

# Height and Cognition at Work: Labor market productivity in a low income setting

Daniel LaFave  
Colby College

Duncan Thomas  
Duke University

May 2016

## Abstract

Taller workers earn more, particularly in lower income settings. It has been argued that adult height is a marker of strength which is rewarded in the labor market, a proxy for cognitive performance or other dimensions of human capital such as school quality, a proxy for health status or a proxy for family background characteristics. As a result, the argument goes, height is rewarded in the labor market because it is an informative signal of worker quality to an employer. It has also been argued that the height premium in the labor market is driven by occupational and sectoral choice. This paper evaluates the relative importance of these mechanisms that potentially underly the link between adult stature and labor market productivity. Drawing on twelve waves of longitudinal survey data collected in rural Central Java, Indonesia, we establish that height predicts hourly earnings after controlling education, multiple indicators of cognitive performance and physical health status, measures of family background, and sectoral and occupational choice. The height premium is large and significant in both the wage and self-employed sectors indicating height is not only a signal of worker quality. Since adult stature is largely determined in the first few years of life, we conclude that exposures during this critical period have an enduring impact on labor market productivity.

**Keywords:** Height, Cognition, Productivity, Labor Markets

**JEL Codes:** O15, I15, J24

This research has benefited from discussions with Ryan Brown, John Maluccio, Hani Mansour, John Strauss, T. Paul Schultz and Tom Vogl. Financial support from the National Institutes of Aging is gratefully acknowledged.

Daniel LaFave, Department of Economics, Colby College, Waterville, ME 04901. Email: [daniel.lafave@colby.edu](mailto:daniel.lafave@colby.edu)  
Duncan Thomas, Department of Economics, Duke University, Durham, NC 27708. Email: [dthomas@econ.duke.edu](mailto:dthomas@econ.duke.edu)

# 1. Human Capital and Labor Market Performance

There is abundant evidence that taller people live longer, are healthier, better educated and have higher standards of living (Deaton and Arora, 2009; Fogel, 2012; Strauss and Thomas, 1998). While the precise mechanisms underlying variation in the aggregate across populations and over time is not clear (Deaton, 2006), within populations the fact that taller workers earn more has been widely documented, particularly in lower income settings.<sup>1</sup> This paper uses rich longitudinal survey data from Indonesia to investigate mechanisms that potentially underlie the association between height and productivity in the labor market in a low income setting.

Adult stature is largely determined in early childhood and reflects the combined influence of the early childhood environment including nutrition, disease insults, and investments in health during pregnancy and the first few years of life (Martorell and Habicht, 1986), along with selective early life mortality (Bozzoli, Deaton and Quintana-Domeque, 2009). In an important paper, Case and Paxson (2008) point out height is but one dimension of human capital that captures very early investments, height is likely to be correlated with other early childhood experiences, and possibly later life human capital investments, many of which are difficult to measure. Specifically, height is likely to be correlated with schooling attainment (Case et al., 2009), cognition (Case and Paxson, 2008), non-cognitive traits such as ambition and confidence (Persico et al., 2004), as well as an array of other markers of both the quantity and quality of health and human capital (Thomas and Strauss, 2008). Moreover, a portion of height is genetic and thus almost surely captures the role of family background and investments made across multiple generations. In low income settings, height may be a marker of strength that translates into greater productivity in physically demanding work. Disentangling that effect is complicated since workers likely self-select into occupations that reward their skills and taller, stronger workers are also likely to have better cognitive skills (Vogl, 2014). It is

---

<sup>1</sup> See, among others: Behrman et al. (2013), Case and Paxson (2008a, 2008b), Gao and Smyth (2010), Hoddinott et al (2008), Lundborg et al. (2014), Persico et al. (2004), Sohn (2015), Thomas and Frankenberg (2002), Thomas and Strauss (1997), Yamamura, Smuth and Zhand (2015) and Vogl (2014).

possible that height is a signal of worker quality or used as a screening device by employers or customers (Sohn 2014; Yamamura, Smyth and Zhang, 2015).

This research investigates each of these potential explanations for the association between height and productivity in the labor market as indicated by hourly earnings. Rather than attempt to identify a single mechanism, height is treated as one measure of human capital investments that is likely to be correlated with many others and we explore the relative contributions of the different explanations using data from a single study setting, Central Java, Indonesia. The Work and Iron Status Evaluation (WISE) was designed to provide the evidence necessary to address this question, and analyze the role of human capital in predicting success in both the formal and informal sectors. In addition to measuring height and education, WISE assesses several different domains of cognition, a battery of additional health markers, measures of family background and multiple labor market outcomes including earnings, wages, sectoral and occupational choice.

## **2. Conceptual Framework**

Height, cognition, health, family background and labor market behavior are examined within a framework that recognizes the multidimensionality and dynamics of human capital accumulation over the life course. Given early childhood production functions for multiple types of human capital, parental choices concerning nutrition and other investments interact with environmental factors and child-specific endowments to establish early levels of human capital. These parental choices reflect not only the prices and opportunity costs of differing investments, but also family background characteristics such as available resources and attitudes towards health and human capital. If different markers of human capital share common inputs, correlations between cognitive and physical development naturally arise at a young age due to parental choices geared toward maximizing their own expected future utility, that of the child and possibly the entire family.

As the individual transitions through adolescence, human capital development and skill acquisition continues through choices regarding investments in schooling and skill development along with early labor market experiences. Facing a labor market with multiple sectors and occupations with varying uncertain returns, the decision to remain in school, for example, will be a reflection of liquidity constraints, expectations and the opportunity cost of forgone work which depend on both the individual's choices and parental investments over the life course.

Similarly, earnings in differing occupations and sectors may offer specific returns to different dimensions of human capital which may vary over the short and longer term. As individuals choose to sort across formal wage work and informal self-employment, for example, their comparative advantage is a reflection not only of human capital at that point in time, but of the human capital accumulation process throughout the life course. Individuals may continue to make health and cognitive investments as they age, with opportunities for further development dependent in part on their chosen sectors and occupations.

Thinking in this framework of a life-course production process for multiple dimensions of human capital illuminates several key relationships explored in this research. The well established correlation between height and labor market productivity could be due to a number of simultaneously determined factors. Attained height as an adult reflects one's early life health, disease and nutrition environments, and likely also captures other early life investments as well as family background characteristics including financial, non-financial resources and tastes for human capital of the next generation. Similarly, shared inputs between the height, cognitive, and other human capital production functions may drive a relationship between physical and intellectual capacity. Finally, as the value of different traits may vary over time and differ depending on the sector and nature of work chosen by an individual, so too may the links between height, earnings, and additional human capital markers.

To empirically examine these relationships in the labor market, we exploit rich, longitudinal data on sectoral and occupational choice, formal and informal earnings along with high quality measures of a wide array of human capital markers.

### **3. The Work and Iron Status Evaluation**

WISE, a large-scale longitudinal study conducted in Central Java, Indonesia designed to collect detailed information on human capital and labor market outcomes, is ideally suited to investigate the relationship between height, cognition, education, health, and labor market outcomes within a population. After a listing survey in late 2001, a population-representative sample of households living in Purworejo *kabupaten* was interviewed every four months beginning in 2002 and continuing through 2005. Longer-term follow-ups were conducted five and seven years after the start of the survey in 2007 and 2009. All twelve waves of the survey are included in this study.

As the analysis relies on following individuals over time, it is imperative that selective attrition does not contaminate inferences. Attrition is extremely low in WISE: ninety-four percent of households from the 2002 baseline were re-interviewed seven years later in the 2009 wave (see Thomas et al. (2011) for a fuller discussion of tracking and attrition). We focus on 5,304 men between the ages of 25 and 65 who reported income during the survey period; there are over 38,000 person-wave observations in our panel sample.<sup>2</sup>

*Hourly earnings* Labor market outcomes are measured with great care in WISE. Each household member age 15 and older is individually interviewed to obtain detailed information on work status, employer and occupation, tenure, nature of work (tasks), and earnings in each job. Hourly earnings from wage work are calculated as total earnings from work in the market sector during the previous

---

<sup>2</sup> 7% of 25-65 year old men report no earnings and are not included in these analyses. As shown in column 1 of the Appendix table, these men are more likely to have difficulty running a kilometer, perform worse on two cognitive assessments and are very slightly better educated than those included in the analyses. There are no differences in the heights of those who do and do not report earnings in WISE.

four months divided by hours worked during the same time period. Similarly, hourly earnings from self-employment are calculated as net profits from self-employment during the prior four months divided by the number of hours worked during that time. Total hourly earnings is the sum of earnings from all jobs divided by the total number of hours worked in all jobs during the previous four months. The four-month periodicity of the survey waves in WISE is selected to coincide with the rice growing cycle, the dominant crop in the area.<sup>3</sup>

Means and standard errors of key variables are displayed in Table 1. Column 1 includes all workers and, in the other columns, the sample is stratified into those who only work in the wage sector (column 2), those who only work in the self-employed sector (column 3) and those who work in both sectors during the study period (column 4).

The first three rows of Table 1 report hourly earnings in Rp 10,000 (approximately 1 USD at the time of the survey). The average male earns Rp3,500 per hour, those working only in the wage sector earn about Rp5,000 per hour and those working only in the self-employed sector earn about Rp4,000 per hour. Those who work in both sectors earn the least: Rp2,900 per hour on average. This reflects the fact that relative to those who never switch sectors, those who do switch sectors or work in both simultaneously during the study period earn less when they work in the wage sector (row 2) but more when they work in the self-employed sector (row 3). These switchers, which accounts for over half the workers, are extremely valuable for this research as they provide an opportunity to directly address selection bias when comparisons of the height premium are drawn between those who choose to work in the wage sector relative to those who are self-employed. There are at least two reasons that working in both sectors is common in these data. First, rice farming is the dominant activity in the area and plots are, on average, less than half an acre in size.

---

<sup>3</sup> WISE collects self-reported net profits for the prior four months for those who are self-employed in a work and earnings module in the individual interview. In addition, in a separate household enterprise module, detailed information is collected about business revenues and expenditures for the prior four months. The match between the two sources of information on total profits for all household members working in the enterprise is very close. For the sake of consistency, we use the earnings data for self-employment and wage work from the individual work and earnings module.

Many farmers supplement income by working both on and off the farm during the year. Second, during the study period, there was considerable variation in weather that affected crop output with one year being a severe drought and, again, many farmers supplemented income by working off the farm.

*Height* The height of every household member is assessed at each survey wave. Adult stature is fixed until older ages, when individuals begin to shrink and so each assessment in our study sample of males age 25 to 65 should be the same. To minimize the impact of measurement error, we use the mean of measured height.<sup>4</sup>

As shown in the fourth row of Table 1, the height of the average male in the sample is 161.6 cm with those who only work in the wage sector being positively selected, those who are only self-employed being negatively selected and those in both sectors falling between those two groups. This parallels the pattern observed for hourly earnings.

*Measures of cognition* A key strength of WISE for this research is that cognitive achievement is assessed using three different, complementary instruments that have been well-validated and are designed to measure different domains of cognitive performance. In addition, each instrument was assessed in more than one survey wave so that it is possible to mitigate the impact of measurement error in the assessments.

First, the Raven's Colored Progressive Matrices pattern recognition test is designed to provide a non-verbal measure of abstract reasoning that has been interpreted as indicative of intelligence (Raven, 2000). The assessment, which does not require literacy or numeracy, involves identifying the missing part of a progression of designs from among four different options. The

---

<sup>4</sup> Height is well measured in this study. For example, one assessment of the quality of measurement is to examine the distribution of the final digit (which is mm of height). It should be uniformly distributed across all integers from 0 to 9. Stacking on preferred digits (0 and 5, for example) would indicate poor measurement protocols. There is no evidence of such stacking. 9.9% of heights end in 0 mm and 11.2% end in 5mm. Using the average of all measured heights for a respondent provides protection against transcription errors and line shifting in recording.

assessment, first developed in 1938 for clinical and general use in the U.K., is thought to be free of cultural bias and has been implemented in a very large number of studies across the globe. In WISE, different subsets of the Ravens test were administered three times to respondents age 15 and older; since performance on the assessment is unlikely to vary during adulthood, we use the average score on all the assessments in the models.

Second, an adaptation of the Philippines National Intelligence Test developed by Guthrie et al. (1977) is utilized to assess fluid intelligence. The test is similar to the Columbia Mental Maturity Scale and was originally designed for settings that are similar to the WISE study site. Like the Raven's test, the assessment is non-verbal and does not require literacy or numeracy; it differs from Raven's in that it uses images of familiar objects and activities of daily life in order for it to be more reflective of experience, logical thinking, and the ability to recognize real world patterns than the abstract Raven's matrices. Specifically, each respondent is shown a series of 5 images and asked to identify the odd man out among the 5: that is, the respondent has to identify the common elements that bind four of the images but are not shared by the fifth image. Figure 1 illustrates a sample question. The assessment has been used in other population-based surveys including the Cebu Longitudinal Health and Nutrition Survey (see Mendez and Adair, 1999; Glewwe and King, 2001; Daniels and Adair, 2005). As with the Raven's assessment, performance on this instrument is unlikely to vary during adulthood and so we use the average score on four assessments to maximize the signal in the measurement.

Third, working memory is assessed using a word recall test in which each respondent is read ten common words in Indonesian from a predetermined list. The respondent is asked to repeat back the words in any order immediately and the number the respondent remembers without prompting is recorded as the immediate word recall. The survey continues with questions about health status and, after about five minutes, the respondent is asked to recall as many words from the list as possible. The number recalled is the delayed word recall. Working memory is a core executive



function and is thought to play an important role in reasoning and decision-making and, is, therefore, potentially related to labor market success. In our models, we use the number of words recalled immediately and after a delay, averaging across surveys for the same respondents. The first time the assessment was conducted, the list of words read to the respondent was randomly assigned so that household members who are present for another members' assessment do not hear the words multiple times. Thereafter, the list of words was selected to assure that each respondent received a different assessment across waves. We use the average number of words recalled in assessments conducted in three waves of WISE. The assessment is used widely in studies of cognitive aging including, for example, the Health and Retirement Survey and related global studies in Europe and low income countries (McArdle et al., 2011; Lei et al., 2012).

Table 1 reports the within-person averages (and standard errors) for each of the cognitive assessments. The average male completed slightly over half the Ravens assessments correctly and more than 60% of the fluid intelligence assessments correctly. He remembered 4.6 of the 10 words immediately and one less word after a delay. For all of the assessments, the average is highest among those who specialize in the wage sector, lowest among those who are only ever self-employed. In order to draw comparisons across the assessments, in the regression models, all of the assessments are standardized to z-scores using the overall sample mean and standard deviation.

*Additional health assessments* In addition to measures of the attained height of individuals, the survey includes several health markers that are potentially related to labor market productivity. First, body mass index (BMI), weight (in kg) divided by height (in m) squared, is an indicator of nutritional status that, unlike height, varies throughout the life course. While extreme values of BMI are predictive of mortality and morbidity, 8% of respondents in this sample are overweight (BMI>25) and less than 0.5% are obese (BMI>30). The BMI of the average male is 20.9 m/kg<sup>2</sup> and 17.3% have BMI<18.5 and so, in this sample, lower BMI is indicative of poorer health while higher BMI is likely associated with elevated VO<sub>2</sub> max and work capacity. Higher BMI is therefore likely to be an

indicator of physical strength and endurance that is potentially valued in the labor market.

Resting blood pressure is measured for each respondent in WISE using an Omron portable automatic blood pressure monitor with upper arm cuffs of different sizes. As in many developing country settings, there are high levels of undiagnosed hypertension in Indonesia (Frankenberg et al., 2016) and very few of the WISE respondents take medication to control hypertension. We examine systolic blood pressure (SBP) along with pulse pressure which is the difference between systolic and diastolic blood pressure. Whereas systolic blood pressure is a measure of the maximum pressure on the arteries, pulse pressure is an indicator of the force that the heart generates each time it contracts. Elevated systolic and diastolic blood pressure are predictive of cardiovascular disease and pulse pressure is also indicative of hardening of artery walls (e.g. Blacher et al., 2000; Franklin et al., 1999; Mattance-Raso et al., 2004; Panagiotakos et al., 2005). The SBP of the average respondent is 125 mm Hg with 16% having SBP above 140 mm Hg, the standard cut-off for hypertension. Pulse pressure of the average respondent is 47 mm Hg and 12% have pulse pressure above 60 mm Hg which is thought to be a risk factor for elevated heart disease.

Each respondent also provides information on a battery of self-assessed Activities of Daily Living (ADLs). We focus on whether the respondent has difficulty running a kilometer to capture a key indicator of physical function that is likely related to strength and endurance. Seventeen percent of males report such difficulty, with fewer than one in ten of those who work only in the wage sector and more than one in four of those who work only in self-employment reporting difficulty running a kilometer.

#### **4. Descriptive Analyses**

All of the human capital markers are likely to be positively correlated. However, it is possible that the correlations are so high that it will not be possible to isolate independent associations with productivity in the labor market. These issues are investigated in Table 2.

Panel A reports pairwise correlations and jackknife standard errors for (log) height and the four cognitive assessments (in z scores). Taller males score significantly better on each cognitive assessment. A 5% increase in height (which is about a standard deviation increase) is associated with about a quarter of a standard deviation increase in the Raven’s and fluid intelligence scores and about a sixth of a standard deviation increase in the working memory assessments. Figure 2 displays locally weighted smoothed scatterplots with a biweight kernel and 20% bandwidth of the relationship between each cognitive assessment and (the logarithm of) height in order to assess whether there are important non-linearities in these associations. There are not. In fact, all of the correlations are positive and statistically significant. The Ravens and fluid intelligence assessments are strongly correlated and the two working memory assessments are also strongly correlated which is consistent with the pairs capturing related domains of functioning.

Panel B of the table reports regressions relating height to the four cognitive assessments (in the first column) and the Raven’s score to the other three cognitive assessments (in the second column). The regressions establish there is independent variation in height and the cognitive assessments: all but delayed working memory predict  $\ln(\text{height})$  although only 8% of the variation in  $\ln(\text{height})$  is attributable to these markers. All three cognitive assessments are significant predictors of the Raven’s score although only about one third of its variation is accounted for by the other assessments.

## 5. Human Capital and Earnings

We turn next to investigating how the association between hourly earnings and height varies in models that control education, cognition and health. Specifically, we estimate models that relate the logarithm of hourly earnings,  $\ln w_{it}$ , of individual  $i$  during the 4 months prior to the survey interview at time  $t$ :

$$\ln(w_{it}) = \beta_0 + \beta_1 \ln(ht_i) + \beta_2 cog_i + \beta_3 age_{it} + \beta_4 ed_i + \beta_5 \theta_{it} + \mu_t + \varepsilon_{it} \quad [1]$$

to  $\ln(ht_i)$ , the logarithm of average measured height of individual  $i$  and  $cog_i$  the within-person mean z score of each of the four cognitive assessments. All models include age of the respondent (specified as indicator variables for each of five-year birth cohorts). We also control education,  $ed_i$ , measured as the number of years needed for an individual to complete the highest level of schooling achieved by the respondent and, in some of the models, we will adjust for markers of health,  $\theta_{it}$ , that are potentially related to height. These include the logarithm of BMI, whether the respondent reports having difficulty running 1km, and the two blood pressure measures, SBP and pulse pressure. All of the health markers are time-varying. Wave fixed effects,  $\mu_w$ , are included to capture common aggregate conditions including seasonality and variation over time in prices and wages in the survey area. All estimates of variance-covariance matrices allow for heteroskedasticity and take into account clustering at the person level; allowing for clustering at the *desa* (village) level does not change any inferences.

Table 3 reports results from estimating equation [1] using earnings from wage work and self-employment for all male workers in columns 1 through 5; results for hourly earnings from working in the wage sector and in the self-employed sector are reported in columns 6 and 7, respectively.

We first establish the magnitude of the associations between earnings, height and cognition in the WISE sample. As shown in column 1, conditional on age and time effects, taller individuals earn more: on average, a 1 percent increase in height is associated with a 3.6 percent increase in hourly earnings. This estimate is slightly smaller than parallel estimates for monthly earnings of Indonesian males based on the Indonesia Family Life Survey, IFLS (Thomas and Frankenberg, 2002, who use the 1993 and 1997 waves of IFLS and Sohn, 2015, who uses the 2007 wave of IFLS). The difference between the IFLS and WISE estimates is driven by three features of the data. First, WISE collects detailed information on earnings and hours of work and so we are able to focus on a good measure of productivity, hourly earnings, to test hypotheses about the impact of human capital on productivity. The IFLS estimates are based on monthly earnings and therefore reflect the

relationship between height and hourly earnings as well as the relationship between height and hours of work. Taller men tend to work more hours in Indonesia. Second, the WISE sample is rural, whereas IFLS includes rural and urban respondents and the correlation between height and earnings is slightly larger among males living in urban areas in Indonesia. Third, the sample in WISE is far more homogenous than the IFLS sample which includes males from across the entire Indonesian archipelago. Differences in the organization of local labor markets, price and wage variation, heterogeneity in school access and quality, as well as genetic variation, are much less of a concern in analyses using the WISE sample relative to IFLS.

As shown in column 2, the cognitive assessments are also significant predictors of wages, both taken together and individually. Holding the other cognitive assessment scores constant, a standard deviation increase in the test score is associated with a 16 percent increase in hourly earnings for both the Raven's and fluid intelligence assessments. Working memory is an independent predictor of hourly earnings: a standard deviation increase in the number of words recalled is associated with a 12 and 5 percent increase in hourly earnings for the immediate and delayed assessments, respectively. As shown above, the cognitive assessments are all correlated and in models that do not control for other assessments, the estimated associations rise to between a 25 and 30 percent increase in hourly earnings which highlights the likely value-added for this research of multiple assessments that are designed to capture different domains of cognitive achievement.

Height, cognition and education are all included, along with age and wave effects, in the model reported in column 3 of the table. About one half of the height premium (in the first column) can be attributed to superior education and cognitive performance of taller men. Somewhat more than half the association between the cognitive scores and earnings can be attributed to education and height with the vast majority of this effect being captured by education. The cognitive assessments remain statistically significant for all but the longer term working memory assessment. Education is a significant predictor of hourly earnings with each year of education being associated

with an 8 percent increase in earnings, holding height and cognition constant. (Without those controls, the education effect is 10 percent.)

Importantly, height continues to be a significant predictor of hourly earnings even after controlling education, cognition and age. It is possible that height is capturing non-linear effects of education: allowing the effect of education to differ for each year of education in the model in column 3 has an imperceptible impact on the coefficient on  $\ln(\text{height})$ . In the model with education specified as linear in years, the estimated coefficient is 1.9 (with a standard error of 0.3) and it is 1.8 (s.e.=0.3) in a semi-parametric specification with an indicator variable for each year of education.

It is also possible that height and education are complements. This would arise, for example, if height were a proxy for strength that is of greatest value to those with the least education. The evidence is not consistent with this interpretation. Restricting the sample to those males who completed primary school or less (about half the sample), the coefficient on  $\ln(\text{height})$  is 1.77 (s.e.=0.4); the coefficient on the other half of the sample, who are better educated, is 1.94 (s.e.=0.5). If the sample is restricted to those men who completed primary school (exactly 6 years of schooling), the coefficient on  $\ln(\text{height})$  is 1.99 (se=0.4).

We have investigated whether there are important non-linearities in the relationships between height and productivity. Non-parametric estimates of the relationship between  $\ln(\text{hourly earnings})$  and  $\ln(\text{height})$  without controls, displayed in Figure 3, indicates it is well approximated by the log-log model. We do not find important non-linear effects of height in the multivariable regression models.

In sum, variation in education and the four markers of cognitive achievement is able to explain no more than half the association between hourly earnings and height in our sample of Indonesian men. We conclude that height is not simply a proxy for cognition in this setting.

Taller people tend to be healthier. In these models, height may be serving as a proxy for other health markers including, for example, strength. Time-varying health status indicators are

included in the model in column 4 and the estimated impact of height is slightly larger than in the model without health controls. As (the logarithm of) BMI rises, so do hourly earnings. This effect is primarily driven by the impact of weight, holding height constant.<sup>5</sup> Men who have difficulty with strenuous exercise earn less as do men with elevated pulse pressure. Elevated systolic blood pressure is positively associated with productivity after controlling weight and pulse pressure; this likely reflects the influence of a more sedentary lifestyle among those who have higher hourly earnings.

### **Sectoral Choice**

It is possible that employers use height as a signal of worker quality. It has also been suggested that customers use height as a signal of quality of services purchased from the self-employed. In rural Central Java, the vast majority of the self-employed are rice farmers and the quality of their product is unlikely to be deduced from their height. If height is used as a signal in our setting, then height should be more highly rewarded in the wage sector than among the self-employed. To investigate this issue, the model in column 4 is estimated separately for the (logarithm of) hourly earnings from wages (in column 5) and from self-employment profits (in column 6). Height is not just a signal in the formal sector. In fact, the effect of height on hourly earnings is substantially and significantly greater in the self-employed sector.

In part, this likely reflects selection of workers into each sector. Over half the men report working in both sectors at some point during the study. The vast majority of these men switch sectors during the study, with many making more than one switch, while some men hold multiple jobs and work in both sectors simultaneously. Panel B of Appendix Table 1 reports results from estimating a multinomial logistic regression of the probability of working in the wage sector, the self-employed sector or in both sectors. Odds ratios and associated standard errors relative to the excluded males who work in both sectors are reported in the table. Taller and better educated males

---

<sup>5</sup> In a model that includes the logarithms of height, weight and height squared, the coefficient on  $\ln(\text{weight})$  is 1.0 (s.e.=0.9) and the coefficient on the square of  $\ln(\text{height})$  is small and not statistically significant.

are more likely to work in only one sector rather than in both sectors, and they are most likely to work in the wage sector (although the height differences between those in the wage and self-employed sectors are not statistically significant). These patterns may reflect differences in the signaling value of these attributes or differences in the values of the attributes in each sector. Males with higher scores on the Raven's test and immediate word recall are less likely to work only in the self-employed sector although, conditional on these scores, those with superior longer-term working memory are more likely to be only self-employed. Those who have difficulty running a kilometer are less likely to work in both the wage and self-employed sectors.

Clearly, human capital characteristics matter for sectoral choice. We can exploit the fact that more than half the men work in both the wage and self-employed sector to assess the extent to which selection explains differences in returns to human capital across sectors. The final column of Table 3 reports estimates of a model for this sub-sample of males that includes individual fixed effects and interacts each covariate with an indicator for whether the male's earnings are from self-employment. The coefficient estimates reflect the premium (or discount) received by that male in the self-employed sector relative to the wage sector. There is no difference in the effect of height on the productivity of the same individual whether his earnings are from the wage or self-employed sector. Nor is there a difference in the return to education although both fluid intelligence and (immediate) working memory have higher returns in the self-employed sector. Neither of those attributes is easily observed and so may be more difficult to reward. There is also a higher return to BMI in the self-employed sector that likely reflects the effect of strength over and above height. The height premium in the self-employed sector in column 6 is apparently driven by selection of higher productivity, taller males into that sector. This issue is directly addressed in the next sub-section which focuses on occupational rather than sectoral choice.



## Occupational Choice

Occupational choice has been shown to depend on education, cognitive skills and height. Vogl (2014) argues that occupational sorting plays a key role in explaining the relationship between height, cognition, and earnings in Mexico. We investigate this issue directly exploiting the fact that WISE records detailed descriptions of each individual's tasks which have been classified into specific occupations at the two digit level.

As a first step, we investigate whether height and the other human capital attributes are associated with selecting into occupations in which strength is likely to be rewarded, specifically, agriculture, production work such as masonry and manual transportation operation (which is mostly bicycle rickshaws). These occupations account for 65 percent of the sample. The first column of Table 4 reports coefficients from a linear probability model with an indicator for working in one of these occupations as the dependent variable. Males who are taller, more educated, and score higher on Raven's exams are less likely to work in occupations that likely reward strength.

Second, we re-estimate the models of hourly earnings including occupation fixed effects. Occupations have been aggregated into the following categories: professional, teachers, administrative, clerical, sales, services, agriculture, manual production, transport operation, military, and students. Results are reported in columns 2 through 4 of Table 4 which add occupation fixed effects to the specifications reported in columns 4 through 6 in Table 3. The occupation fixed effects are statistically significant, but they explain only a small part of the height premium. Occupational selection is slightly more important among the self-employed, for whom the height premium is reduced by 12%, than those in the wage sector, for whom the height premium is reduced by only 4%. Occupational choice does not explain the height premium in this setting.<sup>6</sup>

---

<sup>6</sup> Panel C of Appendix Table 1 repeats these analyses without controlling other dimensions of health which may be correlated with unobserved factors that also affect occupational choice. The estimated effects of height, and our conclusions, are not substantially affected.

## **Family Background**

All of the dimensions of human capital that we have examined – height, cognition, education, and health – have been shown to be related to earnings of males in rural Indonesia. There is at least one key difference between height and all the other human capital markers: height is largely determined in the first few years of life and depends critically on inputs during that period. All the other markers of human capital likely depend on those inputs but also inputs through the rest of childhood and adolescence, as well as adulthood in some cases.

In an effort to investigate the role family background plays in the relationships described above, we adopt two complementary approaches. First, we include controls for parental human capital and, second, we estimate models that include family background fixed effects. These approaches exploit several features of the design of WISE. First, all adults report the education of their parents in the survey, whether or not the parent is alive, so that controlling parental education does not impose any selection rule on the sample. Second, when a respondent moves out of a baseline household, the respondent is followed to his/her new location and interviewed there. (Attrition in WISE is less than 2%.) Third, when a person joins a household, he or she is interviewed (and measured) as part of the survey. As a result of the second and third design features, there are a large number of adult siblings in the study sample although it is important to recognize that they are not a random sub-sample of the population.

Our first approach, reported in column 5 of Table 4, extends the model of hourly earnings to include parental education (which is recorded for all respondents in WISE). Overall, the effects of own human capital are little affected. The estimated return to own education is reduced by about 10% as is the effect of the Raven's score. The effect of height is not reduced.

A complementary strategy is to examine collections of siblings who have shared genetic and environmental backgrounds. To assess the selectivity of this sample for interpretation of the link between human capital and hourly earnings, we re-estimate the model in column 5 with the reduced

sample of siblings. The differences in the effects of height, cognition and education in the two models are modest and they are slightly smaller in the selected sample. A key genetic trait shared by the siblings is parental height which is also a strong predictor of child height. Parental heights are included in the model in column 7: the estimated own height premium is not substantially changed. Finally, we absorb all shared genetic and background differences between siblings by adding a fixed effect for each mother in the final column of the table. (Divorce and remarriage is very uncommon in Central Java and so this approach amounts to including parent fixed effects in the model.)

Identification depends on differences between siblings in human capital and labor market outcomes. These differences are much smaller than those in the sample overall. A substantial part of the height premium is driven by shared background although hourly earnings are 1% higher for each percentage difference in height between brothers. This effect is significant at a 10% size of test. Differences in education and the Raven's score are also significant predictors of differences in the earnings of siblings. These siblings differences likely reflect differences in the environment, resources and parental investments during the child life, with the height differences driven by exposures in the first few years of each child's life.

## **6. Conclusions**

This research has established that, in rural Indonesia, over and above height, multiple dimensions of human capital are rewarded in the labor market for males. These include educational attainment, several different indicators of cognitive performance, and other dimensions of health status. After controlling these indicators of human capital, we have shown that taller men are more productive as measured by hourly earnings in both the wage and self-employed sectors, and that these returns are greater among those who select to work in the self-employed sector. We have also shown that while height does predict occupational choice, taller men earn a premium within occupations. To wit, in our study setting, height does not appear to be a proxy for cognition or strength, to the extent they

are well-measured in our study. This contrasts with evidence from more advanced economies where returns to height in the labor market are much smaller in magnitude and height has been shown to be a proxy for cognition (Case and Paxson, (2008).

In a household-based survey, it is difficult to distinguish the influence of height from that of other dimensions of family background, including genetic traits; to the extent we are able to take family background into account, height appears to have an independent impact on productivity. This is, perhaps, not entirely surprising. Not only are taller men happier, healthier and more productive, but they likely benefited from a reduced burden of infection and inflammation during the first few years of life that construed many benefits (Crimmins and Finch, 2006). These early life benefits are likely to be especially important in low resource settings and they are not likely to be fully captured by the choice of type of work of an individual in adulthood or by even more extensive measures of cognitive performance, educational attainment, and health status measured in adulthood than those that are used in this study. Height likely reflects the combination of the environment, resources and choices of parents during the first few years of a child's life that have long lasting effects on health and well-being even as they are likely correlated with other investments over the life course.

## References

- Barham, T., K. Macours, and J. A. Maluccio. More schooling and more learning? effects of a 3-year conditional cash transfer program in Nicaragua after 10 years. Working Paper, October 2012.
- Behrman, J. R., J. Hoddinott, J. A. Maluccio, and R. Martorell. Brains versus brawn: Labor market returns to intellectual and health human capital in a poor developing country. Working Paper, May 2013.
- Blacher, J., J. A. Staessen, X. Girerd, J. Gasowski, L. Thijs, L. Liu, J. G. Wang, R. H. Fagard, and M. E. Safar. Pulse pressure not mean pressure determines cardiovascular risk in older hypertensive patients. *Archives of Internal Medicine*, 160(8): 1085, 2000.
- Bozzoli, C., A. Deaton and C. Quintana-Domeque. Adult height and childhood disease. *Demography*, 46.4:647-69. 2009.
- Case, A. and C. Paxson. Height, health, and cognitive function at older ages. *American Economic Review: Papers and Proceedings*, 98(2): 463–467, 2008.
- Case, A. and C. Paxson. Stature and status: Height, ability, and labor market outcomes. *The Journal of Political Economy*, 116(3): 499, 2008.
- Case, A., C. Paxson, and M. Islam. Making sense of the labor market height premium: Evidence from the British household panel survey. *Economics Letters*, 102(3):174–176, 2009.
- Crimmins, E. and C. Finch. Infection, inflammation, height and longevity. *Proceedings of the National Academy of Sciences*, 103.2:498-503. 2006.
- Daniels, M. and L. Adair. Breast-feeding influences cognitive development in Filipino children. *The Journal of Nutrition*, 135(11): 2589–2595, 2005.
- Deaton. A. Height, health and development. *Proceedings of the National Academy of Sciences*. 104.33:13232-7. 2007.
- Deaton, A. and R. Arora. Life at the top: The benefits of height. *Economics and Human Biology*. 7:133-136. 2009.
- Fogel, R. W. *Explaining Long-term Trends in Health and Longevity*. Cambridge: Cambridge University Press. 2012.
- Frankenberg, E., J. Y. Ho and D. Thomas. Biological health risks and economic development. In J. Komlos and I. R. Kelly (eds.), *The Oxford Handbook of Economics and Human Biology*. Oxford University Press. 2016.
- Franklin, S., S. Khan, N. Wong, M. Larson, and D. Levy. Is pulse pressure useful in predicting risk for coronary heart disease? The Framingham Heart Study. *Circulation*, 100(354–360), 1999.
- Gao, W. and R. Smyth. Health human capital, height and wages in China. *Journal of Development Studies*, 46(3):466–484, 2010.
- Glewwe, P. and E. King. The impact of early childhood nutritional status on cognitive development: Does the timing of malnutrition matter? *The World Bank Economic Review*, 15(1): 81–113, 2001.
- Guthrie, G. M., A. H. Tayag, and P. J. Jacobs. The Philippine nonverbal intelligence test. *The Journal of Social Psychology*, 102(1):3–11, 1977.
- Hausman, J. and D. McFadden. Specification Tests for the Multinomial Logit Model. *Econometrica*, 52(5): 1219-1240, 1984.

- Huang, W., X. Lei, G. Ridder, J. Strauss, and Y. Zhao. Health, height, height shrinkage and SES at older ages: Evidence from China. *American Economic Journal: Applied Economics*, 5 (2), 86-121, 2013.
- Hoddinott, J., J. Maluccio, J. R. Behrman, R. Flores, R. Martorell. Effect of a nutrition intervention during early childhood on economic productivity in Guatemalan adults. *The Lancet*. 371.9610:411-16. 2008.
- Lei, X., Y. Hu, J. J. McArdle, J. P. Smith, and Y. Zhao. Gender differences in cognition among older adults in China. *Journal of Human Resources*, 47(4): 951–971, 2012.
- Lindqvist, E. and R. Vestman. The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment.” *American Economic Journal: Applied Economics*, 3(1): 101–28, 2011.
- Lundborg, P., P. Nystedt, and D.-O. Rooth. Height and Earnings: The role of cognitive and noncognitive skills. *Journal of Human Resources*, 49(1): 141-166, 2014.
- Martorell, R., Habicht, J.-P. Growth in early childhood in developing countries. In: Falkner, F., Tanner, J.M. (Eds.), *Human Growth, vol. 3, Methodology. Ecological, Genetic, and Nutritional Effects on Growth*. second ed. Plenum, New York, pp. 241–262, 1986.
- Mattace-Raso, F., T. van der Cammen, N. van Popele, D. van der Kuip, M. Schalekamp, A. Hofman, M. Breteler, and J. Witteman. Blood pressure components and cardiovascular events. *Journal of the American Geriatric Society*, 52:1538–1542, 2004.
- McArdle, J. J., J. P. Smith, and R. Willis. Cognition and economic outcomes in the Health and Retirement Survey. In D. Wise, editor, *Explorations in the Economics of Aging*, pages 209–236. University of Chicago Press, 2011.
- Mendez, M. and L. Adair. Severity and timing of stunting in the first two years of life affect performance on cognitive tests in late childhood. *The Journal of Nutrition*, 129(8): 1555–1562, 1999.
- Panagiotakos, D. Kromhout, A. Menotti, C. Chrysohoou, A. Dontas, C. Pitsavos, H. Adachi, H. Black-burn, S. Nedeljkovic, and A. Nissinen. The relation between pulse pressure and cardiovascular mortality in 12,763 middle-aged men from various parts of the world: A 25-year follow-up of the seven countries study. *Archives of Internal Medicine*, 165: 2142–2147, 2005.
- Persico, N., A. Postlewaite, and D. Silverman. The effect of adolescent experience on labor market outcomes: The case of height. *Journal of Political Economy*, 112(5): 1019–1053, 2004.
- Pitt, M., M. Rosenzweig, and N. Hassan. Human capital investment and the gender division of labor in a brawn-based economy. *American Economic Review*, 2012.
- Raven, J. The Raven’s progressive matrices: Change and stability over culture and time. *Cognitive Psychology*, 41:1–48, 2000.
- Safar, ME, S. St Laurent, A. Safavian, B. Pannier, G. London, Pulse pressure in sustained essential hypertension: a haemodynamic study, *Journal of Hypertension*: 5(2): 213-8, 1987.
- Shih, R. A., J. Lee, and L. Das. Harmonization of cross-national studies of aging to the Health and Retirement Study. RAND Labor and Population Working Paper 94-08, July 2011.
- Sohn, K. The height premium in Indonesia. *Economics and Human Biology*, 16:1-15. 2015.
- Spears, D. Height and cognitive achievement among Indian children. *Economics and Human Biology*, 10: 210–219, 2012.

- Strauss, J. and D. Thomas. Health over the life course. In T. P. Schultz and J. Strauss, editors, *Handbook of Development Economics*, volume 4. Elsevier, 2008.
- Thomas, D. and E. Frankenberg. Health, nutrition and prosperity: A microeconomic perspective. *Bulletin of the World Health Organization*, 80.2:106-13. 2002.
- Thomas, D. and J. Strauss. Health and wages: Evidence on men and women in urban Brazil. *Journal of Econometrics*, 77(1):159–185, 1997.
- Yamamura, E., R. Smyth and Y. Zhang. Decomposing the effect of height on income in China: The role of market and political channels. *Economics and Human Biology*, 19:62-74, 2015.
- Vogl, T. Height, skills, and labor market outcomes in Mexico. *Journal of Development Economics*. 107:84-96. 2014.

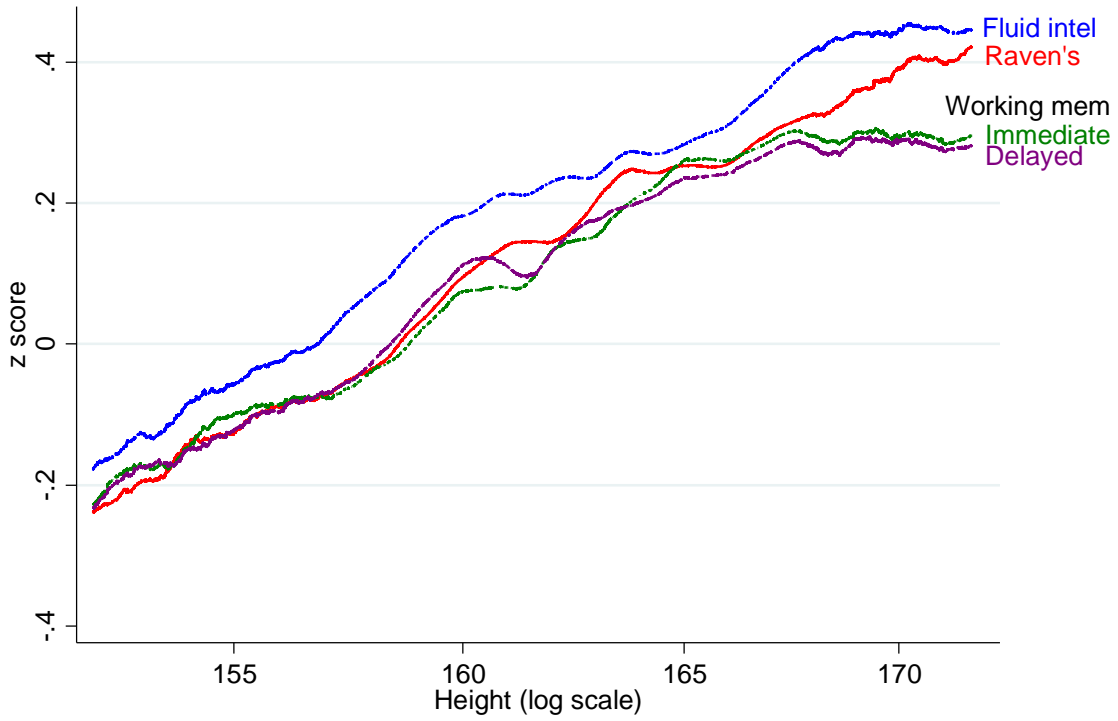
**Figure 1:** Example of a question from the assessment of fluid intelligence .



Note: Each respondents is asked to circle the picture that does not belong with the others.  
(The correct answer is the middle picture.)

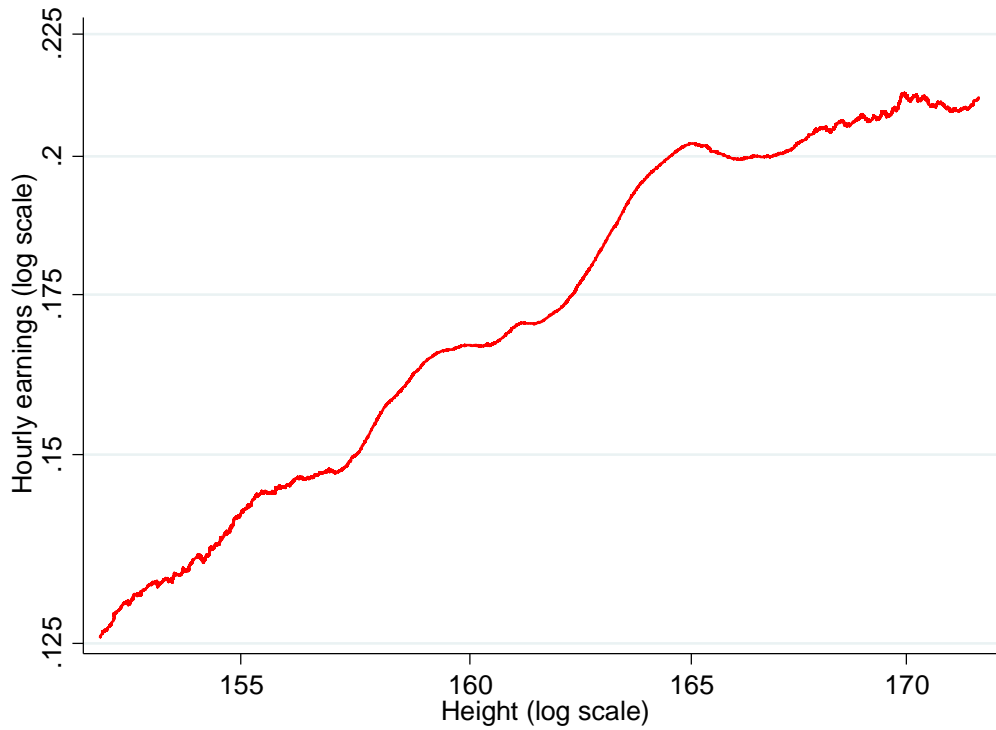


Figure 2. Relationship between cognitive assessments and height



Note: Locally weighted smoothed scatterplots with biweight kernel and 20% bandwidth.

Figure 3. Relationship between hourly earnings and height



Note: Locally weighted smoothed scatterplots with biweight kernel and 20% bandwidth.

**Table 1. Sample description**

	<i>Individual Works in [...]</i>			
	<b>All</b>	<b>Wage Sector</b>	<b>Self-employed</b>	
	<b>(1)</b>	<b>Only</b>	<b>Sector Only</b>	<b>Both Sectors</b>
		<b>(2)</b>	<b>(3)</b>	<b>(4)</b>
Hourly earnings (Rp0,000)				
All work	0.35	0.50	0.40	0.29
	(0.03)	(0.02)	(0.10)	(0.01)
From work in wage sector	0.40	0.50		0.37
	(0.03)	(0.02)		(0.03)
From self-employment	0.44		0.40	0.48
	(0.05)		(0.10)	(0.06)
Height (cm)	161.63	163.70	160.59	161.47
	(0.09)	(0.21)	(0.17)	(0.12)
Raven's Test (% correct)	53.56	65.28	46.52	52.99
	(0.36)	(0.87)	(0.69)	(0.46)
Fluid Intelligence (% correct)	61.26	68.53	55.04	61.98
	(0.34)	(0.90)	(0.66)	(0.42)
Working memory: Immediate (correct out of 10)	4.62	5.26	4.26	4.62
	(0.02)	(0.05)	(0.04)	(0.03)
Working memory: Delayed (correct out of 10)	3.56	4.27	3.20	3.53
	(0.02)	(0.06)	(0.04)	(0.03)
Age	41.36	31.93	48.17	41.26
	(0.17)	(0.30)	(0.35)	(0.21)
Years of Education	8.25	10.63	7.23	7.93
	(0.06)	(0.12)	(0.11)	(0.08)
Body Mass Index	20.87	20.99	20.68	20.92
	(0.04)	(0.09)	(0.08)	(0.05)
Difficulty Running 1km (%)	16.89	9.00	26.80	14.71
	(0.51)	(0.91)	(1.17)	(0.66)
Systolic BP (mm Hg)	125.18	123.59	127.88	124.36
	(0.25)	(0.51)	(0.55)	(0.33)
Pulse Pressure (mm Hg)	46.91	45.22	49.04	46.40
	(0.18)	(0.40)	(0.39)	(0.24)
N. Individual-Wave Obs.	38,430	4,521	8,576	34,274
N. Individuals	5,304	1,000	1,429	2,875

**Table 2. Correlations among human capital markers**

<b>A. Pairwise correlations between height and cognitive assessments</b>				
	<b>Cognitive assessments (z scores)</b>			
	<b>Raven's Prog Matrices (1)</b>	<b>Fluid Intelligence (2)</b>	<b>Working memory Immediate (3)</b>	<b>Working memory Delayed (4)</b>
ln(height)	0.241*** (0.013)	0.243*** (0.013)	0.183*** (0.014)	0.167*** (0.014)
Raven's score		0.55*** (0.010)	0.405*** (0.011)	0.394*** (0.011)
Fluid intelligence			0.418*** (0.012)	0.395*** (0.012)
Working memory: Immediate				0.777*** (0.007)
<b>B. Multivariable correlations between height and cognitive assessments</b>				
	<b>ln(height) (1)</b>	<b>Raven's score (2)</b>		
Raven's score	0.549*** (0.066)			
Fluid intelligence	0.599*** (0.070)	0.482*** (0.014)		
Working memory: Immediate	0.257*** (0.091)	0.128*** (0.018)		
Working memory: Delayed	0.048 (0.090)	0.124*** (0.018)		
R <sup>2</sup>	0.079	0.345		

**Notes:** Sample is 5,304 males. Cognitive scores are averages for each respondent across assessments and converted to z scores using the overall the sample mean and standard deviation. Standard errors in parentheses robust to arbitrary forms of heteroskedasticity \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table 3. Labor market returns to human capital**

	ln(hourly earnings)				In hourly earnings from wage work	In hourly earnings from self-employment	Self-employment premium rel to wages
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Height (logarithm)	3.636*** (0.345)		1.874*** (0.304)	1.942*** (0.299)	1.554*** (0.285)	2.266*** (0.401)	0.331 (0.539)
Cognitive assessments (z scores)							
Raven's progressive matrices		0.166*** (0.017)	0.077*** (0.015)	0.073*** (0.015)	0.056*** (0.015)	0.062*** (0.021)	-0.002 (0.030)
Fluid intelligence assessment		0.164*** (0.017)	0.056*** (0.016)	0.049*** (0.016)	0.008 (0.016)	0.084*** (0.022)	0.105*** (0.030)
Working memory: Immediate		0.127*** (0.021)	0.052*** (0.019)	0.045** (0.019)	0.038** (0.020)	0.049* (0.026)	0.077** (0.036)
Working memory: Delayed		0.045** (0.020)	0.028 (0.018)	0.026 (0.018)	0.016 (0.019)	0.024 (0.024)	-0.030 (0.034)
Completed education (years)			0.083*** (0.003)	0.075*** (0.003)	0.070*** (0.003)	0.059*** (0.005)	0.009 (0.006)
Health indicators							
BMI (logarithm)				1.016*** (0.096)	0.791*** (0.101)	1.139*** (0.130)	0.344* (0.200)
(1) if difficulty running 1km				-0.049** (0.024)	-0.035 (0.026)	-0.016 (0.030)	0.030 (0.035)
Systolic BP				0.025*** (0.009)	0.024*** (0.009)	0.009 (0.011)	-0.005 (0.014)
Pulse Pressure				-0.045*** (0.011)	-0.043*** (0.010)	-0.022 (0.015)	0.004 (0.017)
(1) if self-employment earnings							-2.893 (2.870)
Age Controls	y	y	y	y	y	y	
Wave FE	y	y	y	y	y	y	y
Individual FE							y
Observations	38,430	38,430	38,430	38,430	21,119	26,252	17,856

Note: Standard errors in parentheses clustered at the individual level. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table 4. Occupational choice and family background**

	<b>A. Occupational Choice</b>				<b>B. ln(hourly earnings) and family background</b>			
	<b>LPM</b>	<b>Returns to Human Capital</b>			<b>All males</b>	<b>Males w/ at least one sibling in sample</b>		
	<b>Occupation that likely rewards strength</b>	<b>ln(hourly earnings)</b>	<b>ln(hourly wages)</b>	<b>ln(hourly earnings from self-emp)</b>	<b>Include parental education</b>	<b>Include parental education</b>	<b>Include parental education &amp; height</b>	<b>Include mother fixed effects</b>
	<b>(1)</b>	<b>(2)</b>	<b>(3)</b>	<b>(4)</b>	<b>(5)</b>	<b>(6)</b>	<b>(7)</b>	<b>(8)</b>
Log Height	-0.528*** (0.159)	1.730*** (0.276)	1.494*** (0.267)	1.983*** (0.380)	1.958*** (0.297)	1.867*** (0.549)	1.813*** (0.546)	1.190* (0.716)
Raven's Score	-0.016** (0.008)	0.057*** (0.014)	0.053*** (0.014)	0.056*** (0.020)	0.065*** (0.015)	0.051** (0.024)	0.058** (0.023)	0.114** (0.048)
Fluid Intelligence Score	0.003 (0.008)	0.050*** (0.015)	0.009 (0.015)	0.083*** (0.021)	0.054*** (0.016)	0.050** (0.024)	0.046* (0.024)	-0.010 (0.062)
Work memory: Immed	-0.013 (0.009)	0.029* (0.018)	0.030 (0.018)	0.031 (0.024)	0.044** (0.019)	0.030 (0.029)	0.029 (0.029)	0.001 (0.053)
Work memory: Delayed	-0.017* (0.009)	0.019 (0.017)	0.013 (0.018)	0.019 (0.023)	0.023 (0.018)	0.046 (0.028)	0.049* (0.029)	0.051 (0.069)
Years of Education	-0.036*** (0.002)	0.047*** (0.003)	0.045*** (0.003)	0.041*** (0.004)	0.067*** (0.003)	0.054*** (0.006)	0.058*** (0.005)	0.019** (0.008)
Log BMI	-0.405*** (0.047)	0.758*** (0.090)	0.638*** (0.098)	0.855*** (0.125)	0.979*** (0.095)	0.501*** (0.151)	0.509*** (0.150)	0.469 (0.417)
(1) if difficulty run 1km	-0.024** (0.010)	-0.063*** (0.023)	-0.053** (0.024)	-0.034 (0.029)	-0.042* (0.024)	0.069 (0.044)	0.066 (0.044)	0.109*** (0.041)
Systolic Bp	-0.016*** (0.004)	0.010 (0.008)	0.015* (0.008)	0.000 (0.011)	0.023*** (0.009)	0.012 (0.015)	0.013 (0.015)	0.022* (0.013)
Pulse Pressure	0.033*** (0.005)	-0.018* (0.010)	-0.029*** (0.010)	-0.001 (0.014)	-0.041*** (0.011)	-0.031* (0.016)	-0.035** (0.016)	-0.049*** (0.013)
Age controls	y	y	y	y	y	y	y	y
Wave FE	y	y	y	y	y	y	y	y
Occupation FE		y	y	y				
Family background controls					y	y	y	
Mother FE								y
Sample size	38,430	38,430	21,119	26,252	38,430	11,789	11,789	11,789

Note: Standard errors in parentheses clustered at the individual level. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Appendix Table 1. Sample selection, sectoral choice and occupational choice excluding health status**

	A. Selection into the sample	B. Sectoral choice		C. Adjusting for occupational choice		
	Linear probability model	Multinomial logit (Odds ratios relative to both sectors)		Effects of human capital on ln(hourly earnings) without controlling health		
	Pr(no earnings) (1)	Only wages (2)	Only self-emp (3)	All (4)	Wages (5)	Self-emp (6)
Log Height	-0.022 (0.087)	1.017** (0.008)	1.012** (0.006)	1.681*** (0.277)	1.428*** (0.266)	1.920*** (0.382)
Raven's Score	-0.027*** (0.005)	1.059 (0.059)	0.909* (0.044)	0.058*** (0.014)	0.053*** (0.014)	0.059*** (0.020)
Fluid Intelligence Score	-0.009* (0.005)	0.964 (0.060)	0.924 (0.050)	0.055*** (0.015)	0.013 (0.015)	0.087*** (0.021)
Work memory: Immediate	-0.004 (0.005)	1.014 (0.078)	0.894* (0.056)	0.034* (0.018)	0.036* (0.019)	0.034 (0.024)
Delayed	-0.002 (0.005)	1.117 (0.080)	1.125* (0.070)	0.020 (0.017)	0.013 (0.018)	0.018 (0.023)
Years of Education	0.005*** (0.001)	1.122*** (0.015)	1.023** (0.011)	0.051*** (0.003)	0.047*** (0.003)	0.045*** (0.005)
Log BMI	-0.045 (0.030)	0.624 (0.228)	1.191 (0.362)			
(1) if difficulty run 1km	0.043*** (0.010)	1.297* (0.186)	1.206** (0.112)			
Systolic Bp	0.000 (0.003)	1.017*** (0.004)	1.002 (0.003)			
Pulse Pressure	0.007 (0.004)	0.987** (0.005)	0.992* (0.005)			
Age controls	y		y	y	y	y
Wave FE	y		y	y	y	y
Occupation FE	y		y	y	y	y
Observations	5,741		5,304	38,430	21,119	26,252
R-squared	0.218			0.228	0.300	0.159

Note: Standard errors in parentheses clustered at the individual level except in MNL model. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1