I use activity-specific work time data from Ethiopia to study the relation between the times devoted to three work activities and worker characteristics. The three work activities are household agricultural work, nonagricultural household business work¹ and wage work. Previous studies, e.g. Malathy (1994) and Kimmel and Connelly (2007), studied time allocation between household production, leisure and work in India and the United States, respectively. They estimated Tobit models for the cases in which the dependent variables were censored. Glick and Sahn (1997) studied choice of sector of employment and earnings in Guinea. They utilize a parametric multiple discrete choice model with selectivity. The models in these previous studies rely on functional form assumptions that may be unrealistic. If the functional forms assumed in these previous studies

¹ This corresponds to non-agricultural self-employment in Fields' (2011) classification.

are not adequate characterizations of the data, their estimates may be biased and inconsistent. Since economic theory does not provide guidance on the appropriate functional form, this paper employs a less restrictive nonparametric framework.² To make comparisons I also estimate Tobit models for hours of work. Specification tests indicate that the Tobit models are misspecified, providing support for the less restrictive nonparametric models. The nonparametric models reveal patterns in the effects of schooling on hours of work that are not captured by parametric models. Predictions based on the nonparametric models differ significantly from those based on the Tobit models in the majority of the cases considered. All in all, nonparametric regression is a superior method of modeling labor supply by type of work.

FRAMEWORK

Consider an individual who allocates time to various work activities in order to maximize a utility function whose arguments are consumption and leisure. The individual faces budget constraints that reflect potential hourly wage rates in each type of work activity adjusted for the probability of being employed in each type of work activity. I will treat the probabilities of being employed in each type of work activity as exogenous. An additional constraint faced by the individual is that time devoted to all work activities must add up to total time the individual could potentially devote to work. From the first order conditions of this constrained maximization problem, supply functions may be derived of hours per week the individual devotes to each type of work activity as a function of expected wages in that activity.

In the empirical implementation of the model there are several issues with the measurement of expected wages: wages for the self-employed are not reported in the Ethiopian data, wages for those employed in wage work may be measured with error³ and the individual probability of being employed in each type of work activity is not observed. To overcome these issues I adopt a reduced form approach and use highest grade of schooling completed, measured in years, and age as regressors in the labor supply functions for each type of work activity.⁴ Both regressors are observed for everyone in the sample and are likely to be correlated with productivity in a given activity. The regressors may be regarded as capturing income and substitution effects on labor supply. They are also likely to be associated with ease of access to particular work activities. The regressors are those one would find in a semi-logarithmic wage equation. Finding valid instruments that are uncorrelated with work hours for predicting wages is a challenge, thus the reduced form approach is preferable.

For a given type of work activity the model to be estimated is $y_i = f(x_i) + \varepsilon_i$ with $E(\varepsilon_i | x_i) = 0$ where y_i denotes the outcome variable, x_i is the vector of regressors and $i = 1, \dots, N$ denotes observations in the data. From the two expressions it follows that $E(y_i | x_i) = f(x_i)$. Thus $f(x_i)$ is the conditional mean function of the outcome.

Theory offers no guidance on the specific form the conditional mean function may take. Since the data on work hours are censored at zero, a common practice in the past literature is to use a Tobit model to estimate a supply function of work hours. However, the Tobit model incorporates distributional and functional form assumptions that may be unrealistic. Specifically departures from normality, linearity and homoscedasticity may result in the Tobit model being an unrealistic characterization of the data. To investigate this issue further, I model hours devoted to various types of work activity by nonparametric regression.

² Other studies that use nonparametric methods to estimate labor supply functions include Blomquist and Newey (2002) and Liang (2012) who use data from Sweden and model hours of work in the face of nonlinearity of the budget set arising from the tax system and employ a power series estimator, and Thakral and Tô (2021) who study the labor supply of New York City cabdrivers and estimate local linear nonparametric regressions for the probability of stopping work. In contrast to these studies that use high-income country data, the present study uses data from a developing country and examines hours of work for three distinct types of economic activity.

³ Wages are reported over different time frames by different workers and in order to convert them to a uniform time period (e.g. hourly) they must be divided by hours worked during the relevant time period which are likely to be measured with error.

⁴ I do not include regressors that describe household structure because they are likely to be endogenous.

METHOD

I estimate models that relate hours of work in each of three work activities to worker characteristics by local linear nonparametric regression (Cameron and Trivedi, 2005: ch. 9.6.2; Li and Racine, 2007). The conditional mean function $f(x_i)$ is not specified *a priori* and local linear nonparametric regression involves using the data to estimate it. Specifically $E(y_i|x_i=x)$ is estimated at all data values x . This involves minimization of the following:

$\min \beta \sum_{i=1}^n [y_i - \beta_0 - \beta_1'(x_i - x)]^2 K(x_i, x, h)$ with $\beta = (\beta_0, \beta_1)'$. The conditional mean of the outcome at the value x is β_0 and the derivative of the mean function with respect to x is β_1 . K is a kernel function that weights the observations so that those with x_i closer to x receive more weight. The bandwidth h imparts smoothness to the estimated function. This procedure is repeated for each value x in the data and the conditional mean function and its derivatives are estimated.

In implementing the local linear nonparametric regression, the observations around a data point are weighted using an Epanechnikov kernel function whose bandwidth is determined by cross validation.⁵ Nonparametric regression makes it possible to estimate the conditional mean of the regressand (work hours) and derivatives of the mean function at specific values of the regressors. I present the estimates in summary form as mean predictions and average marginal effects of each regressor on hours of work in an activity. I estimate separate models by gender in order to explore differences in work activity.

Nonparametric regression is robust to functional form misspecification. In particular it does not assume *a priori* a functional form that relates the regressors to the regressand. If a functional form as in Tobit or linear regression models is assumed but is not an appropriate characterization of the data, the regression estimates may be biased and inconsistent (Henderson and Souto, 2018). Nonparametric regression also makes it possible to detect patterns in the regression relation that may be difficult to detect with a parametric framework. In particular, nonparametric regression may reveal nonlinearity relationships that may be difficult to capture by commonly used functional forms in parametric regressions such as the quadratic. I obtain estimates of standard errors by bootstrapping. For comparison purposes I also obtain Tobit estimates and compare them to the nonparametric estimates.⁶ I carry out tests of assumptions that must be valid for the Tobit model to be consistent.

RESULTS

Data and Variables

I utilize data from the 2015–2016 Ethiopian Socioeconomic Survey. The survey is a joint venture of the Ethiopian Central Statistical Agency and the World Bank and it includes both rural and urban households.

The survey includes information on individuals' time allocation during the past week.⁷ It measures time devoted to five work activities: work on a household farm or caring for livestock; work on a non-agricultural household business; work for wages (not temporary); temporary, casual or part time work; and, unpaid apprenticeship. The two latter categories are relatively sparse⁸ which poses a problem with nonparametric regression in that there are too few observations for some values of the regressors. In view of this problem, I aggregate temporary, casual, part time work,

⁵ The smoothing parameter is chosen so that it minimizes the mean squared error of the dependent variable minus the mean function evaluated leaving out that observation.

⁶ I did not estimate parametric selectivity models (unlike Glick and Sahn, 1997) because appropriate instruments for the selection equation were not available. Identification of selectivity models relying on functional form only is often fragile and it is unclear if such models are an improvement relative to models that do not control for selectivity.

⁷ The survey includes information on time devoted to five work activities and a limited number of household tasks. A fuller examination of time devoted to all household production and leisure is not feasible with these data.

⁸ Seven percent of men and 3.3 percent of women report nonzero hours devoted to temporary, casual, part time work and 0.6 percent of men and 0.3 percent of women report nonzero apprenticeship hours, respectively.

apprenticeship and wage work into one category. From this point on I will refer to the combined category as wage work. An apprenticeship typically leads to regular paid work and temporary or part time wage work may later turn into full-time wage work. Nonparametric regression is feasible for the three categories of work for both genders.⁹ The survey reports work time data as zeroes in cases of non-participation in an activity. The work time data are heavily censored at zero, as reported in Table 1, and zero work hours as an outcome are included in the samples used in the estimation. The edited sample includes individuals age 24-64. I do not include younger and older individuals in order to reduce censoring due to school attendance at younger ages or withdrawal from the labor force at older ages.

The edited sample includes 3875 men and 4566 women. Table 1 reports summary statistics. I do not explicitly deal with engaging in more than one of the three work activities in the models. I include the hours that multiple-activity workers devote to each activity in the same manner as hours of those engaging in one activity only. However, multiple work activity is not a major phenomenon in the data with 6.9 percent of individuals¹⁰ reporting spending time on more than one work activity. The rate of non-employment is overall 33.9 percent (22.1 for men, 44 for women).¹¹ It is higher by over ten percentage points in urban than in rural areas¹² reflective of the phenomenon of higher urban than rural non-employment that has been documented in several developing countries (Fields, 2011). A notable characteristic of the sample is the low level of schooling which, on average, is less than completing the primary level and is lower by over two years for women than for men.¹³ Also notable is that 38 percent of men and 64 percent of women report having no schooling.

Table 1. Summary statistics

Variable	Men	Women
Schooling (years)	5.091 (5.314)	2.871 (4.632)
Age	39.136 (10.841)	38.480 (10.846)
Hours household agricultural work	10.925 (16.671)	6.585 (12.673)
Proportion censored (zero)	0.557	0.665
Hours non-agricultural household business work	4.854 (14.865)	4.088 (13.508)
Proportion censored (zero)	0.865	0.904
Hours wage work	7.781 (18.361)	3.630 (12.944)
Proportion censored (zero)	0.799	0.871
Sample size	3875	4566

Note: Standard deviations are in parentheses.

Estimates and Predictions

Table 2 reports nonparametric estimates of average marginal effects of the two slope regressors (schooling, age) on hours of work for each of the three work activities separately by gender. Table 2 also reports mean predicted hours for each activity derived from the nonparametric models. Table 3 reports the corresponding estimates derived from Tobit models with the exact same regressors as the nonparametric models. Purpose of the Tobit estimates is

⁹ Forty-four point three (44.3) percent of men report nonzero hours of agricultural work, 13.6 percent report nonzero hours of work in a non-agricultural household business and 20.1 percent report nonzero hours in wage work. The corresponding proportions for women are 33.5, 12.9 and 9.6 percent.

¹⁰ Eight point nine (8.9) percent of men, 5.1 of women. Glick and Sahn (1997) report a multiple work activity rate of under 5 percent in urban Guinea; a rate somewhat lower than that for Ethiopia.

¹¹ Glick and Sahn (1997) report nonparticipation rates in the labor force of about 40 percent for women age 15-65 and for men 21-30 in Guinea.

¹² In rural areas 17.8 percent of men and 39.5 percent of women report spending zero hours in all three work activities. The corresponding percentages for urban areas are 32 for men and 52.9 for women.

¹³ Considering the low level of completed schooling, I will treat it as primarily determined by decisions made by the families of the individuals when they were growing up and thus not primarily determined by choices made by the individuals themselves.

to compare them to the nonparametric estimates. Tobit models, the most commonly used parametric models for censored endogenous variables, are computationally less demanding than the nonparametric models. However, the assumptions necessary for the Tobit estimates to be consistent may not hold and nonlinearities may be harder to detect than with the nonparametric models.

Based on the nonparametric models I estimate that, on average, an increase in schooling decreases hours supplied to household agricultural work by men. An increase in schooling decreases less, by a factor of approximately forty percent, hours supplied by women. I estimate that an increase in age does not significantly affect hours of household agricultural work for men and that it decreases hours for women. The estimated effects are consistent with men moving out of household agricultural activity at a faster rate than women as their schooling increases. The estimates of Tobit average marginal effects of schooling, in Table 3, reveal a generally similar pattern to the nonparametric ones and exceed them in absolute value.

On average, I estimate that an increase in schooling increases hours of work supplied to non-agricultural household businesses by men. I do not find a significant effect of age on work in non-agricultural household businesses for men. I also estimate no significant effect of schooling or age on hours of work supplied to non-agricultural household businesses by women.

The estimates of the average marginal effects of schooling derived from the Tobit models are positive for men and for women. The Tobit estimated average marginal effects are smaller than those derived from the nonparametric models. The Tobit estimated average marginal effects of age are not significantly different from zero for both genders.

For work for wages an increase in schooling implies, on average, an increase in hours of work for men and for women. Both average marginal effects of schooling are greater than the corresponding ones for non-agricultural household business work. Hours of wage work increase with age for men¹⁴ but not for women. The corresponding average marginal effects of schooling derived from the Tobit models are similar to those derived from the nonparametric models for both genders. In the Tobit models there are no significant age average marginal effects for either gender.

Table 2. Estimates of average marginal effects of slope regressors, Nonparametric hours of work models

	Men	Women
Hours household agricultural work		
Schooling	-1.055 (0.057)	-0.613 (0.030)
Age	-0.030 (0.028)	-0.041 (0.020)
Mean predicted hours	10.910 (0.274)	6.582 (0.192)
Hours non-agricultural household business work		
Schooling	0.340 (0.051)	0.353 (0.221)
Age	0.002 (0.023)	0.034 (0.035)
Mean predicted hours	4.598 (0.233)	4.020 (0.199)
Hours wage work		
Schooling	1.052 (0.064)	0.861 (0.064)
Age	0.048 (0.027)	0.014 (0.016)
Mean predicted hours	7.702 (0.297)	3.681 (0.198)

Note: Bootstrapped standard error estimates are in parentheses (based on 400 replications).

Table 3. Estimates of average marginal effects of slope regressors and mean predicted hours conditional on positive hours, Tobit hours of work models

	Men	Women
Hours household agricultural work		
Schooling	-1.359 (0.101)	-1.010 (0.092)
Age	-0.050 (0.024)	-0.056 (0.018)
Mean predicted hours	10.755 (0.562)	6.542 (0.420)
Hours non-agricultural household business work		
Schooling	0.232 (0.059)	0.172 (0.051)
Age	-0.009 (0.022)	-0.013 (0.019)
Mean predicted hours	4.633 (0.369)	3.976 (0.296)
Hours wage work		
Schooling	0.899 (0.081)	0.589 (0.060)
Age	-0.013 (0.029)	-0.008 (0.016)
Mean predicted hours	7.338 (0.490)	3.406 (0.268)

Note: Standard error estimates adjusted for clustering by enumeration area are in parentheses (326 clusters). The predictions are $E[y^*|y>0]$.

¹⁴ At the 10 percent significance level.

A potential distinction between the nonparametric and Tobit models is that the former may reveal patterns in the effects of schooling on hours of work that are not apparent with the latter. Predictions of hours of work in each activity for a range of values of years of schooling (zero to 19 years, the range of observed values), averaged across the data, for each gender are helpful in a fuller characterization of the relationship between schooling and hours of work in each activity. Figures 1-12 show such predictions. The vertical axis of the Figures shows hours of work and the horizontal axis shows schooling measured as highest grade completed. The order of the Figures is in pairs: one Figure is based on the nonparametric model the immediately following Figure is based on the corresponding Tobit model. Each Figure also depicts 95 percent confidence intervals estimated by bootstrapping.¹⁵

The relationship of hours of household agricultural work and years of schooling predicted with the nonparametric models is nonlinear. Hours of household agricultural work decrease with schooling and asymptotically approach zero at higher levels of schooling for both genders, as shown in Figures 1 and 3. Increases in schooling are likely to open up more attractive opportunities outside agriculture, reducing the hours devoted to household agricultural work. The predictions based on the Tobit model, shown in Figures 2 and 4, are that hours of household agricultural work decline monotonically and nonlinearly with schooling for both genders. The nonparametric models predict a less steep decline in hours devoted to household agricultural work at intermediate levels of schooling relative to the Tobit models.

For work in non-agricultural household businesses, the nonparametric models predict an inverted U pattern of hours with the maximum at 10 years of schooling for men and 8 years of schooling for women, as shown in Figures 5 and 7. These predicted effects suggest increasing access and rewards to non-agricultural household business work, at the expense of household agricultural work, up to these indicated schooling levels but easier access to more desirable wage employment at higher levels of schooling. This reflects the pattern of welfare-improving work activity shifts described by Fields (2011) taking place in Ethiopia as schooling increases. In contrast, the Tobit models predict monotonically increasing hours of work with schooling for both genders, but a relatively flatter hours profile for women, as shown in Figures 6 and 8.

For work for wages the nonparametric models reveal nonlinearities that are not apparent with the Tobit models. For example, the hours-schooling relationship for men levels off at higher levels of schooling (Figure 9). I predict a similar pattern for women (Figure 11). The implication is that at post-secondary schooling levels income and substitution effects on hours of wage work offset each other. In contrast, the Tobit model predicts monotonically increasing hours of wage work with schooling for men and monotonically decreasing hours for women, as shown in Figures 10 and 12.

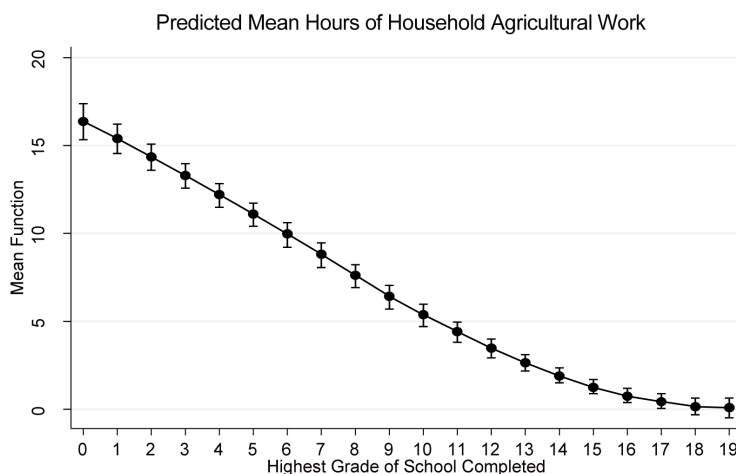


Fig. 1. Men household agricultural work, Nonparametric
Note: Ninety-five (95) percent confidence intervals are included in the Figure.

¹⁵ There are too few observations with schooling of seventeen years or more for the reliable estimation of prediction standard errors at the highest levels of schooling.

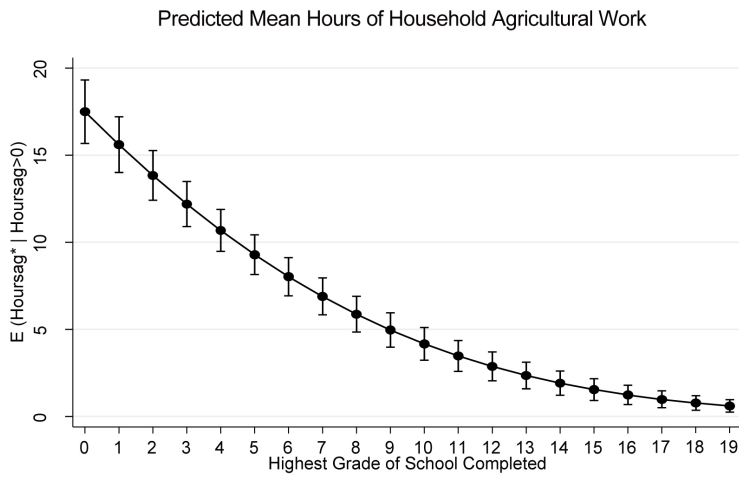


Fig. 2. Men household agricultural work, Tobit, $y^*>0$

Note: Ninety-five (95) percent confidence intervals are included in the Figure.

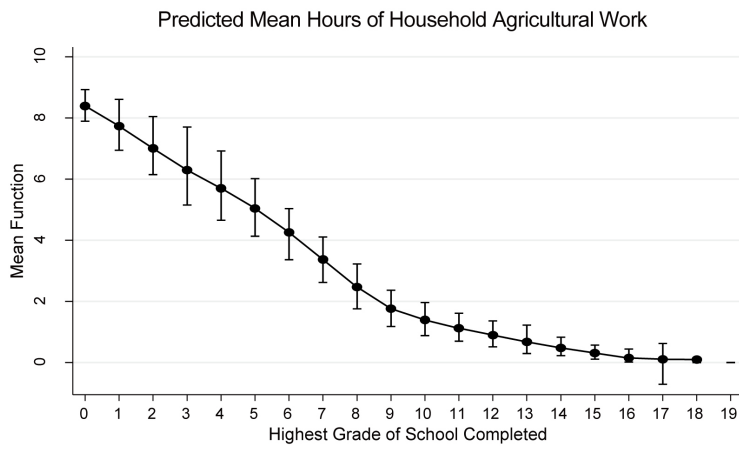


Fig. 3. Women household agricultural work, Nonparametric

Note: Ninety-five (95) percent confidence intervals are included in the Figure.

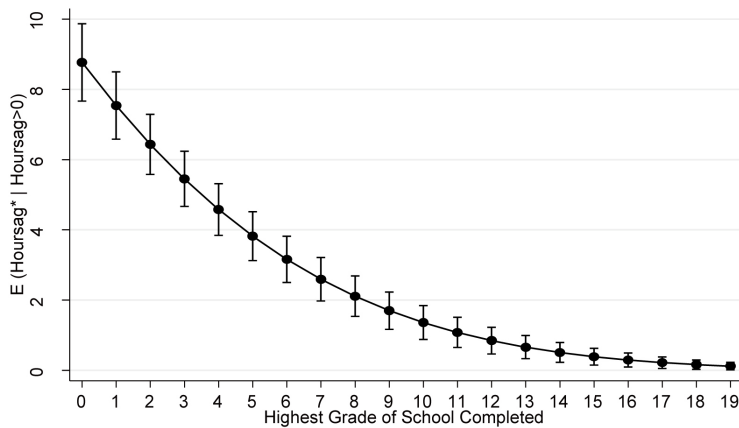


Fig. 4. Women household agricultural work, Tobit, $y^*>0$

Note: Ninety-five (95) percent confidence intervals are included in the Figure.

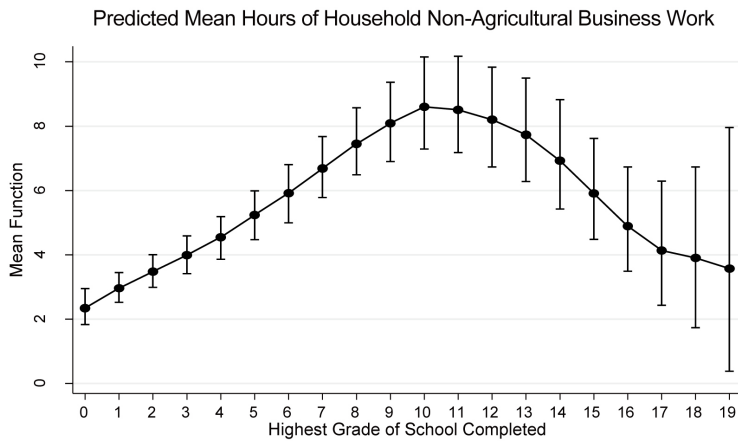


Fig. 5. Men non-agricultural household business work, Nonparametric

Note: Ninety-five (95) percent confidence intervals are included in the Figure.

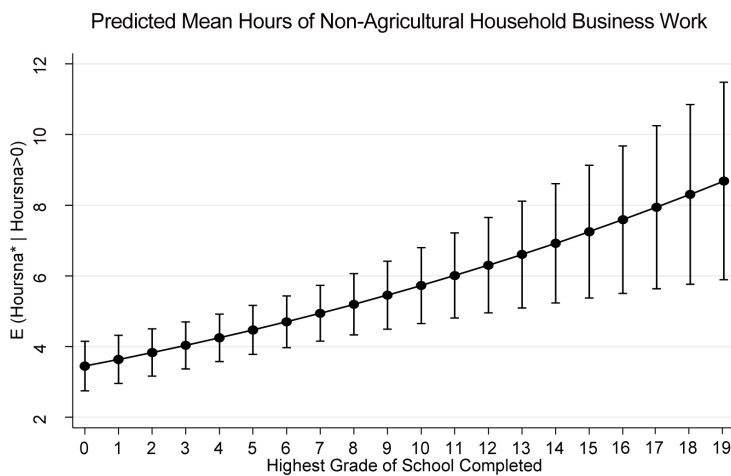


Fig. 6. Men non-agricultural household business work, Tobit, $y^* > 0$

Note: Ninety-five (95) percent confidence intervals are included in the Figure.

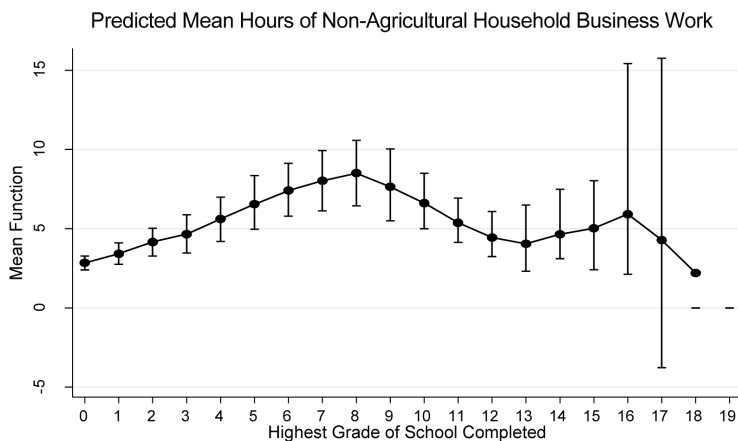


Fig. 7. Women non-agricultural household business work, Nonparametric

Note: Ninety-five (95) percent confidence intervals are included in the Figure.

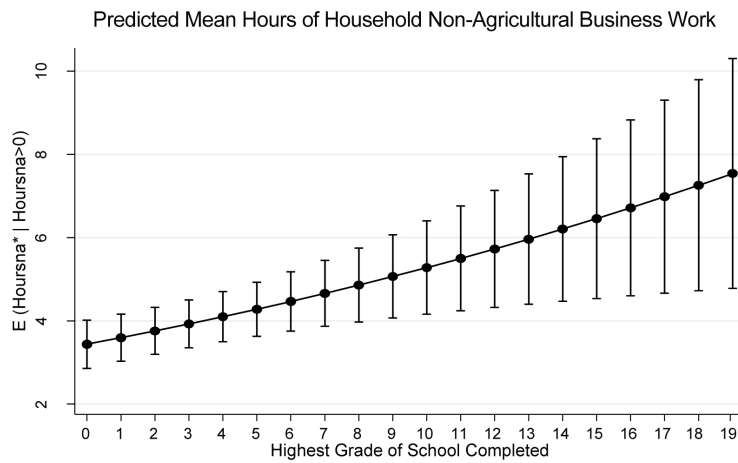


Fig. 8. Women non-agricultural household business work, tobit, $\gamma^* > 0$

Note: Ninety-five (95) percent confidence intervals are included in the Figure.

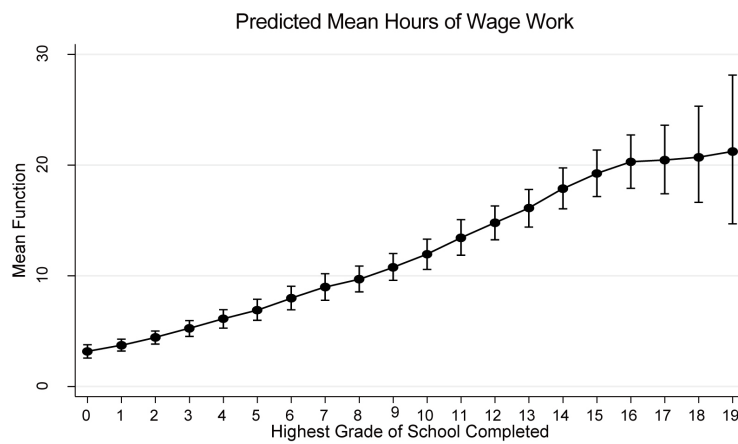


Fig. 9. Men, wage work, Nonparametric

Note: Ninety-five (95) percent confidence intervals are included in the Figure.

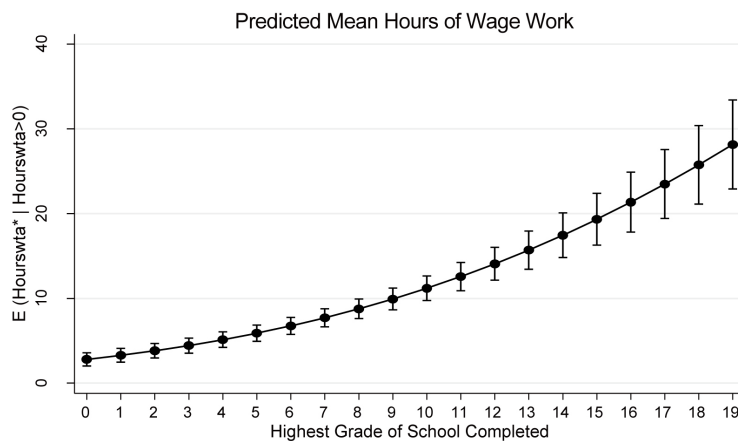


Fig. 10. Men, wage work, Tobit, $\gamma^* > 0$

Note: Ninety-five (95) percent confidence intervals are included in the Figure.

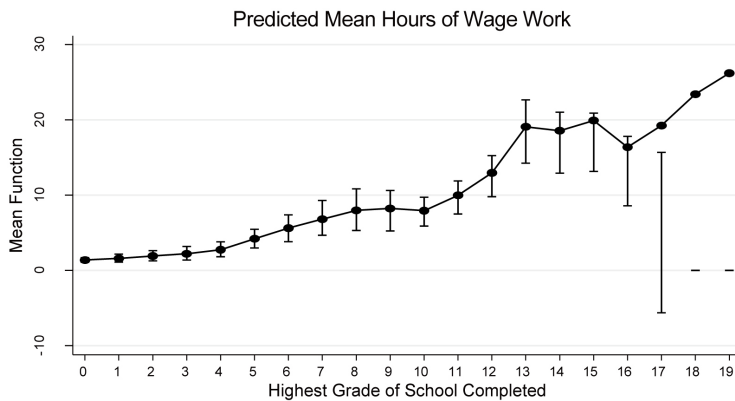


Fig. 11. Women, wage work, Nonparametric
Note: Ninety-five (95) percent confidence intervals are included in the Figure.

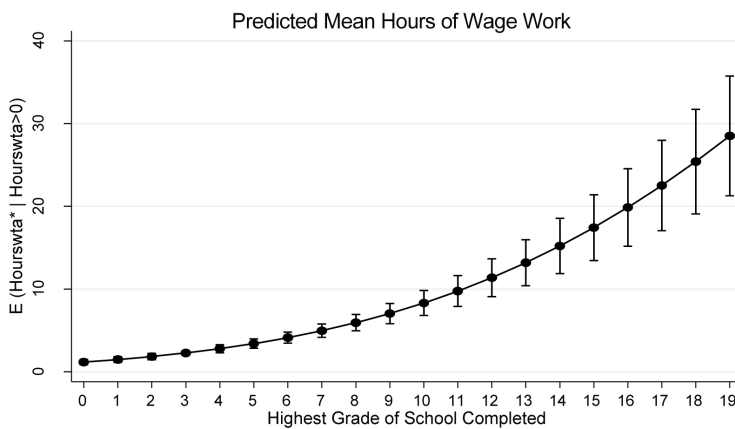


Fig. 12. Women, wage work, Tobit, $y^* > 0$
Note: Ninety-five (95) percent confidence intervals are included in the Figure.

Further insight into how the nonparametric and Tobit models perform can be gained by comparing predictions based on these models at various specified schooling levels. Predictions such as these may be useful in predicting work hours devoted to various work activities as average schooling of the labor force changes in response to deliberate policy changes that increase access to schooling or driven by higher incomes and changing work opportunities over the normal course of economic development. Work time allocation also has implications for individuals' economic well-being in that household agricultural work generally produces low incomes that vary significantly with weather. Non-agricultural household business work is likely to be less affected by weather variability relative to household agricultural work. Wage work is likely to have the highest and least variable economic returns.

Tables 4 and 5 report predictions of hours of work by activity at four levels of schooling based on the nonparametric and Tobit models, respectively. The levels of schooling are zero, four, six, twelve years ranging from no schooling to secondary schooling. Table 6 reports test statistics for testing the equality of predictions based on the Tobit model and the point estimates of the predictions based on the nonparametric model for each gender and type of work activity. Predictions differ significantly at the five percent significance level in 13 out of the 24 cases examined and in one additional case at the ten percent significance level. Focusing only on non-agricultural household business work and wage work, activities that are expected to expand as the economy develops further,

the predictions of the nonparametric and Tobit models differ significantly in 9 out of 16 cases at the five percent significance level and in one additional case at the ten percent level. Thus the two models result in significantly different predictions in the majority of the cases considered.

An example illustrating the differences in predictions is hours of wage work for men and for women with six years of schooling. The nonparametric model predicts 7.969 hours for men and the Tobit model 6.760 hours. For women the nonparametric model predicts 5.623 hours for women and the Tobit model 4.122 hours. For men the nonparametric model predicts 17.9 percent more hours and for women 36.4 percent more hours than the Tobit model.

Table 4. Average predicted hours of work evaluated at specified levels of schooling, Nonparametric models.

Schooling years	Gender	Household agricultural work	Household non-agricultural business work	Wage work
0	Men	16.375 (0.540)	2.345 (0.278)	3.172 (0.314)
0	Women	8.395 (0.280)	2.851 (0.231)	1.374 (0.164)
4	Men	12.216 (0.348)	4.551 (0.342)	6.120 (0.437)
4	Women	5.700 (0.665)	5.616 (0.723)	2.750 (0.490)
6	Men	9.977 (0.356)	5.922 (0.467)	7.969 (0.564)
6	Women	4.260 (0.425)	7.411 (0.862)	5.623 (0.869)
12	Men	3.477 (0.273)	8.205 (0.774)	14.794 (0.768)
12	Women	0.893 (0.236)	4.446 (0.746)	12.973 (1.442)

Note: Mean of predictions for each observation with schooling at the specified value. Bootstrapped standard error estimates are in parentheses (based on 400 replications).

Table 5. Average predicted hours of work evaluated at specified levels of schooling conditional on positive hours, Tobit models

Schooling years	Gender	Household agricultural work	Household non-agricultural business work	Wage work
0	Men	17.503 (0.929)	3.446 (0.358)	2.810 (0.397)
0	Women	8.768 (0.561)	3.436 (0.296)	1.180 (0.161)
4	Men	10.679 (0.611)	4.247 (0.343)	5.129 (0.469)
4	Women	4.580 (0.373)	4.096 (0.308)	2.791 (0.249)
6	Men	8.028 (0.556)	4.701 (0.374)	6.760 (0.512)
6	Women	3.159 (0.334)	4.463 (0.364)	4.122 (0.339)
12	Men	2.874 (0.421)	6.301 (0.688)	14.084 (0.987)
12	Women	0.845 (0.193)	5.726 (0.718)	11.377 (1.167)

Note: Mean of predictions for each observation with schooling at the specified value. The predictions are $E[y^*|y>0]$. Standard error estimates adjusted for clustering by enumeration area are in parentheses (326 clusters).

Table 6. Wald tests of equality of Tobit average predicted hours evaluated at specified levels of schooling and point estimates of average predicted hours based on Nonparametric models

Schooling years	Gender	Household agricultural work	Household non-agricultural business work	Wage work
0	Men	1.47 Pr.=0.225	9.48 Pr.=0.002	0.83 Pr.=0.361
0	Women	0.44 Pr.=0.506	3.90 Pr.=0.048	1.46 Pr.=0.227
4	Men	6.33 Pr.=0.012	0.79 Pr.=0.375	4.46 Pr.=0.035
4	Women	9.01 Pr.=0.003	24.30 Pr.=0.000	0.03 Pr.=0.869
6	Men	12.28 Pr.=0.001	10.68 Pr.=0.001	5.59 Pr.=0.018
6	Women	10.87 Pr.=0.001	65.60 Pr.=0.000	19.63 Pr.=0.000
12	Men	2.06 Pr.=0.152	7.65 Pr.=0.006	0.52 Pr.=0.470
12	Women	0.06 Pr.=0.805	3.18 Pr.=0.075	1.87 Pr.=0.172

Note: Standard error estimates adjusted for clustering by enumeration area are in parentheses (326 clusters).

Predictions of hours of household agricultural work based on nonparametric models reveal a flat relationship between hours of work and age¹⁶ up to the mid-50s and declining hours at older ages for both genders. The hours profile is uniformly lower for women relative to that for men. The reduction in hours of agricultural work for older people may be reflective of the physical demands of the work. Life expectancy in Ethiopia in 2015 was 63 years for men and 67 for women.¹⁷ It is likely that there is a relatively higher frequency of health problems among older members in the sample that affects their ability to perform physically demanding work. A similar pattern is predicted for wage work for men and nonagricultural household business work for women with flat age profiles in all other cases. The corresponding Tobit model predicts monotonically declining hours with age for all three work activities and both genders.

In addition to the Tobit models in Table 3, I also estimated Tobit models with quadratic terms in schooling and age.¹⁸ Estimates of the marginal effects of these models are in Appendix Table 1. The estimated marginal effects (Appendix Table 1) of schooling are relatively closer to the marginal effects derived from the nonparametric models (Table 2) than those derived from the Tobit models without quadratic terms of the regressors (Table 3). Age has positive and significant marginal effects for non-agricultural household business work for both genders and for wage work for men. For work in non-agricultural household businesses an inverted U pattern of predicted hours of work as a function of schooling is estimated.¹⁹ However, the maximum of hours for men are predicted at a lower level of schooling relative to the nonparametric model (at 9 years). The maximum hours for women are predicted at 8 years of schooling, the same as for the nonparametric model. The predictions based on the Tobit models with quadratic terms²⁰ are closer to those based on the nonparametric models (in Table 4) but still differ significantly from them in 5 out of the 24 cases considered.

Predictions of hours of each of the three work activities and for both genders based on Tobit models with quadratic schooling and age terms trace an inverted U pattern of hours at various ages. Maximum hours are estimated at

¹⁶ Figures not reported and available on request.

¹⁷ Source: World Health Organization (2022).

¹⁸ The regressors include schooling and its square and age and its square.

¹⁹ Figure not reported here and available on request. The Figures for predicted hours of other work activities by schooling based on Tobit models with quadratic terms are generally similar to the corresponding Figures based on Tobit models without quadratic terms.

²⁰ Not reported and available on request.

earlier ages (in the 40s) than implied by the nonparametric models. Predictions, of course, are only one aspect of the performance of the various models. Additional insight into the appropriateness of the Tobit models may be gained by performing formal specification tests.

Specification Tests of the Tobit Models

In this section I present the results of formal specification tests of the Tobit models. Estimates of Tobit models are consistent only under certain conditions. The validity of these conditions is testable. With that objective, Table 7 reports Lagrange Multiplier tests of the joint hypothesis that the Tobit models whose estimates are reported in Table 3 have normally distributed error terms, that they are linear in the regressors and that their error terms are homoscedastic (Vincent, 2010). Critical values for the tests are obtained by bootstrapping (Drukker, 2002). In all cases the (joint) null hypothesis is rejected and thus the Tobit models are misspecified. The Lagrange Multiplier tests indicate that the Tobit models are not appropriate characterizations of the data.

It is possible to investigate in more detail one aspect of the specification of the Tobit models by carrying out conditional moment tests of the normality assumption that is fundamental to the Tobit model. I report test statistics of conditional moment tests of normality (Skeels and Vella, 1999) of the Tobit models in Table 8.

Table 7. Lagrange multiplier specification tests of the Tobit models

	Men	Women
Hours household agricultural work	692.64 5% cv=3.408	386.36 5% cv=3.814
Hours household non-agricultural business work	289.15 5% cv=3.804	235.23 5% cv=3.693
Hours wage work	584.46 5% cv=4.182	345.82 5% cv=3.962

Table 8. Conditional moment normality tests for the Tobit models

	Men	Women
Hours household agricultural work	12.177 Pr.=0.002	3.508 Pr.=0.173
Hours household non-agricultural business work	541.71 Pr.=0.000	235.17 Pr.=0.000
Hours wage work	30.527 Pr.=0.000	15.569 Pr.=0.000

According to the conditional moment tests, normality is rejected in five of the six cases: household agricultural work, men, non-agricultural household business work, both genders and wage work, also both genders. Normality is not rejected in one case: household agricultural work, women. In short, normality is rejected for models of the hours of work in sectors of the economy that are expected to expand as the economy develops as well as the sector where men in rural areas spend the majority of their work time. Taken together the Lagrange Multiplier and Conditional Moment and tests indicate that the less restrictive nonparametric models are preferable to the Tobit models.

According to Lagrange Multiplier (LM) Specification tests in Appendix Table 2²¹, the Tobit specification with quadratic terms in schooling and age is rejected for all three work activities and both genders. Based on the LM tests, modeling the effect of schooling on hours of work by means of the less restrictive nonparametric models is preferable.

DISCUSSION

Nonparametric regression estimates reveal patterns in the relation of schooling and time allocation to three different work activities in Ethiopia that are not apparent with Tobit models that are more conventionally used to model labor supply. I compared the nonparametric regression estimates to Tobit estimates. The nonparametric and Tobit models result in significantly different predictions in the majority of the cases considered. Formal statistical

²¹ Conditional moment normality tests (Appendix Table 3) closely parallel the tests reported in Table 8; Normality is rejected for non-agricultural work activities and for both genders and for agricultural work for men.

tests indicate that the Tobit models are misspecified. Thus, the nonparametric models are a superior way of modeling the relationship of hours by type of work and worker characteristics.

CONCLUSIONS

I estimated nonparametric regressions of hours of work in three categories of work activities on individual worker characteristics with data from Ethiopia. The regressions are useful in investigating how policies that result in changes in individuals' schooling are likely to affect their type of work activity which has implications for poverty reduction and their economic well-being. The nonparametric models are the preferred way of modeling the relationship of hours by type of work and worker characteristics.

Acknowledgements

I thank Charles Link and James Mulligan for their comments.

REFERENCES

- Blomquist, S., and Newey, W. (2002). Nonparametric estimation with nonlinear budget sets. *Econometrica*, 70(6), 2455-2480. <https://doi.org/10.1111/j.1468-0262.2002.00445.x>
- Cameron, A. C., and Trivedi, P. K. (2005). *Microeconometrics: Methods and applications*. New York, NY: Cambridge University Press.
- Drukker, D. M. (2002). Bootstrapping a conditional moments test for normality after Tobit estimation. *The Stata Journal*, 2(2), 125-139. <https://doi.org/10.1177/1536867X0200200202>
- Fields, G. S. (2011). Labor market analysis for developing countries. *Labour Economics*, 18, S16-S22. <https://doi.org/10.1016/j.labeco.2011.09.005>
- Glick, P., and Sahn, D. E. (1997). Gender and education impacts on employment and earnings in West Africa: evidence from Guinea. *Economic Development and Cultural Change*, 45(4), 793-823. <https://www.jstor.org/stable/10.1086/edcc.1997.45.issue-4>
- Henderson, D. J., and Souto, A.-C. (2018). An introduction to nonparametric regression for labor economists. *Journal of Labor Research*, 39(4), 355-382. <https://doi.org/10.1007/s12122-018-9279-6>
- Kimmel, J., and Connelly, R. (2007). Mothers' time choices: caregiving, leisure, home production, and paid work. *Journal of Human Resources*, 42(3), 643-681. <https://www.jstor.org/stable/40057322>
- Li, Q., and Racine, J. S. (2007). *Nonparametric econometrics: theory and practice*. Princeton, NJ: Princeton University Press.
- Liang, C.-Y. (2012). Nonparametric structural estimation of labor supply in the presence of censoring. *Journal of Public Economics*, 96(1-2), 89-103. <https://doi.org/10.1016/j.jpubeco.2011.08.009>
- Malathy, R. (1994). Education and women's time allocation to nonmarket work in an urban setting in India. *Economic Development and Cultural Change*, 42(4), 743-760. <https://www.jstor.org/stable/1154294>
- Skeels, C. L., and Vella, F. (1999). A Monte Carlo investigation of the sampling behavior of conditional moment tests in Tobit and Probit models. *Journal of Econometrics*, 92(2), 275-294. [https://doi.org/10.1016/S0304-4076\(98\)00092-X](https://doi.org/10.1016/S0304-4076(98)00092-X)
- Thakral, N., and Tô, L. T. (2021). Daily labor supply and adaptive reference points. *American Economic Review*, 111(8), 2417-2443. <https://doi.org/10.1257/aer.20170768>
- Vincent, D. (2010). *Bootstrap LM tests for the Box Cox Tobit model*. In BOS10 Stata Conference (Vol. 9), Stata Users Group. Retrieved January 28, 2022 from www.stata.com/meeting/boston10/boston10_vincent.pdf
- World Health Organization (2022). *Ethiopia*. Retrieved January 28, 2022 from <https://www.afro.who.int/countries/ethiopia>

Appendix 1. Estimates of average marginal effects of slope regressors and mean predicted hours conditional on positive hours, Tobit hours of work models with quadratic schooling and age terms

Hours household agricultural work	Men	Women
Schooling	-0.894 (0.153)	-0.620 (0.191)
Age	-0.012 (0.027)	-0.030 (0.020)
Mean predicted hours	10.754 (0.567)	6.538 (0.421)
Hours household non-agricultural business work		
Schooling	0.350 (0.047)	0.502 (0.064)
Age	0.058 (0.024)	0.045 (0.022)
Mean predicted hours	4.633 (0.371)	3.974 (0.290)
Hours wage work		
Schooling	0.903 (0.081)	0.595 (0.064)
Age	0.064 (0.031)	0.033 (0.022)
Mean predicted hours	7.338 (0.490)	3.411 (0.271)

Notes. Standard error estimates adjusted for clustering by enumeration area are in parentheses (326 clusters). The predictions are $E[y^*|y>0]$.

Appendix 2. Lagrange multiplier specification tests of the Tobit models with quadratic terms in schooling and age

	Men	Women
Hours household agricultural work	621.32 5% cv=4.427	376.51 5% cv=4.100
Hours household non-agricultural business work	266.48 5% cv=4.138	245.17 5% cv=4.285
Hours wage work	584.79 5% cv=3.794	353.23 5% cv=3.547

Appendix 3. Conditional moment normality tests for the Tobit models with quadratic schooling and age terms

	Men	Women
Hours household agricultural work	7.834 Pr.=0.020	4.351 Pr.=0.114
Hours household non-agricultural business work	236.69 Pr.=0.000	90.279 Pr.=0.000
Hours wage work	27.827 Pr.=0.000	13.714 Pr.=0.000