## Supplementary Appendix A

**1. Overview of recent micro-level research on the effect of religion on economic variables**

**2. Data collection and sampling**

**3. Description of the Heckman (1979) two-step procedure to check for possible selectivity biases arising from unobserved variables and self-selection into religiosity**

**1. Overview of recent micro-level research on the effect of religion on economic variables**

| **Study** | **Synopsis of results with respect to effects on religion on economic variables** |
|---|---|
| Guiso et al. (2003) | Using data from 66 countries, the authors find that religious upbringing, religiosity and religious practice promote economically conducive attitudes. |
| Cuesta (2004) | Religious affiliation does not have an effect on basis needs satisfaction in Nicaragua. |
| Steen (2004) | Focusing on men in the USA, ‘the paper finds evidence that both men raised as Catholics and men raised as Jews have higher earnings.’ |
| Sakwa (2006) | Among Catholic university students in Nairobi, Kenia, religious attitudes towards poverty correlate with specific poverty alleviation objectives. |
| Arano and Blair (2008) | There is a ‘bicausal relationship between religion and income’ in Mississippi, USA. |
| Chiswick and Huang (2008) | Among Jewish men in the USA, ‘religious involvement is associated with more favorable labor market outcomes,’ but ‘beyond some point religious practice has a negative effect.’ |
| Bettendorf and Dijkgraaf (2010) | Using data from 25 countries, ‘church membership is found to have a positive effect on income for high-income countries,’ while ‘this effect is negative for low-income countries.’ |
| Bettendorf and Dijkgraaf (2011) | Religious attendance does not have an effect on household income in the Netherlands. |
| Permani (2011) | There is a positive effect of religious social capital on earnings in Indonesia. |
| Cornelissen and Jirjahn (2012) | In Germany, ‘being raised by two religious parents, but having no current religious affiliation is associated with higher earnings.’ |
| Audretsch et al. (2013) | ‘Religions like Islam and Jainism are more favorable for self-employment,’ while ‘Hindus are less likely to be self-employed.’ |

## 2. Data collection and sampling

The data collection process followed a two-phase design, a qualitative preparatory study followed by a quantitative household survey. We chose the two-phase approach in order to be able to contextualise the key analytical concepts and to pre-test the survey instrument. In the first phase, semi-structured interviews and focus group workshops were conducted in various villages. We gathered information on religious communities and developed contextually relevant categorizations with representatives of the local population. Moreover, we collected data on income sources such as informal income-generating activities and agricultural production patterns such as livestock-breeding, small-scale horticultural activities and the cultivation of field crops (cf. survey questionnaire in supplementary appendix B). The data was used to develop the survey instrument and later to perform consistency checks on the quantitative data.

In the second phase, 221 households were sampled from the universe (all households in Fetakgomo Municipality) in a two-stage cluster sampling process. We used a geographical approach. As the area is entirely rural, in nearly all instances one household inhabits one dwelling. The primary sampling units (clusters) were formed on the basis of the subplace delimitations by StatsSA. Thirty of 61 clusters were randomly selected with equal probability of selection. The size of the clusters varied between 25 and 2066 households.

The secondary sampling units are the households.[1] The sampling frame was recent Google Maps (2011) satellite imagery, in which all dwellings were clearly visible. Cluster delimitations were plotted on the aerial map, facilitating an allocation of households/ dwellings to clusters. In each cluster, the households to be interviewed were selected through

[1] In order to ensure compatibility with data from official statistics, we used StatsSA's (2010) definitions of a household as 'a group of persons who live together and provide themselves jointly with food and/or other essentials for living, or a single person who lives alone' and of a household member as 'a person that resides with the household for at least four nights a week.' The household head was operationalised as the household member who bears the responsibility in the household.

fixed rate sampling (1 in 55), ensuring that each household in the universe had the same probability of selection given the differing size of the clusters. The satellite image proved to be an accurate frame. In the rare case of inaccuracies encountered, the frame was adjusted accordingly by adding those household or removing them, respectively. If the household head was absent, at least four re-visits were done at different hours and at least two different days. Of the 221 sampled households, 14 either refused to participate or were repeatedly unavailable. Interviews were conducted in 207 cases, yielding a response rate of 93.7%. Of these, due to missing values and the removal of outliers, 180 are used in the empirical part. All interviews were conducted with the household head in Northern Sotho, using the questionnaire presented in supplementary appendix B.

**3. Description of the Heckman (1979) two-step procedure to check for possible selectivity biases arising from unobserved variables and self-selection into religiosity**

In the first step, we run two probit-models, on the household head's probability of being member of the ZCC and on the probability of practicing ATR (as those are the religion categories with significant coefficients in the estimation of (2), see section 5 of the main article). The selection of the household into religion category *r* (*zcc* or *atr*, respectively) is modelled as follows.

$$r_i^* = \mathbf{z}_i\gamma + u_i \tag{A.1}$$

where $r_i^*$ is a latent variable and religion category $r_i = 1$ if $r_i^* > 0$ and $r_i = 0$ otherwise. $\mathbf{z}_i$ is the respective vector of the variables explaining the decision to actively practice *zcc* or *atr*. Symbol $\boldsymbol{\gamma}$ is a vector of the respective coefficients in the probit model and $u_i$ is the error term, assumed to be normally distributed. This probit model needs to contain "at least one nontrivial determinant" of $r_i$, that is, a variable uncorrelated with household income except through its correlation with the religion category (Cameron & Trivedi, 2005:870). In section 5 of the main article, we identify such variables from the probit model (A.1) in combination with the results of equation (2) and provide a justification why we consider this exclusion restriction to hold.

We compute the inverse Mill's ratio (Heckman's λ) from the generalised residual of the probit estimates. The inverse Mill's ratio is given as

$$\lambda_1 = \frac{\varphi(z_i\boldsymbol{\gamma})}{\Phi(z_i\boldsymbol{\gamma})} \quad \text{and} \quad \lambda_0 = \frac{\varphi(z_i\boldsymbol{\gamma})}{1-\Phi(z_i\boldsymbol{\gamma})} \tag{A.2}$$

for the $r$=1 and the $r$=0 cases, respectively. $\varphi(z_i\boldsymbol{\gamma})$ is the probability density function of the standard normal distribution and $\Phi(z_i\boldsymbol{\gamma})$ the cumulative distribution function.

In the second step, Heckman's $\lambda$ is included as an additional regressor in the income equation. It is interacted with the dummy of the religion category $r$:

$$\ln(income_i) = \boldsymbol{x}_i\boldsymbol{\beta} + \beta_r r_i + \beta_a \lambda_{1i} r_i + \beta_b \lambda_{0i}\left(1 - r_i\right), \qquad \text{(A.3)}$$

where $\mathbf{x}_i$ summarises the regressors, $\boldsymbol{\beta}$ is the corresponding vector of coefficients, $r_i$ denotes the religiosity dummy variable with $\beta_r$ its coefficient. The terms $\beta_a\lambda_{1i}r_i$ and $\beta_b\lambda_{0i}(1–r_i)$ switch on and off depending on whether a household is in the $r_i$=1 category or not. The coefficient $\beta_r$ is the ‘true’ effect of the dummy variable. The coefficients of Heckman's $\lambda$, $\beta_a$ and $\beta_b$ are the estimated covariances between the unobserved variables in the error term of the probit estimate and the unobserved variables in the error term of the income equation (1). If there is no significant correlation between the error terms, we can rule out selection bias from unobserved variables (Cameron & Trivedi, 2005). To statistically test for selection bias we perform a t-test on the coefficients of $\lambda_1$ and $\lambda_0$ (cf. Vella & Verbeek, 1999).

**References**


Arano, KG, & Blair, BF, 2008. Modeling Religious Behavior and Economic Outcome. Is the Relationship Bicausal? Evidence From a Survey of Mississippi Households. Journal of Socio-Economics 37, 2043–53.

Audretsch, DB, Bönte, W, & Tamvada, JP, 2013. Religion, social class, and entrepreneurial choice. Journal of Business Venturing 28, 774–89.

Bettendorf, L, & Dijkgraaf, E, 2010. Religion and Income: Heterogeneity Between Countries. Journal of Economic Behavior and Organization 74, 12–29.

Bettendorf, L, & Dijkgraaf, E, 2011. The Bicausal Relation between Religion and Income. Applied Economics 43, 1351–63.

Cameron, CA, & Trivedi PK, 2005. Microeconometrics. Methods and applications. Cambridge University Press, Cambridge.

Chiswick, BR, & Huang, J, 2008. The Earnings of American Jewish Men. Human Capital, Denomination, and Religiosity. Journal for the Scientific Study of Religion 47, 694–709.

Cornelissen, T, & Jirjahn, U, 2012. Religion and earnings. Is it good to be an atheist with religious parental background? Economics Letters 117, 905–8.

Cuesta, J, 2004. From Economicist to Culturalist Development Theories. How Strong is the Relation between Cultural Aspects and Economic Development? The European Journal of Development Research 16, 868–91.

Google Maps, 2011. maps.google.com. Accessed 6 June 2011.

Guiso, L, Sapienza, P, & Zingales, L, 2003. People’s Opium? Religion and Economic Attitudes. Journal of Monetary Economics 50, 225–82.

Heckman, JJ, 1979. Sample selection bias as a specification error. Econometrica 47, 153–61.

Lipford, JW, & Tollison, RD, 2003. Religious participation and income. Journal of Economic Behavior & Organization 51, 249–60.

Permani, R, 2011. The presence of religious organisations, religious attendance and earnings. Evidence from Indonesia. The Journal of Socio-Economics, 40 (3), 247–258.

Sakwa, MM, 2006. Bible and Poverty in Kenia. An Empirical Exploration. Dissertation: Radboud Universiteit.

StatsSA (Statistics South Africa), 2010. Concepts and definitions for Statistics South Africa v3. Pretoria: Statistics South Africa.

Steen, TP, 2004. The relationship between religion and earnings. Recent evidence from the NLS Youth Cohort. International Journal of Social Economics 31, 572–81.

Vella, F, & Verbeek, M, 1999. Estimating and interpreting models with endogenous treatment effects. Journal of Business & Economic Statistics 17, 473–78.