Economics Bulletin

Volume 36, Issue 3

# A fresh look at the labor market height premium in Germany

Frieder Kropfhäußer
*Halle Institute for Economic Research*

## Abstract

I use data from the German Socioeconomic Panel Study (SOEP) to analyze the relationship between height and wages in a sample of young German workers. My results show that the crude height wage premium documented in the literature is explained by unobserved heterogeneity on the sibling level. This contradicts the findings of a labor market height premium in Germany using OLS and Hausman-Taylor estimators as well as the Swedish finding of a height effect remaining after controlling for sibling fixed effects.

E-Mail: frieder.kropfhaeusser@iwh-halle.de. Research of Frieder Kropfhäußer was funded by the European Regional Development Fund through the program 'Investing in your Future'. There are no conflicts of interest.
**Citation:** Frieder Kropfhäußer, (2016) "A fresh look at the labor market height premium in Germany", *Economics Bulletin*, Volume 36, Issue 3, pages 1376-1383
**Contact:** Frieder Kropfhäußer - Frieder.Kropfhaeusser@iwh-halle.de.
**Submitted:** May 18, 2016. **Published:** July 17, 2016.

# 1. Introduction

It is a robust empirical finding that there is a positive association between height and socioeconomic status in developed economies. As shown in Figure 1, there is also a strong positive correlation between height and the share of men and women in high skilled occupations, height and average earnings and a strong negative correlation between height and the share of men and women in low skilled occupations in the SOEP.

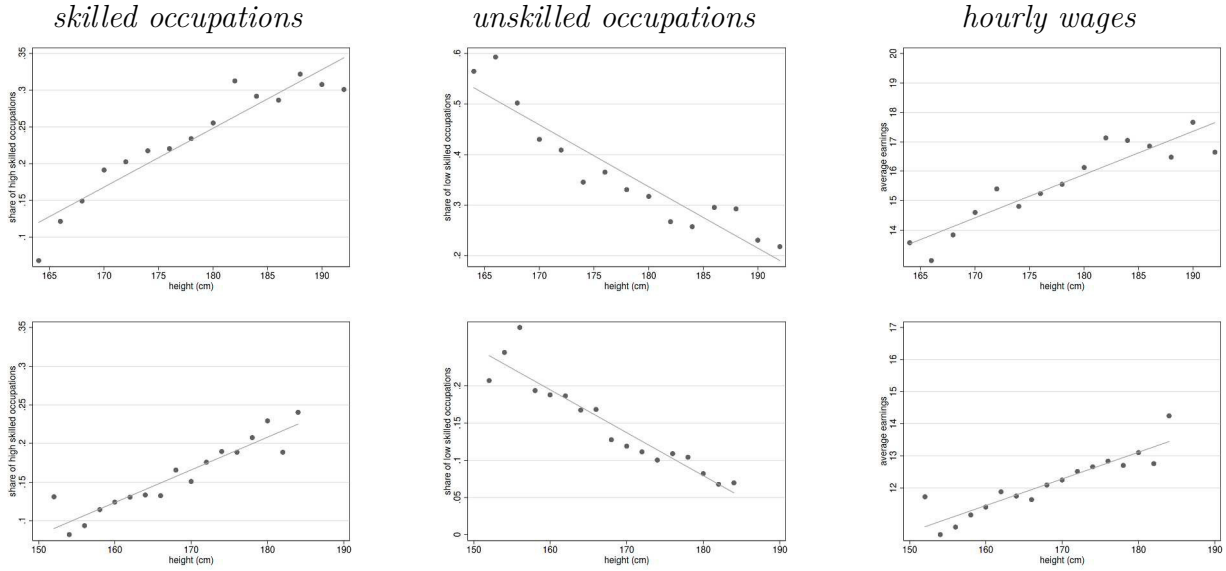


Figure 1: Shares of workers in skilled and unskilled occupations, and average hourly wages by height: men (first row) and women (second row). *High skilled occupations are defined as occupations that fall into ISCO88 major group one (managers) and eight (professionals) while low skilled occupations fall into major group eight (plant and machine operators, and assemblers) and nine (elementary occupations).*

The crude labor market height premium has been largely explained by cognitive and noncognitive ability as well as health status, whereas the remaining effect has been attributed to preferential treatment of tall workers Lundborg et al. (2014). An association between height and cognitive ability is given by early life health and nutrition, which strongly affect whether the genetic potential in height and ability given will be reached as an adult (Case and Paxson, 2008). Furthermore, height affects how people are perceived and how people behave and therefore contributes to the buildup of noncognitive abilities (Persico et al., 2004).

Studies using German data have mainly focused on the identification of a crude height-wage association using OLS and Hausman-Taylor estimation (Heineck, 2005; Rietveld et al., 2015). The mediating role of noncognitive ability and health status has been largely unexplored in theses studies. I follow a different strategy, showing that shared family background explains all the crude height effect for young German workers. The sibling fixed effect picks up shared genetic endowment (biological siblings share on average 50% of their genes) and nongenetic components such as parenting style and parental socioeconomic status.

Table I: Descriptive Statistics

| | Full Sample | | | | Sibling Sample | | | |
|---|---|---|---|---|---|---|---|---|
| | Men | | Women | | Men | | Women | |
| Variable | Mean | S.d. | Mean | S.d. | Mean | S.d. | Mean | S.d. |
| *real hourly wage* | 16.127 | 10.609 | 12.077 | 8.046 | 13.665 | 6.746 | 11.366 | 8.919 |
| *real monthly earnings* | 3026.339 | 1966.093 | 1649.261 | 1198.890 | 2518.018 | 1413.967 | 1617.681 | 981.270 |
| *height (cm)* | 179.146 | 7.045 | 166.364 | 6.262 | 179.714 | 7.000 | 167.536 | 6.248 |
| baseline controls: | | | | | | | | |
| *age* | 42.855 | 10.388 | 42.618 | 10.253 | 32.976 | 7.864 | 30.729 | 6.855 |
| *GDR* | 0.252 | 0.434 | 0.284 | 0.451 | 0.216 | 0.412 | 0.270 | 0.444 |
| *migration background* | 0.114 | 0.317 | 0.103 | 0.303 | 0.139 | 0.346 | 0.111 | 0.314 |
| *disabled* | 0.068 | 0.251 | 0.052 | 0.222 | 0.036 | 0.187 | 0.026 | 0.159 |
| cognitive skills: | | | | | | | | |
| *ISCED 1* | 0.071 | 0.257 | 0.088 | 0.283 | 0.115 | 0.319 | 0.092 | 0.289 |
| *ISCED 2* | 0.521 | 0.500 | 0.528 | 0.499 | 0.537 | 0.499 | 0.504 | 0.500 |
| *ISCED 3* | 0.069 | 0.254 | 0.106 | 0.308 | 0.093 | 0.290 | 0.135 | 0.342 |
| *ISCED 4* | 0.094 | 0.291 | 0.068 | 0.251 | 0.081 | 0.273 | 0.078 | 0.268 |
| *ISCED 5* | 0.245 | 0.430 | 0.210 | 0.407 | 0.174 | 0.379 | 0.191 | 0.393 |
| *mother Abitur* | 0.066 | 0.249 | 0.060 | 0.237 | 0.056 | 0.230 | 0.078 | 0.269 |
| *father Abitur* | 0.117 | 0.322 | 0.112 | 0.316 | 0.091 | 0.287 | 0.144 | 0.351 |
| noncognitive skills (big 5): | | | | | | | | |
| *openness* | 0.078 | 0.872 | 0.089 | 0.932 | 0.370 | 0.312 | 0.426 | 0.327 |
| *conscientiousness* | 0.113 | 0.887 | 0.106 | 0.936 | 0.402 | 0.252 | 0.424 | 0.257 |
| *extraversion* | 0.074 | 0.878 | 0.106 | 0.942 | 0.405 | 0.306 | 0.447 | 0.329 |
| *agreeableness* | 0.063 | 0.872 | 0.094 | 0.935 | 0.378 | 0.263 | 0.437 | 0.257 |
| *neuroticism* | 0.028 | 0.857 | 0.109 | 0.940 | 0.299 | 0.315 | 0.470 | 0.330 |
| health: | | | | | | | | |
| *health satisfaction* | 7.019 | 1.931 | 6.957 | 1.998 | 7.448 | 1.784 | 7.291 | 1.877 |
| Observations | 58,331 | . | 59,597 | . | 6,201 | . | 4,912 | . |
| Persons | 10,178 | . | 11,374 | . | 964 | . | 894 | . |

# 2. Data and method

Summary statistics are given in Table I. The sample is an unbalanced panel of employed individuals aged 21-65 years from the 2001-2014 SOEP waves. Pensioners, civil servants and self-employed are disregarded. To minimize the impact of random reporting errors in self-reported height, its means are calculated for each respondent in the sample.

The height-wage association is estimated based on the following equation:

$$\begin{aligned} ln(w_{ij}) = \beta_0 \quad &+ \quad \beta_1 height_{ij} + \boldsymbol{x_{ij}}'\boldsymbol{\delta_1} + (\beta_2 height_{ij} + \boldsymbol{x_{ij}}'\boldsymbol{\delta_2}) \times female_{ij} \\ &+ \quad \beta_3 female_{ij} + \mu_j + \varepsilon_{ij}, \end{aligned} \tag{1}$$

where $height_{ij}$ stands for body height of respondent $i$ of family $j$, $\boldsymbol{x_{ij}}$ is a vector of exogenous baseline control variables plus year and region dummies and $\mu_j$ is a family fixed effect absorbing unobserved time invariant family-specific characteristics common to all siblings within the same family (e.g. food and nutrition supply in the home, parental practices, and preferences) and shared genetic endowment. As pointed out by Lundborg et al. (2014), the height effect would be biased if there are genes that affect height and wages simultaneously. Since biological siblings share on average 50% of their genes, the sibling fixed effect partly controls for this and because genes are inherited randomly, it is to expect that the part of the genetically determined variation in height that remains when controlling for sibling fixed effects, is exogenous across siblings.

In the literature, height has been shown to be correlated with cognitive and noncognitive ability as well as health status. Hence, $\boldsymbol{x_{ij}}$ is sequentially appended by proxies for these variables. Due to the shared family background and genetic endowment of siblings,

the sibling fixed effect and educational attainment as proxy for cognitive ability should be highly correlated. And since educational attainment is only a rough outcome measure of cognitive ability, I expect the sibling fixed effect to additionally capture domains of cognitive ability that educational attainment does not reflect.

The natural logarithms of real gross monthly wages and real gross hourly wages serve as dependent variables. Hourly wages are calculated from labor income divided by hours worked. Monthly and hourly wages are deflated by the consumer price index included in the SOEP.

Men and women are pooled in order to not to restrict the sibling sample to same-sex siblings. I allow for gender-specific effects of the controls by interacting height and all explanatory variables in $\boldsymbol{x_{ij}}$ with the $female_{ij}$ dummy. Note that $female_{ij}$ not only accounts for the respondent's gender, but also for the gender composition of the corresponding sibling groups in the dataset.

The literature has consistently shown that cognitive ability is one of the key drivers of the labor market height premium (Case and Paxson, 2008; Case et al., 2009; Lundborg et al., 2014). Ordered logit models are used to show that the correlation between height and educational attainment (measured in ISCED categories) also persists in the SOEP data. Because educational attainment changes little between the survey waves, I use only the first observation per person in the ordered logit estimations. The Blow-up and cluster estimator is employed to allow for sibling fixed effects in the ordered logit framework (Baetschmann et al., 2015).

# 3. Results

Table II presents the results for the association between height and educational attainment. Accordingly, an increase in height of one centimeter is associated with an increase in the odds of being in a higher ISCED category of 2.5% for women (column (2)) and 3.19% for men (column(3)). Results for the pooled sample show that this difference is statistically significant at the 10% level. When controlling for sibling fixed effects, height effects become insignificant pointing to shared family background as a channel of transmission of the effect of height on education. To clarify that the finding of an insignificant height coefficient is not due to the reduction in sample size, I re-estimate the ordered logit model in the sibling sample without sibling fixed effects (column 5). However, the height coefficient for men in columns (1) and (5) remain remarkably similar.

Results for the height premium are given by Table III. The first column presents the crude height effect on log real hourly and monthly wages for the full sample, whereas from column (2) onwards the estimation sample is restricted to the sibling sample. Column (1) shows a highly significant marginal effect of height on hourly wages for men.[1] With the height interaction on hourly wages being statistically significant (5% level) and negative, the effect for women reduces to about two thirds of the effect for men. With respect to monthly wages, there is no statistically difference in the height effect for men and women. In magnitude, the height coefficient for hourly wages for men is almost identical to the height coefficient reported by Case and Paxson (2008) for the NCDS sample, but slightly lower for women.

[1] As a robustness check I also tried a second order polynomial of height which did not turn out to be statistically significant.

Table II: Ordered logit results for educational outcomes

| | (1) *pooled* | (2) *women* | (3) *men* | (4) *pooled* | (5) *pooled* |
|---|---|---|---|---|---|
| educ | | | | | |
| *height* | 1.0318*** | 1.0247*** | 1.0319*** | 1.0269 | 1.0311*** |
| | (0.0026) | (0.0029) | (0.0027) | (0.0179) | (0.0120) |
| $height \times female$ | 0.9932* | | | 1.0251 | 1.0103 |
| | (0.0038) | | | (0.0221) | (0.0175) |
| *year dummies* | Yes | Yes | Yes | Yes | Yes |
| *region dummies* | Yes | Yes | Yes | Yes | Yes |
| *sibling FE* | – | – | – | Yes | – |
| *pseudo* $R^2$ | 0.0494 | 0.0442 | 0.0537 | 0.1851 | 0.0668 |
| *Observations* | 24,371 | 12,704 | 11,667 | 1,270 | 1,270 |

*Notes:* */**/*** indicates statistical significance at the 10/5/1% level. Coefficients in odds ratios. All regressions include an intercept as well as controls for age and and indicator for living in the GDR before 1989. Regressions without sibling FE additionally control for mother, father Abitur and migration background. The estimation sample is restricted to first observations of each individual.

Column (2) depicts a lower effect for men, while the height interaction turns insignificant.[2] Proxies for cognitive ability reduce the crude height coefficient by about 39% (hourly wages) and 27% (monthly wages), respectively. Big five personality traits as measure for noncognitive ability have only little effect on the height coefficient, but decrease its significance level. Health satisfaction seems to play no role.[3]

Columns (6) - (9) present the results for the models with sibling fixed effects. Including sibling dummies into the regression reduces the crude height effect by the factors of three (hourly wages) to four (monthly wages), while turning the height effect insignificant. The impact of the inclusion of further controls for (non)cognitive ability and health satisfaction on the already insignificant height coefficient is very low. This might be due to the high intra sibling correlation in education and the fact that the sibling fixed effect also captures parental education, common to all siblings.[4]

# 4. Discussion

Shared family background (including genetics) explains the labor market height premium among young German workers, which contradicts previous studies using SOEP data (Heineck, 2005; Rietveld et al., 2015). However, this does not contradict the claim that the largest share of the height premium can be explained by the association between height and cognitive function through childhood living conditions since the effects of education are partly absorbed by the sibling fixed effect. International comparisons have shown that among OECD countries, the impact of family background on student performance is particularly high in Germany (Schütz et al., 2008).

[2]Respondents in the sibling sample are on average younger. Lindqvist (2012) indicates that for job-status the height effect is increasing with age. This can be also shown for wages (results available upon request). I interpret this as cumulated height gains in terms of wages.

[3]Health satisfaction may only partly captures fitness and strength. To test this, I conducted a robustness analysis using the waves in which grip strength is included in the SOEP (2006, 2008, 2010, 2012 and 2014 ). Results show no confounding effect of grip strength on the height coefficient.

[4]The correlation of the minimum and the maximum in years of education of each sibling group is about 0.67.

Table III: Wage regression results for male and female workers

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| *Hourly wages:* | | | | | | | | | |
| *height* | 0.0092*** | 0.0065*** | 0.0044** | 0.0042** | 0.0043** | 0.0027 | 0.0023 | 0.0023 | 0.0024 |
| | (0.0007) | (0.0020) | (0.0019) | (0.0019) | (0.0019) | (0.0021) | (0.0021) | (0.0020) | (0.0020) |
| *height* × *female* | –0.0026** | –0.0001 | –0.0007 | –0.0006 | –0.0006 | 0.0048 | 0.0038 | 0.0037 | 0.0037 |
| | (0.0011) | (0.0031) | (0.0028) | (0.0028) | (0.0028) | (0.0030) | (0.0029) | (0.0029) | (0.0029) |
| *adj.* $R^2$ | 0.1806 | 0.2459 | 0.3184 | 0.3225 | 0.3240 | 0.5154 | 0.5314 | 0.5346 | 0.5350 |
| *Monthly wages:* | | | | | | | | | |
| *height* | 0.0112*** | 0.0093*** | 0.0072*** | 0.0070*** | 0.0071*** | 0.0031 | 0.0023 | 0.0021 | 0.0022 |
| | (0.0009) | (0.0026) | (0.0024) | (0.0025) | (0.0025) | (0.0028) | (0.0028) | (0.0027) | (0.0027) |
| *height* × *female* | –0.0023 | –0.0020 | –0.0036 | –0.0038 | –0.0038 | 0.0014 | 0.0006 | 0.0004 | 0.0002 |
| | (0.0016) | (0.0047) | (0.0045) | (0.0045) | (0.0044) | (0.0050) | (0.0049) | (0.0048) | (0.0048) |
| *adj.* $R^2$ | 0.2417 | 0.2360 | 0.3139 | 0.3271 | 0.3280 | 0.5121 | 0.5496 | 0.5548 | 0.5549 |
| *year dummies* | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| *region dummies* | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| *cognitive ability* | – | – | Yes | Yes | Yes | – | Yes | Yes | Yes |
| *noncognitive ability* | – | – | – | Yes | Yes | – | – | Yes | Yes |
| *health* | – | – | – | – | Yes | – | – | – | Yes |
| *sibling FE* | – | – | – | – | – | Yes | Yes | Yes | Yes |
| *Observations* | 117,928 | 11,113 | 11,113 | 11,113 | 11,113 | 11,113 | 11,113 | 11,113 | 11,113 |

*Notes:* Cluster robust standard errors in parentheses. */**/*** indicates statistical significance at the 10/5/1% level. All regressions include an intercept and baseline controls. Since migration background and parental education are time-invariant, these are not included in the estimations with family fixed effects.

Because my results do not reinforce the findings of a significant height premium remaining even after controlling for the sibling fixed effect (Lundborg et al., 2014), the conclusions of this study are quite different: My estimates leave no room for psychological explanations of the height-wage association such as preferential treatment of tall workers.

It remains unclear and an avenue for further research, why Hausman-Taylor based estimates using similar waves of the same dataset show a significant labor market height premium.

# References


Baetschmann, G., K. E. Staub and R. Winkelmann (2015) "Consistent estimation of the fixed effects ordered logit model" *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **178**, 685–703.

Case, A. and C. Paxson (2008) "Stature and status: Height, ability, and labor market outcomes" *Journal of Political Economy* **116**, 499–532.

Case, A., C. Paxson and M. Islam (2009) "Making sense of the labor market height premium: Evidence from the British Household Panel Survey" *Economics Letters* **102**, 174–176.

Heineck, G. (2005) "Up in the skies? The relationship between body height and earnings in Germany" *Labour. Review of Labour Economics and Industrial Relations* **19**.

Lindqvist, E. (2012) "Height and leadership" *The Review of Economics and Statistics* **94**, 1191–1196.

Lundborg, P., P. Nystedt and D.-O. Rooth (2014) "Height and earnings: The role of cognitive and noncognitive skills" *The Journal of Human Resources* **49**, 141–166.

Persico, N., A. Postlewaite and D. Silverman (2004) "The effect of adolescent experience on labor market outcomes: The case of height" *Journal of Political Economy* **112**, 1019–1053.

Rietveld, C. A., J. Hessels and P. van der Zwan (2015) "The stature of the self-employed and its relation with earnings and satisfaction" *Economics and Human Biology* **17**, 59–74.

Schütz, G., H. W. Ursprung and L. Wößmann (2008) "Education policy and equality of opportunity" *Kyklos* **61**, 279–308.

# Appendix

Table AI: Definition of variables

| Variable | Definition |
|---|---|
| *outcomes:* | |
| real hourly wages | logarithm of real monthly labor income divided by hours worked per month; nominal labor income deflated by the consumer price index included in the SOEP |
| real monthly earnings | logarithm of real monthly labor income; nominal labor income deflated by the consumer price index included in the SOEP |
| *baseline controls:* | |
| height | mean of self-reported height for each respondent in cm; reported on biennial frequency since 2002 |
| female | 1=female; 0=male |
| age | in years |
| GDR | 1= place of residence before reunification: GDR; 0=place of residence before reunification: FRG |
| migration background[a] | 1= non German or father or mother non German; 0=German and father and mother German |
| disabled | 1=disability status; 0=no disability status |
| year dummies | 2001-2014 |
| region dummies | 16 dummies for each German region ("Bundesländer") |
| *cognitive skills:* | |
| ISCED 1 | 1= highest educational attainment level: general elementary; 0=other |
| ISCED 2 | 1= highest educational attainment level: middle vocational; 0=other |
| ISCED 3 | 1= highest educational attainment level: vocational or Abitur; 0=other |
| ISCED 4 | 1= highest educational attainment level: higher vocational; 0=other |
| ISCED 5 | 1= highest educational attainment level: higher education; 0=other |
| mother Abitur[a] | 1=mother's highest educational attainment level: Abitur or higher; 0= mother's highest educational attainment level: below Abitur |
| father Abitur[a] | 1=father's highest educational attainment level: Abitur or higher; 0= father's highest educational attainment level: below Abitur |
| *noncognitive skills (big 5):* | |
| openness | standardized mean (mean=0, std= 1) of openness to experience score for each respondent; score included in the waves of the years 2005, 2009, 2013 |
| conscientiousness | standardized mean (mean=0, std= 1) of conscientiousness score for each respondent; score included in the waves of the years 2005, 2009, 2013 |
| extroversion | standardized mean (mean=0, std= 1) of extroversion score for each respondent; score included in the waves of the years 2005, 2009, 2013 |
| agreeableness | standardized mean (mean=0, std= 1) of agreeableness score for each respondent; score included in the waves of the years 2005, 2009, 2013 |
| neuroticism | standardized mean (mean=0, std= 1) of neuroticism score for each respondent; score included in the waves of the years 2005, 2009, 2013 |
| *health:* | |
| health satisfaction | subjective evaluation of satisfaction with current level of health; 11 categories: 0=low, . . . , 10=high; included as dummy variables |

*Note:* [a] marks variables that do not vary within sibling-pairs. Since the effect of these variables is absorbed by the sibling fixed effect, theses variables are not included in the fixed effects estimations