Feb. 26 2015

The Effect of Height on Earnings: Is Stature a Proxy for Cognitive and Non-Cognitive Skills?

Laurent Bossavie

Harold Alderman

John Giles

Cem Mete

Abstract: This study investigates the degree to which association of height on earnings in Pakistan is independent of other cognitive and socio-emotional skills. While taller workers are regularly observed to earn more it is also commonly observed that taller individuals have more schooling and higher measures of cognitive ability. Thus, there is some debate concerning the independent contribution of stature on earnings. We find that there is only modest attenuation of the coefficient of height when a measure of cognitive ability – performance on Raven's matrices – is included. Additionally we include an index of socio-emotional capacity based on principal components of the big five indicators and, again, find that the association of height and earnings remains similar to its magnitude when nutritional status is included alone.

The Effect of Height on Earnings: Is Stature a Proxy for Cognitive and Non-cognitive Skills?

Introduction

Earnings and wages are regularly found to be associated with height in both developed economies (Case and Paxson, 2008) and low income settings (LaFave and Thomas, 2013; Schultz, 2003; Thomas and Strauss, 1997; Haddad and Bouis, 1991). In some occupations this may be due to a direct impact of height on physical capacity for work; in others it may reflect the indirect effect of height on schooling or on status or a combination of these (Pitt, Rosenzweig, and Hassan, 2013). Plausibly, however, height may have relatively little direct impact on earnings but may be a proxy for other dimensions of human capital that are less often measured and – for employers – less easily observed at the time of hiring.

Differences in the measured impact of nutrition on earnings or wages may reflect - as is often the case - context. However, differences across studies may also reflect whether the results are net of schooling or learning (Behrman et al. 2103), whether health has been assumed to be exogenous or not (Alderman et al. 1996; Schultz 2003) and whether cognitive skills are included in the analysis (Vogl 2012; LaFave and Thomas 2013). The current study looks at these issues as well as includes a further measure of socio-emotional skills. These latter skills – which are also referred to as non-cognitive skills, particularly in economics literature – have been shown to be an important determinant of labor market outcomes in the United States (Heckman, Stixrud and Urzua 2006) but have only recently been included in studies of earnings in a wider context.

Our results are generally consistent with from the findings of Vogl (2014) and Lafave and Thomas (2014) as well as a similar paper by Bargain and Zeidan (2014) in that height remains a determinant of earnings even when cognitive and non-cognitive measures are included. The current study, however, adds to the small pool from which generalizations can be made. Moreover, unlike the previous studies, the current investigation takes the endogeneity of height into consideration.

Basic Conceptual Framework

In order to view human capital over a lifetime, we consider three periods. In the first period, the foundations for an individual's health and nutrition (H_{i1}) are established as a function of investments (I_{i1}) in that period as well as the individual's own genetic makeup (X_i), his or her family's characteristics (F_1) and community infrastructure (V_1). These latter two categories can be time varying.

1) $H_{i1} = h(I_{i1}, X_i, F_1, V_1)$.

In the following period the child accumulates other forms of human capital (S_{i2}), which can be considered as schooling or learning (Hanushek and Woessmann, 2008) and which reflect health accumulated earlier as well as current inputs and individual, family and community characteristics.

2) $S_{i2} = s(h(H_{i1}), I_{i2}, X_i, F_2, V_2)$

When the individual enters employment his or her wages or earnings reflect both health and learning along with other individual characteristics as well as local market conditions.

3) $W_{i3} = w(h(H_{i1}), s(Si2), X_i, V_3)$.

More detailed models can illustrate how inputs in one period influence the returns to inputs in subsequent periods or can fine tune different period of sensitive investments (Cunha and Heckman, 2007). In addition, the number and types of investment in each period included in forms of this model of inter-period accumulation of human capital depends on the nature of the analysis. However, the model is general and a parsimonious illustration suffices for the study at hand.

Data.

This paper uses data from the second wave of the Labor Skill Survey (LSS) conducted in Pakistan in the last quarter of 2013. The survey was designed to be nationally representative and covers all regions of Pakistan except Balochistan and the Federal Administered Tribal Areas, which jointly represent less than 7% of the total population. The sample was drawn using a stratified three-stage design. Twenty districts (7 in urban and 13 in rural areas) were first selected through a systematic random sampling in each of the urban and rural strata. In the second stage, 100 primary sample units at the union council level were selected within each stratum systematically with probability proportional to size. Finally, in the third stage, a random systematic sampling was used to select 25 target households. A total of 2,500 households in 20 districts and 100 union councils were finally selected. Interviews were completed for 2,354 households in 94 union councils, due to security issues in six union councils of Khyber Pakhtunkhwa (KPK).

The LSS household survey consists of a questionnaire for the household head as well as a separate questionnaire for a subset of one male and one female randomly selected within the household, among all mentally able household members aged between 15 and 64. The household head questionnaire collected general information on all household members including age, gender, height and general education. The male and female questionnaires collect detailed information on individual employment, income, and individual skills reported by the individual himself. Additionally, cognitive abilities were assessed for all males and females aged 15 to 64 in the household. The height variable used in the paper was obtained from actual measurement of all household members aged 5 or more.

Our measure of cognitive ability is from Raven's Progressive Matrices™ [Pearson Publishing Company, NJ, USA]. Raven's Progressive Matrices measure logical reasoning ability and consist of 60 questions of increasing difficulty in which the individual has to find the missing figure in a logical sequence. Raven's Progressive Matrices have been extensively used in the literature to measure cognitive abilities and logical reasoning. We use the final raw score obtained from the Raven's Progressive Matrices as our proxy for cognitive abilities.

Our measure of non-cognitive abilities is based on the Big 5 Personality Test that was included in the male and female questionnaires. The Big 5 personality test consists of a set of 24 questions aimed at measuring 5 different dimensions of non-cognitive abilities, typically referred as "the Big 5" in the literature. Each question is a statement about a given behavior of the respondent in his daily life to which it can be answer to be always true, true most of the time, rarely true or never true. The Big 5 personality traits measured are openness, conscientiousness, extraversion, agreeableness, and neuroticism. The methodology underlying the Big Five taxonomy is the Five Factor Model (FFM) and has been widely used in the psychology and economics literature (Heckman and Kautz, 2012; Heckman and Mosso, 2014).

The Big Five indicators have been proven to show consistency in interviews, self-descriptions and observations (Costa and McCrae, 1987). We construct our indicator of non-cognitive abilities from the score obtained in the 5 components of the Big 5 personality traits, using a principal component analysis. The first principal component of the Big 5 indicators obtained from this procedure is used as our proxy for non-cognitive skills.

The Pakistan LSS surveys a total of 14,254 individuals in 2,354 households. Since income data is non-zero only for individuals who work and who receive payment or for those respondents who report earning a profit, our final sample consists of 1,364 working male individuals that are either self-employed, or paid employees. This paper restricts the analysis to a sample to male workers, given the very small number of females that report receiving earning. We classify individuals who report earning money from their own activity as self-employed, and individuals who receive payment from their employer as wage earners. Self-employed workers were asked the amount of net profits generated over the last period of work, explicitly defined as sales minus expenses to the individual. We use this amount converted into monthly profits as our measure of earnings for the self-employed. Table 1 reports summary statistics on this sample while Table 2 indicates the correlation between the main variables of interest for this analysis, height, skills and schooling. There is a legitimate concern that land owners could conflate rental earnings with crop production minus expenses. Table 3a and 3b presents some indication that this is not an obvious concern with the data. Land owners - roughly 20% of the entire category of self-employed and approximately half of all individuals who reported cultivation as their main employment - reported agricultural profits that were not significantly different than those reported by renters. Furthermore, there is no correlation between reported profits and the amount of land owned as would be the case for rental earnings.

Results.

Table 4 presents OLS regressions of the logarithm of earnings from both wage and own employment. The first three columns focus on the impacts of stature and cognitive ability entered separately. An additional centimeter of height contributes nearly 1% to earnings. Similarly, an additional 1 point of the Raven's (roughly 1 standard deviation) is associated with a comparable 1% increase in earnings. The regression in the fourth column includes all three of these measures jointly. As indicated, there is an attenuation of the magnitude of height and cognition compared to that observed in the first three columns but they remain individually significant. This is consistent with the relatively low correlations reported in Table 2.

Moreover, as reported in column 5 when we add the education of the father, that variable proves significant. As the labor market does not directly reward the ability of the father – many of whom are deceased - this may indicate that there are aspects of the ability of the current generation of workers that are not directly measured by stature or the skills included here yet are recognized in the labor market. Possibly the father's education is a proxy for unmeasured skills that are genetically transmitted. Alternatively, or additionally to this interpretation, the coefficient of father's education may indicate learning that is imparted by parental guidance. Furthermore, the coefficient can also reflect access to networks that an educated father can facilitate.

The regression in column 5 may be considered as the full reduced form impact of these categories of skills. That is, the coefficients capture the indirect impact of these aspects of human capital on wages via schooling and also the impact of stature and skills on labor market choices regarding sector and labor supply as well as measure any direct impact on earnings conditional on these choices. Cawley, Heckman, and Vytlacil (2001) show that the estimated impact of ability on wages is substantially smaller when it is conditional on levels of schooling. This is also observed in our data. Column 6 reports the impact of stature and skills conditional on schooling. The coefficient on height remains significant but the magnitude declines relative to column 5. That it has additional impact conditional on school either might be because it influences learning per year of school or because it conveys abilities that are rewarded in earnings beyond the returns to schooling *per se* or both.

In contrast, both the measure of cognitive skills as well as socio-emotional ability are no longer significant in the regression in column 6. The standard errors for the coefficient of Raven's actually declines relative to the previous estimate, so the loss of significance is driven by the reduction of the point estimate. The impact implied by the coefficients of Raven's and socio-emotional skills in column 5 likely works primarily through the indirect impact of skills on schooling. That is, the coefficients in column 5 can be viewed largely as $\delta W/\delta S * \delta S/\delta H$. At the same time it appears that schooling is not merely a signal for these skills since the inclusion of schooling increases the portion of earnings explained in the regression.

However, as mentioned, wage earnings and earnings from own employment are pooled in the first 6 columns. Columns 7 and 8 indicate how these skills influence earnings in these two sectors respectively. Height is far more important in wage employment than in own employment. This is in partial contrast to Pitt, Rosenzweig, and Hassan (2013) who argue that employment in agriculture may reflect the relative importance of physical capacity in that sector. As 41.1 percent of the individuals in our sample who are self-employed are in agriculture the results in column 7 should largely reflect the role of stature in agriculture. Our measure of socio-emotional skills is not important in wage employment although it is in self-employment. Conditional on schooling, Raven's scores are not significant in either sector, although the coefficient of father's education is significant in both.

The results in table 4, however, tacitly assume that eq. 1 and 2 can not only be considered lagged endogenous but also that they have no common unobservables with equation 3 and, thus, no correlation of errors with estimations based on equation 3. Table 5 reports the same regressions as in table 3 with stature instrumented.[1] The land holding of the father of the current employed individuals (as this individual recollects it) and its square are used as instruments. The first stage regression is reported in the annex. As the survey obtained height information from all adults but earnings for only a subset, the full sample was used for instrumenting height and standard errors for the second stage IV were obtained by bootstrapping rather than running the IV estimates as a simultaneous set of equations. This approach was also motivated by the fact that the first step of the IV corresponds to equation 1 while the wage equation is an estimate for equation 3. Time variant information that is observed at the time of employment does not pertain to the production of skills in equation 1 and thus is not appropriate in the estimation of height. The F statistic for the first stage regression exceeds the rule of thumb for plausible instruments from Stock and Yogo

[1] As columns 2 and 3 in Table 4 are not affected by the IV approach used here these are omitted in Table 4.

(2005). The Hansen's test of over-identification indicated that father's education was not uncorrelated with errors in the equations reported in table 3, hence it was included in the wage equation and not considered an identifying instrument here.

While the magnitude of the contribution of height to explaining wages in table 4 using OLS regressions is somewhat lower than reported elsewhere in the literature (such as in Vogl 2014) the magnitude of the coefficient of instrumented height in table 5 increases substantially.[2] This may reflect a combination of errors in measurement as well as endogenous choices. In contrast to the OLS results the IV coefficient is significant for self-employed (including those engaged in agriculture) while it is not for wage workers. The magnitude Raven's score is not significant conditional on schooling when height is instrumented but it is a significant explanatory variable for wage earnings in the relatively reduced form excluding education.

In principle, schooling is also an endogenous choice and thus it would be desirable to also instrument the education variables in table 5. Duflo (2001), however, finds little difference between instrumented estimates of the impact of years of schooling on wages and OLS results, a result that reinforces an earlier review by Ashenfelter et al. (1999). Similarly, Chou et al. (2010), using an approach similar to that of Duflo, observe little bias in the OLS coefficients on the effect of education on birth weights and infant mortality. Thus, given the limited information available for investments in human capital made a generation ago, the next step in this analysis concentrates on including selection into wage employment in the determinants of earnings.

The initial selection into wage employment is reported in Table 6 while Table 7 reports the results for the regressions explaining earnings conditional on selection into wage employment. The probit equation includes land holding and education of the father of the worker as well as a dummy variable for whether the worker's father is still living and an interaction of that variable

[2] This was also observed by Schultz (2003) in the case of Ghana.

with land holding. The probability of wage employment increases with education only when the individual has a secondary or higher level of education. Conditional on education, wage employment decreases with land holding; conversely, the probability of being self-employed including working in agriculture increases with father's landholding as is logical. The Raven's score is significant in the choice of wage employment even when years of completed schooling are included but neither height nor the index of socio-emotional skills influence selection into wage employment. The absence of a role of height in sectoral choice is in contrast to the results in both Vogl (2014) and LaFave and Thomas (2013). However, the significance of Raven's is consistent with other evidence in the literature including the cited studies.

The OLS results reported in table 7 do not differ appreciably from the corresponding tables which do not account for selection into wage employment. The IV results, however, reflect the challenge of accounting for both selection as well as endogeneity with the limited instruments for lagged decisions on health. Statistical power is lost, for example, on the coefficient of height in self-employment although the point estimate is similar to that in table 5. On the other hand, the final column in table 7 indicates a significant impact of cognitive ability on wages that was not observed in the corresponding regressions in either table 4 or 5.

Conclusion

The results reported in this paper support the view that height provides independent information on labor productivity rather than only serving as a proxy for other measures of human capital. This general point is consistent with other studies that include one or more measure of skills in addition to height. The IV results reported here show a similar impact of height that is similar in magnitude to those in Schultz (2003) as well as the OLS results conditional on the inclusion of cognitive ability in Vogl (2014). Similarly, the portion of cognitive ability that is captured by the Raven's score confirms a pattern that is often observed. The paper also shows that socio-emotional skills have additional explanatory power, a finding that is more commonly noted in studies from high income countries.

However, drilling down a bit into the manner by which height explains labor choices indicates that there also are some differences with previous studies. For example, height seems to have no role in selection into wage work in Pakistan and, of course, given the bivariate choice in the choice of occupations this also implies it has no role in selection into other activities including those in which brawn is assumed to be more central.

The differences with other results in the literature likely reflect context – there are too few studies that include both height and a range of other measures of human capital from which to generalize. In addition, the data required for precise estimation of selection and endogenous choice over the life-cycle are daunting. Still, as the OLS results hint, the three relatively accessible measures of skills studied here confirm not only that ability is multi-dimensional but that insights into these dimensions available with judicious modifications of standard labor force surveys.

References:

Alderman, Harold, Jere Behrman, David Ross, and Richard Sabot. 1996. The Returns to Endogenous Human Capital in Pakistan's Rural Wage Labor Market. Oxford Bulletin of Economics and Statistics. 58(1): 29-55.

Ashenfelter, Orley; Harmon, Colm P. and Oosterbeek, Hessel. A Review of Estimates of the Schooling/Earnings Relationship. Labor Economics, November 1999, *6*(4), pp. 453–70.

Bargain, Oliver and Jinan Zeidan. 2014. Stature, Skills and Adult Life Outcomes: Evidence from Indonesia. Aix Marseilles University. Working paper.

Behrman, Jere, John Hoddinott, John Maluccio, Reynaldo Martorell. 2013. Brains versus Brawn: Labor Market Returns to Intellectual and Health Human Capital in a Poor Developing Country. Middlebury College. Dept. of Economics. Processed.

Case, Anne and Christina Paxson. 2008. Stature and Status: Height, Ability and Labor Market Outcomes. Journal of Political Economy. 116(3): 499-532.

Cawley, John, James Heckman and Edward Vytlacil. 2001. Three Observations on Wages and Measured Cognitive Ability. Labor Economics 8:419-442.

Chou S-Y, Liu J-T, Grossman M, Joyce T. Parental education and child health: evidence from a natural experiment in Taiwan. American Economic Journal: Applied Economics 2010; 2: 33–61.

Costa, Paul and Robert McCrae. 1987. Validation of the five-factor model of personality across instruments and observers. J Pers Soc Psychol. 52(1):81-90.Cunha, Flavio and James Heckman. 2007. The Technology of Skills Formation. American Economic Review**.** 2007 97(2): 31-47.

Duflo, Esther. 2001. Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. American Economic Review 91(4): 795-813.

Haddad, Lawrence and Howarth Bouis. 1991. The impact of nutritional status on agricultural productivity: wage evidence from the Philippines. Oxford bulletin of economics and statistics, 53(1), 45-68.

Hanushek, Eric, and Ludger Woessmann. 2008. The Role of Cognitive Skills in Economic Development. Journal of Economic Literature 46 (3): 607–68.

Heckman, James, Jora Stixrud and Sergio Urzua. 2006. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. Journal of Labor Economics 24(3): 411-82.

Heckman, James and Stefano Mosso. 2014. The Economics of Human Development and Social Mobility. Annual Review of Economics, Annual Reviews 6(1): 689-733, 08.

Heckman, James and Tim Kautz. 2012. Hard evidence on soft skills. Labor Economics, 19(4): 451-464.

LaFave, Daniel and Duncan Thomas. 2013. Height and Cognition at Work. Labor Market Performance in a Low Income Setting. Colby College. Processed.

Pitt, Mark, Mark Rosenzweig, and M.N. Hassan. 2013. Human Capital Investment and the Gender Division of Labor in a Brawn-Based Economy. American Economic Review, 102(7): 3531–60.

Schultz, Paul. 2003. Wage Rentals for Reproductive Human Capital: Evidence from Ghana and the Ivory Coast. Economics and Human Biology. 1: 331-336.

Stock, James and Motohiro Yogo. 2005. Testing for Weak Instruments in IV Regression, in Identification and Inference for Econometric Models: A Festschrift in Honor of Thomas Rothenberg. Donald W. K. Andrews and James H. Stock, eds. Cambridge University Press, pp.80-108.

Thomas, Duncan and John Strauss, 1997: Health and Wages: Evidence on Men and Women in Urban Brazil, Journal of Econometrics, 77(1): 159–187.

Vogl, Tom. 2014. Height, Skills, and Labor Market Outcomes in Mexico. Journal of Development Economics. 107(1): 84-96.

Table 1: Summary Statistics of Male Workers in the Sample

Variable	N	Mean	Std. Dev.	Min	Max
Height in meters	1364	1.665256	0.078722	1.27	2
Monthly earnings (all workers)	1364	14550.26	26906.59	700	600000
Monthly profit (self-employed only)	1034	14295.58	27741.38	0	600000
Monthly wage (employees only)	330	14866.27	23735.49	1000	300000
Dummy for Sindh	1364	0.237276	0.425565	0	1
Dummy for Punjab	1364	0.650896	0.476858	0	1
Dummy for KPK/AJK	1364	0.111828	0.315268	0	1
Age	1364	36.76057	11.43524	15	64
Potential experience	1364	26.29606	12.66625	0	58
Dummy for urban	1364	0.351971	0.477756	0	1
No schooling	1364	0.464516	0.498918	0	1
Primary school	1364	0.170609	0.376302	0	1
Middle school	1364	0.13405	0.340828	0	1
High school and higher	1364	0.230824	0.421511	0	1
Years of schooling	1364	4.670259	5.06476	0	21
Raven score/100	1364	0.218746	0.096328	0.03	0.56
Non-cognitive index	1364	0.007728	1.014704	2.93434	3.039801

Table 2: Correlation between Height, Skills and Years of Schooling

	Height	Ravens score	Years of schooling	Non-cognitive index
Height	1			
Ravens score/100	0.0599	1		
Years of education	0.1125	0.3303	1	
Non-cognitive index	0.036	0.1427	0.1531	1

Table 4: OLS Regressions Results for Earnings

	(1) All Workers	(2) All Workers	(3) All Workers	(4) All Workers	(5) All Workers	(6) All Workers	(7) Self employed	(8) Wage earners
Urban	0.444***	0.437***	0.448***	0.437***	0.381***	0.331***	0.361***	0.242***
	(0.0689)	(0.0664)	(0.0690)	(0.0663)	(0.0590)	(0.0506)	(0.0590)	(0.0847)
Age/100	0.0992	0.220	0.160	0.166	0.315*	0.283	0.0183	0.990**
	(0.176)	(0.178)	(0.175)	(0.182)	(0.181)	(0.173)	(0.199)	(0.385)
Height	0.898***			0.797***	0.689**	0.536**	0.364	1.025**
	(0.284)			(0.278)	(0.272)	(0.256)	(0.314)	(0.512)
Ravens' score/100		1.055***		0.947***	0.592**	0.230	0.222	0.291
		(0.311)		(0.304)	(0.284)	(0.261)	(0.308)	(0.429)
Non-cognitive index			0.0597***	0.0441**	0.0355*	0.0251	0.0574**	-0.0499
			(0.0220)	(0.0214)	(0.0209)	(0.0208)	(0.0261)	(0.0333)
Father's education					0.0394***	0.0273***	0.0296***	0.0219**
					(0.00635)	(0.00638)	(0.00803)	(0.00892)
Primary schooling						0.0555	0.0498	0.109
						(0.0515)	(0.0600)	(0.111)
Middle schooling						0.154**	0.196**	0.0570
						(0.0741)	(0.0859)	(0.0986)
High school or higher						0.411***	0.294***	0.651***
						(0.0681)	(0.0844)	(0.0983)
Constant	7.489***	8.712***	8.961***	7.428***	7.583***	7.834***	8.232***	6.719***
	(0.458)	(0.106)	(0.0659)	(0.457)	(0.447)	(0.420)	(0.507)	(0.850)
N	1364	1364	1364	1364	1364	1364	1034	330
adj. R-sq	0.078	0.086	0.076	0.094	0.123	0.155	0.129	0.291

* p<0.1 **p<0.05 *** p<0.01

Standard errors clustered at the village level are reported in parentheses

Table 5: Instrumental Variable Regression Results for Earnings

	(1)	(2)	(3)	(4)	(5)	(6)
	All Workers	All Workers	All Workers	All Workers	Self employed	Wage earners
Urban	0.446***	0.440***	0.380***	0.325***	0.356***	0.227***
	(0.0696)	(0.0670)	(0.0593)	(0.0511)	(0.0599)	(0.0843)
Age/100	-0.163	-0.0785	0.0783	0.118	-0.126	1.001**
	(0.188)	(0.192)	(0.188)	(0.182)	(0.215)	(0.459)
Height	3.587***	3.333***	3.212**	2.788**	2.885*	0.952
	(1.240)	(1.267)	(1.255)	(1.198)	(1.571)	(1.943)
Ravens' score/100		0.957***	0.593**	0.228	0.197	0.322
		(0.306)	(0.284)	(0.266)	(0.316)	(0.438)
Non-cognitive index		0.0465**	0.0377*	0.0246	0.0584**	-0.0536
		(0.0217)	(0.0211)	(0.0211)	(0.0261)	(0.0355)
Father's education			0.0407***	0.0289***	0.0309***	0.0232**
			(0.00635)	(0.00639)	(0.00801)	(0.00903)
Primary schooling				0.0591	0.0614	0.0707
				(0.0516)	(0.0594)	(0.114)
Middle schooling				0.165**	0.208**	0.0467
				(0.0755)	(0.0880)	(0.102)
High school or higher				0.408***	0.296***	0.633***
				(0.0686)	(0.0841)	(0.0997)
Constant	3.116	3.301	3.473*	4.155**	4.096	6.861**
	(2.027)	(2.077)	(2.059)	(1.968)	(2.586)	(3.140)
N	1364	1364	1364	1364	1034	330
adj. R-sq	0.076	0.094	0.124	0.156	0.132	0.281

* p<0.1 **p<0.05 *** p<0.01

Standard errors clustered at the village level in parentheses

Table 6: Selection into Wage employment – First step of Heckman correction

	(1) Probit Coefficient	(2) Marginal effect
Father died	0.00923	0.0026
	(0.0938)	(0.0267)
Father's area of land	-0.0602***	-0.0171***
	(0.0185)	(0.0052)
Land*Father died	0.0239	0.00680
	(0.0216)	(0.1193)
Age/100	-1.332***	-0.3796***
	(0.414)	(0.1193)
Height	0.0665	0.0189
	(0.507)	(0.144)
Raven's score	1.226**	0.349**
	(0.512)	(0.149)
Non-cognitive index	0.0444	0.0126
	(0.0469)	(0.0134)
Primary schooling	0.182	0.0544
	(0.124)	(0.0379)
Middle schooling	0.144	0.0427
	(0.130)	(0.0399)
High school of higher	0.594***	0.1884***
	(0.117)	(0.0389)
Constant	-0.770	
	(0.833)	
N	1356	1356
Pseudo R-sq.	0.0717	

* p<0.1 ** p<0.05 ***p<0.01
Standard errors clustered at the village level are reported in parentheses.

Table 7: Wage Estimates by Sector using Heckman correction

	Ordinary Least Square				Instrumental Variable			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Profit	Wage	Profit	Wage	Profit	Wage	Profit	Wage
Urban	0.366***	0.268***	0.365***	0.226***	0.364***	0.265***	0.361***	0.220***
	(0.0613)	(0.0811)	(0.0590)	(0.0841)	(0.0615)	(0.0801)	(0.0586)	(0.0832)
Age/100	0.341	1.931***	-0.0946	0.264	0.146	1.957***	-0.208	0.158
	(0.237)	(0.479)	(0.231)	(0.455)	(0.241)	(0.570)	(0.248)	(0.511)
Height	0.477	1.238**	0.374	0.860*	3.285**	0.722	2.326	0.961
	(0.350)	(0.611)	(0.313)	(0.495)	(1.637)	(2.136)	(1.512)	(1.841)
Ravens' score/100	0.0669	0.189	0.350	0.880*	0.0799	0.171	0.334	0.882*
	(0.339)	(0.534)	(0.334)	(0.463)	(0.343)	(0.555)	(0.330)	(0.480)
Non-cognit. index	0.0511*	-0.0815*	0.0617**	-0.0367	0.0531*	-0.0780*	0.0623**	-0.0339
	(0.0280)	(0.0442)	(0.0261)	(0.0440)	(0.0279)	(0.0447)	(0.0257)	(0.0453)
Father's education	0.0350***	0.0269**	0.0285***	0.0241***	0.0350**	0.0279**	0.0294***	0.0256***
	(0.00744)	(0.0101)	(0.00790)	(0.00895)	(0.00747)	(0.0103)	(0.00797)	(0.00868)
Primary schooling			0.0760	0.166			0.0779	0.171
			(0.0603)	(0.145)			(0.0599)	(0.148)
Middle schooling			0.214**	0.106			0.218**	0.119
			(0.0892)	(0.134)			(0.0888)	(0.137)
High school or higher			0.366***	0.909***			0.359***	0.932***
			(0.0906)	(0.139)			(0.0921)	(0.141)
Coeff. of inverse Mills		-0.756		-0.68		-0.764		-0.726
		(0.095)		(0.092)		(0.909)		(0.133)
N	1034	330	1034	330	1034	330	1034	330

* p<0.1 **p<0.05 *** p<0.01

Standard errors clustered at the village level are reported in parentheses

Appendix: Instrumenting Equation for Endogenous Nutritional Status

The first column in Annex Table 1 indicates the regression used to instrument for height and to control for any measurement error. In addition to father's education and its square, the explanatory variables include an indicator of parental consanguinity all of which are statistically significant. Additionally, the model specification allows for the possibility that the younger individuals in the sample may not have finished growing.

The specification presented in the second column is not actually used in the estimates reported in Table 5 and 7 but it is included as it contains a result of interest. In a regression not reported we regressed the height of the young children of the adults currently in the labor survey using measures of household characteristics as well as the child's age and gender. The residual of that equation should reflect the child's genetic makeup in addition to the influence of unobserved community elements and conventional measurement error. While genes flow from parent to child, the association is two way and including the average residual of all measured descendants as a regressor in the equation for the height of the adult provides information not generally available and the coefficient is plausible. Unfortunately, as not all individuals in the labor sample had children young enough to be included in the measurement of height, this restricted the application of this information and the results in column 2 were not used in the estimations reported in table 5.

Annex Table 1: Instrumenting for Height – First Stage

	(1) Height	(2) Height
Parents are related	-0.00807*	-0.0102**
	(0.00418)	(0.00451)
Age/100	0.0276***	0.0407*
	(0.00973)	(0.0214)
Age is below 19	-0.0870***	
	(0.00613)	
Father's area of land	0.615***	0.484**
	(0.210)	(0.228)
Father's area of land squared	-0.971**	-0.721*
	(0.373)	(0.377)
Mean residual of children's height		0.00466***
		(0.00107)
Constant	1.658***	1.668***
	(0.00552)	(0.0110)
N	3715	1339
adj. R-sq	0.137	0.025
F-statistic	55.34	6.16

* p<0.1 ** p<0.05 *** p<0.01

Standard errors clustered at the village level are reported in parentheses