I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12665

Is Consanguinity an Impediment to Child Development Outcomes?

Cem Mete
Laurent Bossavie
John Giles
Harold Alderman

OCTOBER 2019

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12665

Is Consanguinity an Impediment to Child Development Outcomes?

Cem Mete
World Bank

Laurent Bossavie
World Bank

John Giles
World Bank and IZA

Harold Alderman
International Food Policy Research Institute

OCTOBER 2019

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Is Consanguinity an Impediment to Child Development Outcomes?*

Marriages between blood relatives – also known as consanguineous unions – are widespread in North Africa, Central and West Asia and most parts of South Asia. Researchers have suggested that consanguinity has adverse effects on child development, but assessing its impact is not straightforward as the decision to marry a relative might be endogenous to other socio-economic factors. Using a unique dataset collected in rural Pakistan, this paper assesses the extent to which consanguinity is linked to child cognitive ability and nutritional status. As economic benefits of marrying cousins may lead to upward bias to estimates of the effects of consanguinity on child outcomes, prior work likely underestimates the negative impacts of consanguinity on child outcomes. After controlling for current household wealth and parent education, this paper exploits (current and past) grandfather land ownership and maternal grandparent mortality to identify the effect of endogenous consanguinity on child cognitive ability and height-for-age. Children born into consanguineous unions have lower cognitive scores, lower height-for-age, and a higher likelihood of being severely stunted. More importantly, adverse effects are significantly larger after accounting for the endogeneity of consanguineous unions, suggesting that negative impacts on child development are substantial, and likely to be larger than suggested in previous studies. Reducing incentives for consanguineous unions should therefore be of concern among policy makers aiming at improving child development outcomes where marrying cousins is common.

JEL Classification: I15, J12

Keywords: consanguineous marriages, cognitive tests, malnutrition, household decision making, child development

Corresponding author:
John Giles
The World Bank
1818 H Street NW
Washington, DC 20433
USA

E-mail: jgiles@worldbank.org

* The findings, interpretations and conclusions expressed herein are those of the authors, and do not necessarily reflect the views of the International Bank for Reconstruction and Development/the World Bank or the governments they represent.

1. Introduction

Improving the well-being of a population and raising child development outcomes forces policymakers to confront many interrelated constraints and shortcomings simultaneously. The Sustainable Development Goals (SDGs) provide policymakers' in the developing world with a set of objective benchmarks that suggest specific areas for policy reform and investment. For example, addressing the goal of reducing under-five mortality rates lead to an emphasis on increasing the availability of health care centers and encouraging regular preventive health checkups and other healthy behaviors (WHO and UNICEF, 2014). Although strong economic growth and improving institutional environments are at the core of achieving SDGs, these policies work best when designed in accord with a strong understanding of household values and knowledge.

Thus, after first addressing issues related to health care access, further progress on improving child development outcomes may require incentivizing changes to long-established social and cultural practices that are important impediments to achieving child development outcomes. As creating incentives to change any cultural norm or practice may be both costly and controversial, empirical assessments of potential benefits are important. In this paper, we examine the effect of consanguineous marriage, marriage to first or second cousins, to child cognitive ability and height-for-age.

Consanguinity is widespread in North Africa, Central and West Asia, as well as in most parts of South Asia (Kaiser, 2016). Further, consanguineous unions are more common in Islamic societies: in Afghanistan the proportion of consanguineous marriages is estimated to be 46.2 percent (Saify and Saadat, 2012), in Lebanon it is 35.5 percent (Barbour and Salameh 2009), between 20.9 to 32.8 percent in Egypt, 47 to 60 percent in Iraq, 42.1 to 66.7 percent in Saudi Arabia and 40 to 44.7 percent in Yemen (Tadmouri et al. 2009). Outside the Muslim world, consanguinity also prevails in some societies: between

20 to 45 percent of marriages in the Hindu-majority states of South India are contracted between close relatives, and those mostly occur between uncle and niece (Bittles, 1994).

Existing evidence shows that inbreeding increases the risk of neonatal and post-neonatal mortality due to the expression of detrimental recessive genes (Saggar and Bittles 2008; Dorsten, Hotchkiss and King 1999), but the linkages between consanguinity and a broader array of child development outcomes are yet to be studied in depth. Such research on consanguinity tends to suffer from limited sample sizes, few outcomes of interest, or both. Further, to our knowledge no research to date controls for non-random selection into consanguineous unions.

As consanguinity may be linked to a wide range of child development indicators, we use a unified framework to lay out the decision to enter consanguineous marriages and potential effects on child development. We then explore the relationship between consanguinity and two important child development outcomes: child cognitive ability and stature. The paper makes use of a unique household survey from Pakistan that includes information on the marriage patterns of all household members as well as their parents (regardless of whether they are still alive or present in the household). Child development outcomes include height and weight measurements for all household members age 5 and above, cognitive tests administered to children, and a set of questions useful for modelling the marriage decisions of the parents of household members, including grandparent land ownership status and mortality.

The next section briefly reviews motivations for marriage to a relative, most frequently a first or second cousin, and existing evidence on human development impacts. Section 3 presents a modified version of Becker's marriage model that is used to motivate the empirical estimation discussed in Section 4. Data sources are described by Section 5, followed by a discussion of empirical results and conclusions.

2. Consanguineous Marriages

2.1 Why Marry a Relative?

High levels of consanguinity in Pakistan have been well documented. Using the 1990/91 Pakistan Demographic and Health Survey, Hussain and Bittles (1998) estimate that 60 percent of marriages are consanguineous unions, and that the incidence of consanguinity had remained unchanged over the previous three to four decades.[1] The decision to marry a relative may be driven by cultural or economic factors or both.

In some societies there is a belief that compatibility between husband and wife, as well as between the bride and the rest of the household, will be maximized by marriages among kin (Khlat et al. 1986). In addition, in societies in which violence against women is rampant, concerns over the safety and welfare of females may influence the decision to enter a consanguineous marriage.[2] Some contributions in ethnography and sociology have also documented that marrying a cousin can provide higher autonomy and status to women compared to those who are not related to their spouse (Weinreb, 2008). Further, beliefs about the potential negative effects on child outcomes of marriages among kin would likely influence the prevalence of such partnerships. Many individuals may be unaware of potential negative effects of marriages among kin or not fully believe scientific evidence.

Economic explanations for consanguinity range from incentives within agricultural economies to reduce risk in the absence of savings and insurance markets, or alternatively to family wealth maximization strategies. In agricultural societies, parents may prefer to keep productive and responsible

[1]This is consistent with the trends observed in other developing countries: Givens and Hirschman (1994), for example, find that there was a modest increase in marriages between cousins in Iran from 1940s to 1970s. Contrary to other traditional marriage practices that have declined, most notably early marriage (Mensch et al. 2005), consanguinity has remained resilient over time.

[2] To the best of our knowledge, this hypothesis remains untested.

adults in the family, rather than taking a risk on integrating an outsider into the household. More generally, when informal contracting for productive activities is the norm, "trust" is essential but scarce, and thus increases the benefits of consanguinity as it reinforces within-family partnerships (Kuper 2009).

A second set of economic motivations for consanguineous marriage stems from the lack of well-developed savings and insurance markets, which influences both considerations of security in old age as well as the ability to cope with idiosyncratic risks. Support in old age is likely to be higher when the next generation couple, both the husband and the wife, have some relationship to an elderly parent.[3] In Bangladesh, the effect of this expectation leads to lower dowries for consanguineous marriages (Do, Iyer and Joshi, 2013). Caldwell, Reddy, and Caldwell (1983) and Bittles (1994) have also suggested that households marry within the family in South Asia to avoid large dowry payments at the time of marriage. With respect to insurance, Mobarak, Kuhn and Peters (2013) utilize the introduction of a flood protection embankment on one side of a river to test how consanguinity responds to the reduction of flood risk. Consistent with the hypothesis that consanguinity is a response to uninsured risk, flood-protected households are less likely to enter into consanguineous marriages and consequently, when poor households face less flood risk, they may be willing to use more savings for dowry payments to non-relatives at the time of marriage.

Third, and perhaps most important among economic explanations, maintaining family property is often a key motivation for marriages among kin (e.g., Caldwell, Reddy and Caldwell 1988). As land is not portable and division of plots may lead to lower productivity, keeping land within the family may be an important motivation in rural areas. Wealth is more than land ownership and may thus have two effects operating in opposite directions; landownership in rural areas increases the likelihood of consanguineous marriage while wealth overall can decrease the likelihood of marrying a relative. Saedi-Wong, Al-Frayh

[3]See Holy (1989) for a discussion of consanguinity and elderly support in the Middle East.

and Wong (1989) find that, in Saudi Arabia there are higher rates of marriage to a close relative in rural areas and among the poor. In addition, among the Reddis of Chittoor District in South India, Rami Reddy and Chandrasekhar Reddy (1979) show that marriage to a close family member is higher among landowning families.

2.2 What are the "Impacts" of Consanguineous Marriages on Children?

Marrying a biological relative increases the likelihood that offspring will receive two copies of a deleterious gene from parents, with potentially detrimental consequences for children's outcomes. Most research to date on the "impact" of consanguineous marriages on children has focused on child mortality. In Pakistan, without correcting for socioeconomic characteristics, Bittles (1994) finds that as a percentage of all reported pregnancies, total pre- and post-natal mortality rose from 16.4 percent in non-consanguineous progeny, to 20.1 percent in second cousins, 22.1 percent in first cousins and 39 percent in double first cousins. Infant mortality was 5.1 percent among non-consanguineous progeny, 6.9 percent for second cousins, 7.9 percent for first cousins and 12.7 percent for double first cousins. Similarly, Shah, Toney and Pitcher (1998) also document a correlation between first-cousin marriages and child mortality in Pakistan. Other examples include Farah and Preston (1982), who find that 37 percent of women in a sample from Sudan are married to cousins and child mortality among this group is 20 percent higher than among families in which the husband and wife are not related by blood (more distant relatives have intermediate child mortality). The magnitude of the effect is quite large; requiring about six additional years of woman's education to offset the child-survival consequences of marrying a cousin. For Egypt, Shawky et al. (2013) also find higher rates of genetic diseases and higher prenatal, neonatal and child morbidity and mortality among individuals born from consanguineous unions.

Apart from child mortality, consanguinity may also be associated with malnutrition, but there are relatively few studies examining this connection. In one notable exception, Hasnain and Hashmi (2009) survey eight hundred children in rural Sindh, Pakistan and showed that consanguinity (treated as an exogenous variable) is a key predictor of being underweight. Further, descriptive evidence suggests that the offspring of unrelated parents may perform better in cognitive tests than the children of first-cousin marriages. Using a representative sample of 3,203 (grade 4 and grade 6) children from the Arab educational system in Israel, Bashi (1977) finds that the offspring of double-cousin marriages perform the worst on cognitive tests, followed by the offspring of first-cousin marriages and then the offspring of unrelated parents Another contribution from the medical literature by Morton (1978) also finds negative effects of consanguinity on children's cognition. Similarly, Kanaan et al. (2008) report higher rates of mental retardation among children whose parents are first cousins in Lebanon. In Kuwait, Al-Kandari and Crews (2011) show that rates of cognitive disability are significantly higher among children from consanguineous unions. Finally, a recent contribution by Lakhan et al. (2017) finds a positive association between consanguinity and family history of intellectual disability in tribal and nontribal populations from India.

Regarding birth defects, Magnus et al. (1985), Mehrabi and Zeyghami (2005), Temtamy and Aglam (2012) report increased congenital malformations among children from consanguineous unions in Norway, Iran and Egypt, respectively. A review of the medical literature by Jaber et al. (1998) reports that the rate of congenital malformation among the offspring of related parents is about 2.5 times that of the offspring of unrelated parents in many country settings. Zlotogora (2002) findsthat estimates of excess birth defects in first cousin progeny have ranged from 0.7 % to 7.5 %. In Morocco, Jaouad et al. (2009) report that autosomal recessive disorders are strongly associated with consanguinity.

With respect to observed height, which is a proxy for nutritional status in early childhood, and child cognitive ability, there is good reason to believe that the estimated magnitude of any negative effects of consanguineous marriage will be biased toward zero. Economic motives driving consanguineous marriages, particularly improved ability to manage idiosyncratic risk, and increased productivity in agriculture and home production, would likely lead to more inputs available for child nutritional support in households with consanguineous unions between parents than in those in which parents are unrelated.

Although the vast majority of contributions in the literature report a negative association between consanguinity and health and cognitive outcomes, some contributions suggested that consanguinity can also increase the concentration of positive traits, which may have positive implications for other development outcomes not considered here. Denic and Nicholls (2007) for example show that consanguinity can be protective against malaria and potentially other conditions that, in turn, have effects on childhood development/accumulation of human capital.

An important limitation of previous studies on consanguinity, both in social sciences and in the medical literature, is that they do not account for the endogenous decision to marry a relative A main contribution of this paper lies in identifying the consequences of consanguinity separately from biases associated with selection into consanguineous unions. These may be net negative biological effects that are masked by positive economic benefits within consanguineous marriages, or effects due to differences in investment in child human capital with expectations that children, like their parents, may marry first or second cousins.

3. Theoretical Framework

To motivate the empirical analysis that follows, we build on the theoretical models proposed by Becker and Tomes (1976) and Becker (1981). In this setup, spouse characteristics influence the expected utility derived from marriage. Marriage is assumed to be always preferred to remaining single. A household's utility derived from marriage is a function of the quality of offspring Q, and spouse's wealth W expressed as:[4]

$$U = f\,(W, Q) \tag{1}$$

The utility derived from marriage is increasing in both wealth and child quality:

$\partial U(W,Q)/\partial W > 0$, and $\partial U(W,Q)/\partial Q > 0$

Marital Wealth Technology

Spousal assets are assumed to be of two types: joint land wealth, w_l, of the husband and wife, which is assumed to be illiquid (with relatively few observed sale-purchase transactions of land in rural South Asia) and non-land wealth w_o, assumed to be perfectly liquid. Household wealth is a function of land wealth, other wealth and the degree of consanguinity between spouses, δ, or

$$\mathrm{W} = \mathrm{g}(w_l, w_o, \delta) \tag{2}$$

Marital wealth is increasing in both land wealth and non-land wealth ($\partial W(w_l, w_o, \delta)/\partial\, w_l > 0$ and $\partial W(w_l, w_o, \delta)/\partial\, w_o > 0$), and the illiquidity of land implies that land wealth and consanguinity are complements, or $\partial^2 W(w_l, w_o, \delta)/\partial w_l \partial\, \delta > 0$. The assumed complementarity between land wealth and consanguinity, reflected in the positive cross partial derivative, is mainly derived from Do et al. (2013), who argue that short social distance plays the role of social capital in marriage contracting, by making ex-ante commitment between families easier. Similarly, Goody (1973), Agarwal (1994) and Bittles (2001) have also argued consanguinity provides a means to consolidate and maintain family assets and resources, reinforcing incentive-based motives aimed at resisting shirking.

[4]For the sake of simplicity, and contrary to Becker and Tomes (1976), we abstract from decisions on the number of children, and on the interaction between quantity and quality of children.

The model of Do et al. (2013), which is applied to the South Asian context, postulates that the commitment problem is on the bride's side due to the patrilocal exogamy of marriages. They also make the argument that ex-ante commitments made as part of marriages are more easily enforced as informal contracts, and easier to enforce within the extended family. They argue that in a consanguineous or socially close union, the bonds of trust are likely to be stronger, and interests of the bride and groom's families are less likely to diverge. Similarly, Putnam (2000) argues that close relatives have more (verifiable) information about each other, are more likely to exert effort in economic activities, and are less likely to engage in opportunistic behaviors. They are likely to show higher levels of trust, cooperation, and altruism to both their natal and marital families. A related mechanism that can make consanguineous unions more valuable for families with more land wealth is simply property retention within a family (Bittles 1994; Sandridge et al. 2010). According to these two possible mechanisms, consanguineous marriages will be more beneficial to families that are landowners.

The Child-Quality Production Function

Following Becker and Tomes (1976), offspring quality, Q, is assumed to be a function of household investment in children *i*, and such factors as inherited ability, public expenditure on children, "luck" and other unobservables that affect quality. Taken together, the component of offspring quality outside of parental control is captured by *e,* denoting offspring endowment, in the sense of Becker and Tomes (1976). Consanguinity between spouses, δ, may have a detrimental effect on offspring quality because of a higher degree of homozygosity between parents. In addition, the model includes heterogeneity in beliefs about the detrimental effects of consanguinity, ϑ. The child quality production function can be written in general form as:

$$Q = h(i, \delta, \vartheta, e) \tag{3}$$

Beliefs about the detrimental effects of consanguinity on child quality (child development outcomes), $\vartheta = v(s)$, are a function of schooling, s, and these beliefs are increasing in schooling, $d\vartheta/ds > 0$.[5] The reason for assuming a negative linear relationship between parental level of schooling and consanguineous marriage is mainly twofold. First, one can suspect that individuals with higher levels of schooling have access to a larger and more diverse social network, and hence to a larger pool of potential spouses that includes non-relatives. Second, more educated individuals are more likely to be aware of the detrimental effects of consanguinity for children.[6] Consanguinity is assumed to have potential detrimental effects on the physical and mental abilities of children, or $\partial Q(i,\delta)/\partial\delta<0$, and offspring quality is increased through parental investments, $\partial Q(i,\delta)/\partial i>0$.

Consanguinity in the Household Optimization Decision

At the time of marriage, families maximize their expected utility from marriage with respect to δ in (4) subject to a household budget constraint (5):

$$\max_{\delta} U = f\left(W(w_l, w_o, \delta), Q(i, \delta, \vartheta, e)\right) \tag{4}$$

s.t.

$$i \leq w_o \tag{5}$$

[5]A large fraction of Pakistani households might not be aware of the potential detrimental effects of consanguinity. In the Pakistan Labor and Skills Survey 2013, individuals are asked to agree or disagree with the statement that "consanguinity between spouses can negatively affect children's health or abilities." 72.4% of individuals report disagreeing with this statement.

[6]This positive relationship between awareness about the determinantal effects of consanguinity and levels of schooling is supported by the Pakistan Labor Skill Survey data used by this study. We find a very strong and positive association between level of education and awareness about the detrimental effects of consanguinity in the sample. We also find that more educated individuals are less likely to engage in consanguineous unions, holding other observable characteristics fixed. Levels of schooling of both husbands and wives are strongly negatively associated with marrying a relative in the data, although the strength of the association is larger for females in our sample from rural Pakistan.

For the sake of simplicity, we abstract from parental optimization between own consumption and investment in children by assuming that all the liquid wealth w_o is invested in child quality. Applying the Chain rule, families choose the optimal degree of consanguinity, δ^*, where the marginal benefits of consanguinity in terms of wealth production equal the marginal costs associated with the detrimental effects on child quality, or

$$\partial U/\partial W \cdot \partial W/\partial \delta = \partial U/\partial Q \cdot \partial Q/\partial \delta \tag{6}$$

where $\partial^2/\partial \delta \partial\, w_l > 0$.

In this context, when choosing δ^* families optimize by finding the balance between the *wealth effect* of consanguinity and negative effects associated with kinship *(*a *child quality effect*)*.*[7] Since illiquid land wealth increases the marginal benefit to consanguinity without affecting its marginal costs, it follows that:

$$d\delta^*/dw_l > 0 \tag{7}$$

The optimal level of consanguinity chosen by families will be a positive function of illiquid land wealth, implying that levels of consanguinity will be higher in rural areas where land ownership is more common.[8] In addition, since the belief that consanguinity is detrimental to child quality is increasing in schooling, the level of consanguinity chosen by families will decrease with schooling, *ceteris paribus*.

[7]In the context of rural Pakistan, marital decisions are primarily taken by the parents of the bride and groom. In the Pakistan Labor Skills Survey data used by the paper, 87% of the married women interviewed reported that the main decision maker in choosing their spouse was either their mother or father, while the respondent reported that she was the main decision maker in only 10% of the cases. For married men, 17% responded that they were the main decision maker while 81% reported that the main decision maker was the father or the mother. Similarly, Mobarak et al. (2013) report that the vast majority of respondents who married a cousin did so based on their parents' wishes in Bangladesh.

[8]The model assumes a linear and positive relationship between land wealth and consanguinity. While some studies in other contexts assume a non-linear relationship, the data from the Pakistan Labor Skill Survey shows no evidence for strong non-linearities in the rural sample. Once we regress the dummy variable for whether the parents married a relative on the area of land owned by the grand-father, the linear OLS coefficient is positive and highly statistically significant. Once a quadratic term is also included to account for non-linearities, the quadratic term is statistically insignificant and the R-squared is virtually unchanged compared to the specification with the linear term only.

$$d\delta^*/ds < 0 \tag{8}$$

The model suggests distinguishing among four cases:

(1) $w_l > 0$ *and* $\vartheta > 0$: δ has two opposing effects on expected U: a positive effect via w (wealth effect), and a direct negative effect via *Q (child quality effect)*;

(2) $w_l > 0$ *and* $\vartheta = 0$: only the *wealth effect* plays a role in decision making, and the highest degree of consanguinity (*δ)* will be preferred;

(3) $w_l = 0$ *and* $\vartheta > 0$: δ only has an impact on Q and a negative and direct effect on child quality, and $\delta = 0$ will be preferred*;*

(4) $w_l = 0$ *and* $\vartheta = 0$: families will be indifferent between various levels of *δ.*

4. Empirical Estimation of the Effects of Consanguinity

As in Becker and Tomes (1976), equation (3) can be written as an additive function, to express the quality of children produced by the household as:[9]

$$Q = e - \delta v(s) + w_o \tag{9}$$

Recognizing that father's education, child age, gender and geographic location may be independently correlated with child outcomes, we control for these additional covariates, X, and empirically estimate the model as[10]:

$$Q = \alpha + \beta\delta + \gamma w_o + X'\theta + \varepsilon \tag{10}$$

[9]A necessary condition for the reduction of equation (3) into an additive form is a constant marginal product of parental inputs.

[10]We also control for the quadratic term of father's education and the interaction between father's education and geographical location.

where the child's endowment e is captured by the error term ε. Estimating β by simple OLS would yield biased estimates if child endowments are correlated with δ, or more generally, if some unobservables affecting the decision to marry within the family also determine child quality. In this model, beliefs about the detrimental effects of consanguinity, ϑ, are potentially important unobservables which are likely to be associated with both δ and ε. Further, given the potential economic motives for marrying cousins, there may be a positive bias associated with parents who are related if there are fewer disagreements over investing available resources in child quality. For this reason, to identify β we need one or more instruments correlated with the decision to marry a relative, but after controlling for current family wealth and other covariates, uncorrelated with those child endowments that are unrelated to consanguineous marriage of their parents. Within our theoretical framework, the decision to marry within the family is affected by land wealth w_l, of the child's grandparents and parental schooling *s.*

$$\delta = \rho + \pi w_l + \rho s + u \tag{11}$$

Assuming linearity and additivity, and including father's education among exogenous regressors, X the first stage used to identify determinants of consanguineous marriage can be estimated as:

$$\delta = \rho + \pi_1 w_l + \pi_2 w_o + X'\vartheta + u \tag{12}$$

In order to use (12) to identify a causal effect of consanguineous marriage, δ, on child quality, *Q*, we must be confident that, after controlling for total current household wealth, w_o, land ownership of a child's grandparent (in the past if not no longer alive, or at present if still living) is uncorrelated with both child endowments and parent investments in child quality. Contrary to liquid wealth, markets for land wealth are thin in rural South Asia (Griffin, Khan and Ickowitz, 2000). Land wealth is therefore unlikely to be converted into child quality inputs and therefore to directly affect child quality. Since in the model w_l (land ownership) only affects *Q* indirectly via its effect on δ in (12) but does not enter (10) directly (exclusion restriction), it will identify the effect of δ on Q as long as errors in measurement of wealth are not systematically related to both land wealth and child outcomes.

As grandparents may have had a strong influence on the marriage decision of a child's parents (Holy, 1989), we also explore using indicators of whether a grandparent was alive at the time the child's parents were married as instruments. As the dataset does not directly collect information on whether a given grandparent was alive at the time of marriage, we instead use a dummy variable for whether the grand-parent died before the median age of grand-parental death in the sample for rural Pakistan, which is age 65.

One key concern about the identification strategy is the violation of the exclusion restriction underlying the IV approach. One could suspect that grandfather's land ownership affects the nutritional status and cognition of grandchildren through channels other than consanguinity. For example, grandfather's land ownership might be correlated with how much influence he has on the grandchildren's care and thus affect health and cognitive outcomes, independently of consanguinity. The Pakistan Labor Skills survey collects data on decision-making through a set of questions that are asked to women in the household. Specifically, respondents are asked to name the primary and secondary decision makers for a range of decisions, including child's education, fertility and children's discipline. 89% of interviewed women report that the main decision maker about children's education is either the husband (76%) or the wife (13%), while other relatives represent a very small minority. The patterns are very similar for decisions about how a child should be disciplined. Results are also very similar regarding the decision to have another child. The questionnaire further asks whether there is a second decision maker for those issues. In about 50% of the cases, it is reported that there is a second decision maker, and in between 85% to 90% of the cases, it is again reported to be one of the two spouses. Although this does not entirely preclude the grand-father from having some residual influence on decisions related to the child's care, this corroborative evidence significantly alleviates concerns about the validity of the exclusion restriction.

In addition, one may also be concerned about the fact that household non-land wealth is treated as exogenous in our setting. The underlying assumption is that the wealth index captures elements of household long-term wealth, rather than short-term fluctuations in wealth. The index was constructed using principal component analysis by excluding land, and therefore consists only of non-wealth land. The wealth index includes primarily dwelling characteristics as well as households' durable goods which are mostly fixed in the short-run. To check whether the index of household wealth captures short-term fluctuations in household income, we correlate the wealth index to self-reported shocks that affected the household in the past 12 months. The survey collects detailed information on a range of negative shocks including natural, calamities, agricultural shocks, economic shocks, or health shocks, which are expected to generate short-term fluctuations in household income.

The results of regressing the wealth index on a dummy for whether the household has been hit by the corresponding shock in the past 12 months are reported in Table A.1. The p-value of the OLS coefficients on shocks is statistically insignificant for 18 out of 20 shocks tested.[11] In addition, while one would expect the association between a negative shock and the wealth index to be negative if the index is capturing short-term income fluctuations, about half of the coefficients are positive, suggesting that those correlations capture random fluctuations of the wealth index with respects to shocks, rather than

[11] Although information on 42 different shocks was recorded in the original survey, we do not run the test for shocks that hit less than 1% of household, as the estimated correlations would likely be driven by a handful of observations given the small sample test. After applying these restrictions, we can estimate the association between the wealth index and twenty different shocks that could have hit the household in the past 12 months by running separate regressions of the wealth index on a dummy for whether the household was hit by the corresponding shock. We also control for district fixed effects in each individual regression to account for other unobservable factors at the district level that could be correlated with the incidence of shocks.

systematic associations. Among the two coefficients that turn out to be statistically significant, births in the family appear to be negatively associated with the wealth index, but the association is only marginally significant at the 10% level. The only coefficient that is statistically associated with the wealth index at the 5% level is the one associated with flooding in the past 12 months. However, the sign of the association is positive, which is the opposite of one would have expected in the presence of a negative shock.

In sum, there is no robust evidence to suggest that the wealth index is associated with endogenous short-term fluctuations in income, particularly after considering standard multiple inference tests, although one of the tests yields statistically significant results. Anderson (2008), and earlier scholars, emphasize that the likelihood of type I errors (false rejection of the null hypothesis) increases with the number of outcomes tested, and so standard errors on such multiple inference tests need to be adjusted accordingly. Several methods are available to correct standard errors of individual tests under such a multiple inference exercise as conducted here. The simplest and most popular method is the Bonferroni correction, which simply multiplies each p-value of the coefficients of interest by M, the number of tests performed. As twenty different tests are conducted in this case, none of the adjusted p-values for individual tests are below the standard critical levels after applying the Bonferroni correction. These results offer further comfort that the wealth index employed is not capturing short-term endogenous fluctuations in household income.

Finally, the theoretical and empirical model abstract from potentially endogenous fertility decisions for the sake of simplicity, as incorporating multiple types of decisions would make both the model and empirical estimation quite complex. This is acceptable empirically if there is no endogenous relationship between child quality, consanguineous unions and fertility. Although the endogeneity of fertility decisions cannot be directly tested, we assess whether there exist systematic differences in the number of children of women in consanguineous unions versus women who are not married to a relative. When we test for a simple difference in means between the two groups, the mean number of children in the two groups is quite similar, and the null hypothesis of the number of children being equal in the two groups of women is not rejected. We therefore do not find suggestive evidence that women selected into consanguineous unions have more or less children in rural Pakistan.

5. The Pakistan Labor and Skills Survey

This paper uses data from the rural sample of the of the Labor and Skills Survey (LSS) wave 2, conducted in Pakistan in 2013.[12] The survey is representative at the national and provincial level and covers all regions of Pakistan except Balochistan and the Federally Administered Tribal Areas, which represent less than 7 percent of the total population.[13] The final sample used to estimate the impact of consanguinity on cognitive development outcomes consists of 1,411 children aged 5-13 from 60 rural villages. When looking at the effect of the treatment on children's height-for-age, the sample size is 1,285 children.[14]

The LSS consists of a detailed household roster, a female and male questionnaire, and a cognitive assessment taken by all adults in the household and all children aged 5-13.[15] The household roster collects general information on all household members including demographics and education, but also anthropometrics and overall health status, parental land ownership and parental mortality. It also inquired about the degree of consanguinity between not only household members and their spouse, but also of their parents and grandparents. Degrees of consanguinity are captured by 3 different categories:

[12]We also conducted the analysis in the urban sample of the survey. However, the set of instruments used in rural areas lack power to identify the impact of consanguinity in urban areas. Therefore, the causal impact of consanguinity in urban areas could not be identified. As a result, they are not reported in the paper but are available upon request.
[13]Those areas could not be covered by the survey due to security reasons.
[14]Following international standards, children with z scores below -6 or above 6 were excluded from this estimation sample.
[15]The female and male questionnaires were administered to one randomly selected male and one randomly selected female in the household aged 15-64.

1st cousins, 2nd cousins, or not-related. In the rural sample, close to 65 percent of children are born to consanguineous unions (55 percent of children have parents who are 1st cousins and 10 percent of children have parents who are 2nd cousins). These figures are broadly consistent with Hussain and Bittles (1998) estimate of 60 percent consanguinity. Consistently with the model, consanguinity rates are significantly lower among parents that declare being aware of the detrimental effects of consanguinity (55% against 69%), although they are still high among the sub-sample of parents that declare being aware of the detrimental effects of consanguinity.[16]

Based on this information, we construct four different indicators for consanguinity, including three indicator variables: a dummy variable if parents are related, a dummy variable if parents are first cousins, and a dummy variable for second-cousin parents. The data also allow us to construct the "f-coefficient of inbreeding" or kinship in the literature, which measures the probability at which a subject has inherited an identical copy of a gene from both parents (i.e. homozygosity).[17] Using this metric in our context and referring to biological classifications, the f-coefficient for 1st cousins was given the value 0.125, second cousins was given the value 0.031, while the f-coefficient for unrelated individuals takes the value 0.

The measure of cognitive ability is obtained from the Raven's test of progressive matrices, administered to all children aged 5 to 13 in the LSS households. The Raven's test of progressive matrices aims at measuring logical reasoning ability. The instrument consists of 36 questions of increasing difficulty in which the child is asked to identify the missing figure in a logical sequence of colored figures. We

[16] The question on whether parents are aware of the detrimental effects of consanguinity was only asked to a subsample of the original sample, with prevents conducting the main analysis in these two separate subsamples.
[17]More precisely, the f-coefficient of inbreeding expresses the expected percentage of homozygosity arising from a given system of breeding. For a given gene with equally common dominant and recessive variants *A* and *a*, a random-bred stock will be 50% homozygous (25% *AA* and 25% *aa*), while a closely inbred population will be 100% homozygous (100% *AA* or 100% *aa*). The coefficient of inbreeding *f* is thus designed to run from 0 for an expected 50% homozygosis to 1 for an expected 100% homozygosis, $f=2h-1$, where *h* is the chance of finding homozygosis in this gene.

construct the measure of cognitive development from the answers given to each of the thirty-six items using Item Response Theory (IRT). Calculating a score through IRT has the advantage in that it optimally exploits data from a test by weighting questions by contribution to determining differences among respondents. In this sense, the scores provide a more precise measure of ability than a raw score or z-score. Though initially used in psychometrics, Das and Hammer (2005), among several others, have used this approach to optimally calculate test scores for economic applications.[18] IRT assumes that there is an underlying latent random variable, θ, and every question in a test maps this latent variable to a response. The IRT method estimates the relationship between the latent trait of interest, in our case cognitive ability, and the Raven's question items intended to measure the trait using maximum likelihood methods. In this paper, we use a two-parameter logistical model to estimate IRT scores.[19] As the main advantage of this approach, compared to using raw scores or z-scores, is that it takes into account differences in difficulty of the thirty six test questions in the calculation of the cognitive score, it is not surprising that the correlation between Raven's raw scores and IRT scores is quite large at 0.97.[20]

The survey does not report expenditure data but collects detailed information on durable goods and assets owned by the household, and on the characteristics of the dwelling in which the household lives. Following a well-established approach pioneered by Filmer and Pritchett (2001), we use this information to construct an index for the household's contemporaneous wealth based on principal component analysis. The estimate of relative wealth using the PCA is based on the first principal component, which we use as a proxy for household's wealth.[21] Finally, height and weight measurements were obtained from all household members aged five and above.

---

[18]See Rasch (1960), Birnbaum (1967), or Hambleton et al. (1991).
[19]Three-parameter models are also used in the literature but tend to require larger sample sizes than available in the LSS to converge.

[21]This asset-based measure aims at capturing the household's long-run economic status. The Filmer and Pritchett (20-01) results were validated using both household assets and consumption data for a set of low of middle-

[Table 1 about here]

6. Empirical Analysis

6.1. Determinants of Parental Consanguinity

We first discuss determinants of (parental) consanguinity both because the question is of interest in itself, but also because the consanguinity estimates will serve as first-stage equations in estimates of the effect of consanguinity on child height and cognitive ability. Using a linear probability model, we regress three alternative indicators of parental consanguinity on a set of family and individual characteristics that are likely to affect parental consanguinity. The most important factor that increases the likelihood of (parental) consanguinity is grandfather's ownership of land.[22] This is consistent with the (mostly qualitative) literature on the topic, where illiquid wealth, and in particular landholding of the previous generation, is seen as an important determinant of consanguineous marriage.[23] At the same time, contemporary household liquid wealth as captured by an asset index is negatively related to consanguineous marriage.

income countries, including Pakistan, and it was concluded that PCA "provides plausible and defensible weights for an index of assets to serve as a proxy for wealth". Following Filmer and Pritchett (2001), many studies, especially in the fields of economics and public policy, have implemented and recommend the use of PCA for estimating wealth effects.

[23] The extent of the association between grand-father's land ownership on the mother's and father's side and consanguinity is very similar in the data, even though women cannot in practice inherit land in Pakistan. When one distinguishes between grandfather's land ownership on the mother and father's side and estimate the association between these dummy variables and grand-father's land ownership in separate regressions, the magnitude of the coefficients on the mother's side and father's side is very similar. We therefore chose to use a single indicator for grandfather's land ownership, which indicates whether one of the grand-fathers owned land.

We also explore the effect of the early death of one of the grand-parents on consanguinity, as an indirect proxy for the grand-parent being alive at the time of marriage. In a context like Pakistan where parents play a crucial role in marital decisions and are also expected to have a preference for consanguineous marriages for the reasons outlined in Section 3, the death of one of the grand-parents before the parents' marital age is expected to decrease the likelihood of marrying a relative.[24] More specifically, within the set of four grand-parents, the maternal grand-father is expected to play a crucial role due to patrilocality of the Pakistani society highlighted by Holly (1989), and the fact that grooms cannot credibly commit *ex- ante* to treat a prospective wife well (Jacoby and Mansuri, 2010).[25] In Table A.2. of the Appendix, we report the association between the dummy for whether parents are related, and a dummy for the early death of each of the grand-parents. The association between having a consanguineous marriage and the early death of each of the grand-parents is negative for all grand-parents. However, this negative association is much larger in magnitude and statistically significant only for the maternal grand-father, which is consistent with Holly (1989).

We set age 65 as the cut-off age for the early grand-parental death dummy variable, as it is the median age of grand-parental death in the sample. Table A.3 illustrates the logic behind this instrumental variable strategy, where we look at the incidence of consanguinity for different timings of maternal grand-father's death. The table shows a strong pattern between the age at which the maternal grand-father passed away, and the incidence of a consanguineous marriage between the parents. As suspected, the later the death of the maternal grand-father, the higher the incidence of a consanguineous marriages.

[24] Close to 90% of individuals in the sample report that their mother or father was the main decision maker is choosing their spouse.

[25] In patrilocal societies like Pakistan, daughters leave their parental homes while married sons do not. As daughters move to a new household, the monitoring costs of her treatment increase significantly and a marriage to a relative could decrease the likelihood of poor treatment. In this context, the death of the patriarch would diminish the ability of the family to provide protection to a daughter who is then more likely to move to an unrelated household.

When included as regressor in the determinants of consanguineous unions, maternal grandfather's death before age 65 is linked to a statistically significant reduction in parental consanguinity, which is consistent with the hypothesis that the maternal grandparent may prefer a consanguineous union for both economic and cultural reasons. As already suggested by Table A.2., only the coefficient associated with maternal grandfather is significant when the full set of grand-parental survival indicators is included in the regression. As those additional grand-parental death instruments drive down the first-stage F-statistics and would aggravate the weak instrument issue we face, we choose to restrict the grand-parental death instrument to the maternal grand-father only.

[Table 2 about here]

6.2. Consequences of Consanguinity for Children: Ravens Test Results and Stunting

Consanguinity and Cognitive Performance. Children born into consanguineous unions have lower cognitive abilities, as captured by Ravens test scores (Table 3). OLS estimates, shown in column (1) suggest that children with related parents score 0.14 standard deviations lower on a Ravens test than children with unrelated parents. Other measures of consanguinity, the F-coefficient and treating parents as first and second cousins separately, also suggest a negative, but insignificant impact of consanguinity. Further, when including district fixed-effects, to control for unobservables related to local economic development, access to off-farm labor markets and local customs, we note that children with related parents have a statistically insignificant score that is 0.1 standard deviations lower than those with unrelated parents.

As suggested in the discussion above, the economic motives behind consanguineous unions are likely to improve the financial health of households and thus also access to child inputs (food and healthcare). Thus, there will be an upward bias on estimates of the biological effects of consanguinity on child outcomes. In addition, the endogeneity of consanguineous unions could also partly why the negative effects of parents being second cousins are stronger than for parental being first cousins in the OLS estimates, which can be puzzling at first.[26] We thus next introduce an instrumental variables approach, estimated through an instrumental variable limited information maximum likelihood (IV-LIML) approach as it is less susceptible than other IV estimators to weak instrument bias, and show results in column 5 (province fixed effects) and column 6 (district fixed effects). Consistent with our expectation of an upward bias in OLS, the IV-LIML estimates suggest that children born to parents in consanguineous unions score 0.83 and 1.09 standard deviations lower, in models with province and district fixed effects, respectively, on Ravens Progressive Matrices tests. The negative effect of consanguinity on cognitive development is significant at the 5 percent level in both models.

The results of the Hansen J-test for over-identification show no evidence against the identification strategy, but the F-statistics on excluded instruments are 8.27 and 5.30 in models with province and district fixed effects. Although jointly significant at better than one percent, the F-statistics are below 10, the standard rule of thumb at which researchers should be concerned with possible weak instrument bias (e.g., Stock and Yogo, 2005). Note that we are clustering standard errors at the village level, and the "rule-of-thumb" was derived in an environment in which errors can be considered independent, thus standard thresholds may not apply with cluster-corrected F-statistics Cameron and Miller (2011, 2015). Appropriate *F*-statistic thresholds for the presence of weak instrument bias are likely to vary by application and we thus explicitly test for weak instrument bias. Following the suggestion of Cameron and Miller, we use a cluster robust version of Moreira's (2003) Conditional Likelihood Ratio test, derived using the method of moments (Finlay and Magnusson 2009). We also use this test statistic to generate weak instrument robust

[26] Given that the probability to inherit the same version of a gene is higher when parents are more closely related, one would expect the effects of consanguinity to be stronger if parents are first cousins than if they are second cousins.

95 percent confidence intervals around the coefficient estimate on the consanguinity variable. The CLR test shows that the coefficients on parental consanguinity in the IV model are negative and statistically different from zero, in models that alternatively include province and district fixed effects.

The lower bound of the CLR test confidence interval indicates that the negative impact of consanguinity is at least 3 times larger than the estimated effect using Ordinary Least Squares. This underlines the importance of accounting for the endogeneity of consanguinity when attempting to estimate its causal impact on child outcomes, which, to the best of our knowledge, has not been done by previous literature. By failing to control for positive bias due to endogeneity, ordinary least squares approaches underestimate the negative effect of consanguinity on child development outcomes.

Other explanatory variables have expected coefficient signs. The cognitive test scores increase with child age, which is intuitive. Household wealth, as captured by an asset index has a large positive and statistically significant effect on test scores. Male children tend to score higher, although the estimated coefficients are not statistically significant. Location within Pakistan appears to matter as well. In the IV specification, children in KPK score systematically lower than children in Punjab.

[Table 3 about here]

Consanguinity and Stature. The children of related-parents are likely to have smaller height for age z-scores compared to other children, with negative coefficients on regressors indicating that parents are related in both OLS and 2SLS specifications (Table 4). The magnitude of the estimated effect increases significantly in IV-LIML models with province fixed effects, and it is statistically significant at the 10% level. The magnitude of the effect remains with district fixed effects, but the estimates are not statistically significant. After identifying consanguinity, the IV estimates indicate that children's height-for-age z scores are reduced by 1.35 standard deviations.

[Table 4 about here]

Next we examine the impacts of parent consanguinity on the likelihood that children are moderately and severely stunted. Following WHO classifications, moderate stunting is defined as having a height for age z-score below -2, and severe stunting below -3. The incidence of stunting and severe stunting is high in rural Pakistan: in the sample 47% are classified as stunted, and 32% suffer severe stunting.

[Table 5 about here]

The IV estimates reported in Table 5 and 6 indicate that consanguinity increases the likelihood of stunting substantially, with a larger and statistically significant impact on severe stunting (Table 6, columns 5 and 6). Having blood-related parents raises the probability of being extremely stunted by more than 30 percentage points. The CLR test confidence intervals suggest that the point estimate of the effects of stunting are above zero with 95% confidence. While results examining the effect of consanguinity on stature are significant only for severe stunting at the 10th percentile, they are consistent with the cognitive outcome results, and recommend further study examining the causal impact of consanguinity on stunting and anthropometrics in study designs promising more statistical power.

[Table 6 about here]

The magnitude of the estimated negative effects of consanguinity are substantial in the instrumental variable specifications. Regarding the effect on cognitive scores, the estimated coefficient which ranges from about -0.8 to -1.1 implies that marrying a relative decreases children's cognitive test scores by about one standard deviation in cognitive tests, which is very large in magnitude. This is equivalent to moving from the median score of the cognitive test score distribution to the 10th percentile, or to moving from the 90th percentile of the distribution to the 50th percentile. Regarding the effects on nutritional status, the magnitude of the IV estimates also quite large. Height z-scores are estimated to be reduced by about half a standard deviation by consanguineous union, which is equivalent to dropping from 75th percentile of the height for age distribution to the 50th percentile, or from 50th percentile of the distribution to the 30th percentile.

The large magnitude of the IV results may be partly explained by the fact that it reflects a local average treatment effect: the IV results pick up the effect of consanguinity for the subset of the population whose decision may be influenced by grandparent land ownership or early death of a maternal grandfather (after controlling for contemporaneous long-term wealth). Additionally, given the small sample size combined with the fact that the instruments are relatively weak, we cannot rule out that the precision of the IV estimates is an issue in some of the specifications, as suggested by large standard errors. This could explain why the estimates for the effect of consanguinity on height for age and severe stunting are only marginally significant, despite being quite large in magnitude.

At first, our findings appear to contrast with those of Leroy at al. (2016) for Uganda. The authors find an effect on children in the top half of the HAZ distribution, but not on stunting rate and argue that the program might have mostly affected better off children and not those near the cutoff. However, in their paper, the average HAZ was close to the -2 cut off, and since in the normal distribution large numbers are close to the mean, a small shift in the mean could lead to large numbers crossing the cut-off, which is actually not different than our shift in stunting. Second, while the authors found most of the movement in the *relatively* well off, the discussion in that paper also makes it clear that this result was unexpected and might reflect the high incidence of extreme poverty. The authors also make it clear that those later findings are quite specific to the environment of the study.

Awareness about the detrimental effects of consanguinity is highly endogenous and this variable in only available for a sub-sample of the main sample, which reduces the precision of our estimates, especially in the presence of weak instruments. Despite these limitations, we also run the baseline estimates for the effects of consanguineous unions on the two separate subsamples of parents that declare being aware or not aware of the detrimental effects of consanguinity. The results of the IV estimations in the two subsamples are reported in Table A.4. of the Appendix. Our results show stronger negative effects of consanguineous unions among the subsample of children whose parents report not being aware of the detrimental effects of consanguinity. Awareness is highly endogenous and reflect many unobservable parental characteristics, which could partly drive those findings. In addition, one possible explanation could be that in the subsample where individuals are aware of the detrimental effects of consanguinity, parents that still chose to marry a relative may compensate for the expected detrimental effects of consanguinity by providing additional inputs to the child. In contrast, in the group of unaware parents, those compensatory investments would not occur, and the negative effects of consanguinity on child outcomes.

7. Conclusions

Utilizing a unique household survey from Pakistan, we find strong evidence linking consanguinity to reduced cognitive abilities and higher incidence of severe stunting among children. The magnitudes of the estimated effects are much more pronounced in instrumental-variables specifications where we are able to single out causal effects by treating parental consanguinity as an endogenous variable.

Given suspicion that parental consanguinity may be related to a range of negative health outcomes, it is natural to ask what factors drive the decision to marry a relative. In rural Pakistan, we find that if a child's grandfather owns (or owned) land, then the likelihood of his/her parents' being in a consanguineous relationship increases significantly. We also find confirmation of the potential importance of maternal grandfather bargaining power: if the maternal grandfather passes away early, then family influence on parental marriage decision-making "changes" and a child's parents are less likely to be in a consanguineous union.

What do these findings mean for public policy? First, the negative effects of consanguinity go beyond selected severe disabilities and thus there is reason to be concerned about such high prevalence of consanguineous marriages in the developing world. Second, if left-alone, the prevalence of consanguinity is likely to decrease very slowly. With urbanization, increased educational attainment, increased non-land wealth and improvements in both the business and public order environments one can expect to see some decreases in consanguinity, but the effects of these changes may be gradual and relatively modest. Indeed, discussing trends in consanguineous marriages, Saggar and Bittles (2008) also come out as somewhat pessimistic that this practice will end on its own over time, observing that "it could be argued that the ongoing wide-spread popularity of consanguinity makes a rapid decline in its prevalence improbable. In many developing countries, strenuous official efforts are being made to lessen the appeal of close-kin unions, although with no apparent appreciation or acknowledgement of the balancing social and economic benefits."

Outright prohibition of consanguinity through legislation, perhaps through prohibition of first-cousin marriage given its stronger linkage to poor development outcomes might be an option (the practice in Western countries varies: cousin marriages are legally banned in 24 out of 50 states of the United States of America but they are not prohibited under UK law). In Pakistan and many other developing countries, the state's ability to enforce such legislation might be limited. In fact, negative externalities to such policy cannot be ruled out: for example some might avoid formalizing or reporting their marital status, others might avoid getting national ID cards (which is a requirement for establishing Bank accounts, voting, driving, getting a cell phone among others).

Simple dissemination of information to the public about the possible negative results of consanguineous marriages may provide an easier means of reducing their incidence. Furthermore, information interventions (even those about future intentions) are feasible to implement at relatively low cost (e.g. Leibman and Luttmer, 2015). A large share of the adult population, 72.4 percent in the LSS survey, report that they are not aware of negative consequences from consanguinity. As consanguinity is more prevalent among low-income households (except for rural landholding families), such information campaigns might target the poor, who are the targets of new social safety net programs. Indeed, pairing an information intervention with social insurance schemes aiming to reduce exposure to earnings risk may be a useful approach. Further, preconception consultation programs, similar to those implemented in Iran and Saudi Arabia (Hamamy, 2012), might be focused on carrier detection and genetic counseling. While such screening and genetic-counseling might prevent certain types of disabilities, other negative effects of consanguinity on children's cognitive ability might not be necessarily addressed by these interventions.

References

Agarwal, Bina. 1994. *A Field of One's Own: Gender and Land Rights in South Asia*. Cambridge: Cambridge University Press, 1994.

Al-Kandari, Yagoub and Douglas E. Crews. 2011. The effect of consanguinity on congenital disabilities in the Kuwaiti population, *Journal of Biosocial Science* 43 (1): 65-73.

Anderson, Michael. 2008. Multiple inference and gender differences in the effects of early Intervention: A reevaluation of the Abecedarian, Perry Preschool, and early training projects, *Journal of the American Statistical Association* 103(484): 1481-1495.

Barbour, Bernadette and Pascale Salameh. 2009. Consanguinity in Lebanon: prevalence, distribution and determinants, *Journal of Biosocial Science* 41(4): 505-517.

Bashi, Joseph. 1977. Effects of inbreeding on cognitive performance, *Nature*, 266(5601), 440-442.

Becker, Gary S. and Nigel Tomes. 1976. Child Endowments and the quantity and Quality of children, *Journal of Political Economy* 84(4): 143-162.

Becker, Gary S. 1981. *A Treatise on the Family.* 424pp. Cambridge: Harvard University Press.

Birnbaum, Allan. 1967. Some Latent Trait Models and Their Use in Inferring an Examinee's Ability. In Lord, Frederic M. and M.R. Novick, eds., *Statistical Theories of Mental Test Score*. London: Addision-Wesley Publishing Company.

Bittles, Alan H. 1994. The role and significance of consanguinity as a demographic variable, *Population and Development Review* 20(3): 561-584.

Bittles, Alan H. 2001. Consanguinity and its relevance to clinical genetics, *Clinical Genetics* 60(2): 89-98.

Botticini, Maristella and Aloysius Siow. 2003. Why Dowries? *American Economic Review* 93(4): 1385–1398.

Caldwell, John C., Palli H. Reddy and Patricia Caldwell. 1983. Causes of marriage change in south India, *Population Studies* 37(3): 343–361.

Caldwell, John C., Palli H. Reddy and Patricia Caldwell. 1988. *The Causes of Demographic Change: Experimental Research in South India*. Madison: University of Wisconsin Press.

Cameron, Colin A. and Douglas L. Miller. 2011. Robust Inference with Clustered Data, in A. Ullah and D.E. Giles eds., *Handbook of Empirical Economics and Finance*, CRC Press.

Cameron, Colin A. and Douglas L. Miller. 2015. A Practitioner's Guide to Cluster-Robust Inference, *Journal of Human Resources* 50(2): 317-372.

Das, Jishnu and Jeffrey S. Hammer. 2005. Which Doctor? Combining Vignettes and Item Response to Measure Clinical Competence, *Journal of Development Economics* 78(2): 348-383.

Denic, Srdjan and Michael G. Nicholls. 2007. Genetic benefits of consanguinity through selection of genotypes protective against malaria, *Human Biology* 79(2):145-58.

Do, Quy-Toan, Sriya Iyer and Shareen Joshi. 2013. The economics of consanguineous marriages, *The Review of Economics and Statistics* 95(3): 904-918.

Dorsten, Linda Eberst, Lawrence Hotchkiss and Terri M. King. 1999. The effect of inbreeding on early childhood mortality: twelve generations of an Amish settlement, *Demography* 36(2): 263-271.

Farah, Abdul-Aziz and Samuel H Preston. 1982. Child mortality differentials in Sudan, *Population and Development Review* 8(2): 365-383.

Filmer, Deon and Lant H. Prichett. 2001. Estimating wealth effects without expenditure data—or tears: An application to educational enrollments In States of India, *Demography* 38(1): 115–132.

Finlay, Keith and Leandro Magnusson. 2009. Implementing weak-instrument robust tests for a general class of instrumental-variables models, *Stata Journal* 9(3): 398-421.

Givens, Benjamin P. and Charles Hirschman. 1994. Modernization and consanguineous marriage in Iran, *Journal of Marriage and Family* 56(4): 820-834.

Goody, Jack. 1973. Bride wealth and Dowry in Africa and Eurasia, in Jack Goody and Stanley Tambiah (eds.), Bridewealth and Dowry, Cambridge: Cambridge University Press.

Griffin, Keith, Azizur R. Khan, and Amy Ickowitz. 2000. *Poverty and Distribution of Land.* United Nations Development Programme.

Hamamy, Hanan. 2012. Consanguineous marriages. Preconception consultation in primary health care settings, *Journal of Community Genetics* 3(3): 185-192.

Hambleton, Ronald K., H. Swaminathan and H. Jane Rogers. 1991. *Fundamentals of Item Response Theory*. 184pp, London: Sage Publications.

Hasnain, Fariha and S. K. Hashmi. 2009. Consanguinity among the risk factors for underweight in children under five: a study from rural Sindh, *Journal of Ayub Medical College* 21(3): 111-116.

Holy, Ladislav. 1989. *Kinship, honour and solidarity: Cousin marriage in the Middle East.* 143pp. Manchester: Manchester University Press.

Hussain, Rafat. 1999. Community perceptions of reasons for preference for consanguineous marriages in Pakistan, *Journal of Biosocial Science* 31(4): 449-461.

Hussain, Rafat. and Alan H. Bittles. 1998. The prevalence and demographic characteristics of consanguineous marriages in Pakistan, *Journal of Biosocial Science* 30(2): 261-275.

Jaber, Lufti, Gabriella Halpem and Mordechal Shohat. 1998. The impact of consanguinity worldwide, *Community Genetics* 1(1): 12-7.

Jacoby, Hanan G., and Ghazala Mansuri. 2010. Watta Satta: Exchange Marriage and Women's Welfare in Rural Pakistan, *American Economic Review* 100(4): 1804–1825.

Jaouad, Imane, Siham Chafaï Elalaoui,, A. Sbiti, F. Elkerh, L. Elkerh Belmahi and Abdelaziz Sefiani. 2009. Consanguineous marriages in Morocco and the consequence for the incidence of autosomal recessive disorders, *Journal of Biosocial Science* 41(5): 575-581.

Kaiser, Jocelyn. 2016. When DNA and culture clash, *Science* 354(6317): 1217-1221.

Kanaan, Zihaad, Rami Mahfouz and Hala Tamim. 2008. The prevalence of consanguineous marriage in an underserved area in Lebanon and its association with congenital anomalies, *Genetic Testing* 12(3): 367-372.

Khlat, Myriam, Suzan Halabi, Adele Khudr, Vazken M. Der Kaloustian, John M. Opitz, James F. Reynolds. 1986. Perception of consanguineous marriages and their genetic effects among a sample of couples from Beirut, *American Journal of Medical Genetics* 25(2): 299-306.

Kuper, Adam. 2009. *Incest & influence: the private life of bourgeois England*. 304pp. Massachusetts: Harvard University Press.

Lakhan, Ram, Rajshekhar Bipeta, Srinivasa Yerramilli, and Vinayak Nahar. 2017. A Family Study of Consanguinity in Children with Intellectual Disabilities in Barwani, India, *Journal of Neurosciences in Rural Practice* 8(4): 551-554

Leibman, Jeffrey B. and Erzo Luttmer. 2015. Would People Behave Differently if they Better Understood Social Security? *American Economic Journal: Economic Policy* 7(1): 275-299.

Leroy, Jef, Deanna Olney and Marie Ruel. 2016. Tubaramure, a Food-Assisted Integrated Health and Nutrition Program in Burundi, Increases Maternal and Child Hemoglobin Concentrations and Reduces Anemia: A Theory-Based Cluster-Randomized Controlled Intervention Trial, t*he Journal of Nutrition* 146 (8): 1601–1608.

Magnus, Per, Kerri Van der Berg and Tor Bjerkedal. 1985. Association of parental consanguinity with decreased birth weight and increased rate of early death and congenital malformations, *Clinical Genetics* 28(3): 35–342.

Mensch, Barbara S., Susheela Singh, and John B. Casterline. 2005. Trends in the timing of first marriage among men and women in the developing world. in C. B. Lloyd, J. R. Behrman, N. P. Stromquist, and B. Cohen (eds.), *The changing transitions to adulthood in developing countries: Selected studies* (pp. 118–171). Washington, DC: The National Academies Press.

Mehrabi, Ali K. and Bahram Zeyghami. 2005. The effect of consanguineous marriage on congenital malformation, *Journal of Research in Medical Sciences* 10 (5): 298-301.

Mobarak, Ahbed M., Randall Kuhn and Christina Peters. 2013. Consanguinity and other marriage effects of a wealth shock in Bangladesh, *Demography* 50(5): 1845-1871.

Moreira, Marcelo J. 2003. A Conditional Likelihood Ratio Test for Structural Models, *Econometrica* 71(4): 1027-1048.

Morton, Newton E. 1978. Effect of inbreeding on IQ and mental retardation, *Proceedings of the National Academy of Sciences of the United States of America* 75(8):3906–3908.

Putnam, Robert D. 2000. *Bowling Alone: The Collapse and Renewal of American Community.* New York: Simon and Schuster.

Rami Reddy, V. and B.K. Chandrasekhar Reddy. 1979. Consanguinity effects of fertility and mortality among the Reddis of Chittoor District South India, *Indian Journal of Heredity* 11: 77-88.

Rasch, George. 1960. *Probabilistic Models for Some Intelligence and Attainment Tests*. 199pp. Chicago: The University of Chicago Press.

Saedi-Wong, Simin, Abdul R. Al-Frayh and Henry Y. H. Wong. 1989. Socio-economic epidemiology of consanguineous matings in the Saudi Arabian population, *Journal of Asian and African Studies* 24(3): 247-251.

Saggar, Anand K. and Alan H. Bittles. 2008. Consanguinity and child health, *Pediatrics and Child Health* 18(5): 244-249.

Saify, Khyber and Mostafa Saadat. 2012. Consanguineous marriages in Afghanistan, *Journal of Biosocial Science* 44(1): 73-81.

Sandridge, Amy L., Juman Takeddin, Essa Al-Kaabi and Yaakov Frances. 2010. Consanguinity in Qatar: knowledge, attitude and practice in a population born between 1946 and 1991, *Journal of Biosocial Science* 42(1): 59-82.

Shah, Gulzar H., Michael B. Toney and Brian L. Pitcher. 1998. Consanguinity and child mortality: the risk faced by families, *Population Research and Policy Review* 17(3): 275-283.

Shawky, Rabah M., Solaf M. Elsayed, Mouchira E. Zaki, Sahar M. Nour El-Din, Ferihan M. Kamal. 2013. Consanguinity and its relevance to clinical genetics, *Egyptian Journal of Medical Human genetics* 14(2): 157-164.

Stock, James H., and Motohiro Yogo. 2005. Testing for Weak Instruments in Linear IV Regression. In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. James H. Stock, Donald W. K. Andrews, 80–108. Cambridge, England, and New York: Cambridge University Press.

Tadmouri, Ghazi O., Pratibha Nair, Tasneem Obeid, Mahmoud Ali, Hajib Khaja and Hanan Hamamy. 2009. Consanguinity and reproductive health among Arabs, *Reproductive Health* 6(17): 1-9.

Temtamy, Samia and Mona Aglan. 2012. Consanguinity and genetic disorders in Egypt, *Middle East Journal of Medical Genetics* 1(1): 12-17.

Weinreb, Alexander A. 2008. Characteristics of women in consanguineous marriages in Egypt, 1988 – 2000, *European Journal of Population* 24(2): 185-210.

Zlotogora, Joel. 2002. What is the birth defect risk associated with consanguineous marriage? *American Journal of Medical Genetics* 109(1): 70–71.

Tables

Table 1. Descriptive Statistics, Children 5-13 in the LSS Rural Sample

Variable	Mean	Standard deviation	Observations
IRT cognitive score	-0.01	0.91	1411
Height for age (HAZ)	-2.38	2.62	1285
Moderately stunted (HAZ<-2)	0.47	0.50	1285
Severely Stunted (z-score<-3)	0.32	0.45	1285
Parents are related	0.65	0.48	1411
Parents are 1st cousins	0.55	0.50	1411
Parents are 2nd cousins	0.10	0.30	1411
F-coefficient of inbreeding	0.07	0.06	1411
Grand-father owns/owned land	0.53	0.50	1411
Maternal grand-father died before 65	0.30	0.46	1411
Age	8.78	2.51	1411
Being male	0.52	0.50	1411
Household wealth index	-0.40	0.84	1411
Father's years of education	3.76	4.71	1411
Punjab province	0.66	0.47	1411
Sindh province	0.22	0.42	1411
Khyber Pakhtunkhwa (KPK) Province	0.12	0.32	1411

Table 2. Determinants of Parental Consanguinity, Linear Probability Model

	Parents are Related		F-coefficient of Inbreeding		Parents are 1st Cousins	
	(1)	(2)	(3)	(4)	(5)	(6)
Grandfather Owned/Owns land	0.113***	0.064	0.013***	0.008	0.105**	0.068
	(0.039)	(0.047)	(0.005)	(0.005)	(0.043)	(0.044)
Maternal Grandfather Died before Age 65	-0.144***	-0.129***	-0.016***	-0.014**	-0.117**	-0.113**
	(0.043)	(0.041)	(0.006)	(0.006)	(0.050)	(0.049)
Age	0.010**	0.009**	0.001	0.001	0.009	0.008
	(0.005)	(0.005)	(0.001)	(0.001)	(0.005)	(0.005)
Male	0.011	0.012	0.001	0.002	0.017	0.018
	(0.023)	(0.023)	(0.003)	(0.003)	(0.026)	(0.026)
Wealth index	-0.066***	0.004	-0.007**	0.000	-0.053*	-0.006
	(0.027)	(0.025)	(0.004)	(0.003)	(0.030)	(0.028)
Father's Years of Education	0.005	0.006	0.001	0.001	0.005	0.004
	(0.006)	(0.005)	(0.000)	(0.001)	(0.005)	(0.005)
Dummy for Sindh Province	-0.119		-0.003		-0.021	
	(0.100)		(0.011)		(0.081)	
Dummy for KPK Province	-0.164**		-0.030**		-0.233***	
	(0.076)		(0.009)		(0.073)	
Father's years of education*Sindh	-0.014	-0.019**	-0.002*	-0.003**	-0.015	-0.016
	(0.011)	(0.009)	(0.001)	(0.001)	(0.010)	(0.008)
Father's years of education*KPK	-0.011	-0.016	-0.000	-0.001	0.006	-0.006
	(0.015)	(0.016)	(0.001)	(0.002)	(0.013)	(0.016)
District Fixed Effects	No	Yes	No	Yes	No	Yes
F-Test on Instruments	8.27	5.30	6.18	3.90	5.04	3.25
F-Probability	0.001	0.007	0.003	0.025	0.01	0.046
Number of observations	1411	1411	1411	1411	1411	1411

Notes: *: statistically significant at the 10% level; **: statistically significant at the 5% level; ***: statistically significant at the 1% level. Standard errors clustered at the village level are reported in parenthesis. The F-statistic, corrected for clustering at the village-year level, tests the hypothesis that the estimated coefficients on the instruments (Grandfather Owned/Owns Land and Maternal Grandfather Died before 65) are zero.

Table 3. Parental Consanguinity and Children's Ravens Score, Age 5-13

	Ordinary Least Square				LIML Estimator	
	(1)	(2)	(3)	(4)	(5)	(6)
Parents are related	-0.144**			-0.104	-0.832**	-1.09**
	(0.072)			(0.075)	(0.419)	(0.545)
F-coefficient of inbreeding		-0.816				
		(0.576)				
Parents are 1st cousins			-0.105			
			(0.072)			
Parents are 2nd cousins			-0.108			
			(0.086)			
Age	0.086***	0.086***	0.086***	0.087***	0.093***	0.095***
	(0.009)	(0.010)	(0.010)	(0.009)	(0.011)	(0.012)
Male	0.059	0.059	0.059	0.052	0.069	0.067
	(0.045)	(0.045)	(0.045)	(0.043)	(0.046)	(0.048)
Wealth index	0.156***	0.160***	0.159***	0.129***	0.104**	0.129***
	(0.049)	(0.051)	(0.050)	(0.045)	(0.050)	(0.046)
Dummy for Sindh Province	0.266*	0.284**	0.283**		0.146	
	(0.136)	(0.139)	(0.139)		(0.163)	
Dummy for KPK Province	-0.498**	-0.497**	-0.499**		-0.628***	
	(0.232)	(0.242)	(0.242)		(0.225)	
Father's Years of Education	0.017**	0.017**	0.018**	0.018**	0.021**	0.025**
	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.010)
	-0.016	-0.016	-0.017	-0.014	-0.026	-0.033*
Father's Years of Education*Sindh	(0.015)	(0.015)	(0.015)	(0.014)	(0.017)	(0.018)
Father's Years of Education*KPK	0.050*	0.052*	0.051*	0.031	0.044*	0.014
	(0.028)	(0.029)	(0.029)	(0.027)	(0.025)	(0.032)
District Fixed Effects	No	No	No	Yes	No	Yes
Test Statistics						
Hansen J statistic					0.930	0.001
p-value					0.305	0.972
F-statistic of Instruments					8.27	5.30
F-probability					0.001	0.007
CLR Test					5.01	4.87
p-value					0.036	0.04
95% Confidence interval based on CLR Test					[-2.12, -0.30]	[-3.93, -0.30]
Number of observations	1411	1411	1411	1411	1411	1411

Notes: Village-level cluster robust standard errors are reported in parenthesis. Columns (4) and (6) control for factors related to village location with district fixed effects. The Hansen J statistic tests for over-identification. We report cluster-corrected F statistics to test for weak instruments and then implement the conditional likelihood ratio (CLR) test developed by Moreira (2003) to test for weak instrument bias. The CLR test tests the hypothesis that the instruments and the coefficient on the endogenous variable are jointly zero. We also show the 95% confidence interval suggested by the CLR test. The CLR test is generalized for clustered dependence in error terms using the minimum distance approach by Finlay and Magnusson (2009).

Table 4. Parental Consanguinity and Children's Height Z-Score, Age 5-13

	Ordinary Least Squares				IV-LIML estimator	
	(1)	(2)	(3)	(4)	(5)	(6)
Parents are Related	0.013			-0.011	-1.35*	-1.52
	(0.183)			(0.182)	(0.793)	(1.20)
F-coefficient of Inbreeding		-0.040				
		(1.53)				
Parents are 1st cousins			-0.019			
			(0.197)			
Parents are 2nd cousins			-0.216			
			(0.311)			
Age	-0.097***	-0.097***	-0.097***	-0.091***	-0.083***	-0.077***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.021)	(0.023)
Male	0.245**	0.246**	0.245*	0.230*	0.260*	0.246*
	(0.122)	(0.122)	(0.123)	(0.117)	(0.137)	(0.135)
Wealth Index	-0.064	-0.065	-0.070	-0.014	-0.171	-0.013
	(0.127)	(0.128)	(0.127)	(0.118)	(0.142)	(0.114)
Dummy for Sindh Province	-0.295	-0.298	-0.300		-0.542	
	(0.323)	(0.326)	(0.327)		(0.441)	
Dummy for KPK Province	-0.498	-0.502	-0.506		-0.782	
	(0.886)	(0.882)	(0.880)		(0.797)	
Father's Years of Education	-0.013	-0.013	-0.012	-0.030	-0.005	-0.019
	(0.024)	(0.024)	(0.024)	(0.024)	(0.025)	(0.025)
Father's Years of Education*Sindh	0.003	0.003	0.001	0.024	-0.015	-0.003
	(0.033)	(0.033)	(0.033)	(0.032)	(0.040)	(0.037)
Father's Years of Education*KPK	0.035	0.0351	0.034	0.004	0.017	-0.028
	(0.087)	(0.087)	(0.086)	(0.056)	(0.075)	(0.049)
District Fixed Effects	No	No	No	Yes	No	Yes
Test statistics						
Hansen J statistic					0.926	2.21
p-value					0.336	0.136
F-statistic, Instruments					7.85	5.55
F-probability					0.001	0.005
CLR Test					3.55	2.34
p-value					0.074	0.148
95% Confidence interval based on CLR Test					[-3.33, -0.02]	[-9.39, 0.30]
Number of observations	1285	1285	1285	1285	1285	1285

Notes: Village-level cluster robust standard errors are reported in parenthesis. Columns (4) and (6) control for factors related to village location with district fixed effects. The Hansen J statistic tests for over-identification. We report cluster-corrected F statistics to test for weak instruments and then implement the conditional likelihood ratio (CLR) test developed by Moreira (2003) to test for weak instrument bias. The CLR test is a test of the hypothesis that the instruments and the coefficient on the endogenous variable are jointly zero. We next show the 95% confidence interval suggested by the CLR test. As implemented, CLR test is generalized for clustered dependence in error terms using the minimum distance approach by Finlay and Magnusson (2009).

Table 5. Parental Consanguinity and Likelihood to be Moderately Stunted, Age 5-13

	Ordinary Least Squares				IV-LIML Estimator	
	(1)	(2)	(3)	(4)	(5)	(6)
Parents are Related	-0.006			-0.007	0.267	0.225
	(0.049)			(0.049)	(0.177)	(0.249)
F-coefficient of Inbreeding		0.109				
		(0.387)				
Parents are 1st Cousins			0.016			
			(0.049)			
Parents are 2nd Cousins			0.038			
			(0.069)			
Age	0.028***	0.028***	0.028***	0.027***	0.025***	0.024***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)
Male	-0.053	-0.053	-0.053	-0.049	-0.057	-0.051
	(0.034)	(0.034)	(0.034)	(0.033)	(0.036)	(0.035)
Wealth Index	0.000	0.002	0.002	-0.008	0.022	-0.008
	(0.030)	(0.030)	(0.030)	(0.028)	(0.033)	(0.027)
Dummy for Sindh Province	0.064	0.067	0.067		0.114	
	(0.074)	(0.074)	(0.075)		(0.097)	
Dummy for KPK Province	0.128	0.132	0.133		0.180	
	(0.181)	(0.180)	(0.179)		(0.163)	
Father's Years of Education	0.004	0.004	0.004	0.008	0.002	0.006
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Father's Years of Education*Sindh	-0.008	-0.008	-0.008	-0.014*	-0.005	-0.010
	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)
Father's Years of Education*KPK	-0.011	-0.011	-0.011	-0.006	-0.007	-0.001
	(0.019)	(0.019)	(0.019)	(0.012)	(0.017)	(0.013)
District Fixed Effects	No	No	No	Yes	No	Yes
Test statistics						
Hansen J statistic					0.227	1.35
p-value					0.634	0.244
F-statistic, Instruments					7.85	5.55
F-probability					0.001	0.005
CLR Test					2.07	0.99
p-value					0.171	0.351
95% Confidence interval based on CLR Test					[0.01, 0.61]	[0.01, 2.73]
Number of observations	1285	1285	1285	1285	1285	1285

Notes: Village-level cluster robust standard errors are reported in parenthesis. Columns (4) and (6) control for factors related to village location with district fixed effects. The Hansen J statistic tests for over-identification. We report cluster-corrected F statistics to test for weak instruments and then implement the conditional likelihood ratio (CLR) test developed by Moreira (2003) to test for weak instrument bias. The CLR test is a test of the hypothesis that the instruments and the coefficient on the endogenous variable are jointly zero. We next show the 95% confidence interval suggested by the CLR test. As implemented, CLR test is generalized for clustered dependence in error terms using the minimum distance approach by Finlay and Magnusson (2009).

Table 6. Parental Consanguinity and Likelihood of being Severely Stunted, Age 5-13

	Ordinary Least Squares				IV-LIML Estimator	
	(1)	(2)	(3)	(4)	(5)	(6)
Parents are Related	0.026			0.035	0.356*	0.401*
	(0.046)			(0.047)	(0.198)	(0.239)
F-coefficient of Inbreeding		0.109				
		(0.372)				
Parents are 1st Cousins			0.018			
			(0.047)			
Parents are 2nd Cousins			0.062			
			(0.070)			
Age	0.021***	0.022**	0.022***	0.020***	0.018***	0.017***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
Male	-0.014	-0.014	-0.014	-0.012	-0.017	-0.016
	(0.026)	(0.026)	(0.026)	(0.025)	(0.030)	(0.029)
Wealth Index	-0.007	-0.008	-0.007	-0.019	0.019	-0.019
	(0.022)	(0.023)	(0.023)	(0.020)	(0.029)	(0.020)
Dummy for Sindh province	0.084	0.080	0.083		0.144	
	(0.062)	(0.062)	(0.063)		(0.096)	
Dummy for KPK province	0.230	0.229	0.237		0.300*	
	(0.174)	(0.175)	(0.175)		(0.157)	
Father's Years of Education	0.002	0.002	-0.005	0.007	0.000	0.004
	(0.005)	(0.005)	(0.009)	(0.005)	(0.005)	(0.005)
Father's Years of Education*Sindh	-0.003	-0.003	-0.005	-0.009	0.001	-0.002
	(0.007)	(0.007)	(0.008)	(0.006)	(0.010)	(0.008)
Father's Years of Education*KPK	-0.014	-0.014	-0.015	-0.008	-0.010	-0.001
	(0.017)	(0.018)	(0.018)	(0.011)	(0.014)	(0.011)
District Fixed Effects	No	No	No	Yes	No	Yes
Test statistics						
Hansen J statistic					2.40	2.67
p-value					0.121	0.102
F-statistic, Instruments					7.85	5.55
F-probability					0.001	0.005
CLR Test					4.07	6.53
p-value					0.056	0.016
95% Confidence interval based on CLR Test					[0.01, 0.91]	[0.30, 2.42]
Number of observations	1285	1285	1285	1285	1285	1285

Notes: Village-level cluster robust standard errors are reported in parenthesis. Columns (4) and (6) control for factors related to village location with district fixed effects. The Hansen J statistic tests for over-identification. We report cluster-corrected F statistics to test for weak instruments and then implement the conditional likelihood ratio (CLR) test developed by Moreira (2003) to test for weak instrument bias. The CLR test tests the hypothesis that the instruments and the coefficient on the endogenous variable are jointly zero. We also show the 95% confidence interval suggested by the CLR test. The CLR test is generalized for clustered dependence in error terms using the minimum distance approach by Finlay and Magnusson (2009).

Appendix Tables

Table A.1. Correlation of the household wealth index with self-reported shocks that hit the household in the past 12 months.

	Type of shock	Coefficient	p-value
1	Flood	0.391	0.011**
2	Loss of Harvest	0.147	0.460
3	Lack of irrigation water	-0.009	0.927
4	Livestock epidemic / Stolen livestock	0.069	0.648
5	Damage to agricultural land due to flood	0.211	0.203
6	Expecting a job but couldn't find employment		
7	Lowered income of any member	0.052	0.776
8	Fall in prices of products in the household business (including agriculture)	-0.096	0.615
9	Compensation for fire/ accidents (insurance)	0.065	0.806
10	Unusual increase in food prices	0.023	0.322
11	Unusual increase in rent	-0.069	0.715
12	Unusual increase in other prices	-0.034	0.764
13	Unusual increase in other prices	-0.007	0.957
14	Criminal act or theft	-0.138	0.634
15	Land dispute	0.170	0.462
16	Conflict	0.322	0.36
17	Major cost for marriage	-0.12	0.339
18	Births	-0.405	0.087*
19	Death of other members of the household	-0.297	0.335
20	Illness / Accident / Disability of main earning member of the household	0.001	0.992

Note. The table reports the OLS coefficient and corresponding p-value of separate regressions of our wealth index (dependent variable) on a dummy variable indicating whether the household has been hit by the corresponding shock in the last 12 months (dependent variables). All regressions also control for district fixed effects.

Table A2. Correlation between early grand parental death with parents being relatives

Explanatory variable	Dependent variable: Whether parents are related (1)	(2)	(3)	(4)
Maternal grand-father died before 65	-0.172***			
	(0.050)			
Paternal grand-father died before 65		-0.088*		
		(0.050)		
Maternal grand-mother died before 65			-0.018*	
			(0.048)	
Paternal grand-mother died before 65				-0.071
				(0.044)
N. of observations	1,411	1,411	1,411	1,411

Note. Reported OLS coefficients are estimated from separate unconditional regressions of a dummy for whether parents are related on a dummy for whether the corresponding grand-parent died before age 65.

Table A.3. Timing of death of the bride's father and percentage of consanguineous unions

	Maternal grandfather died before age 65	Maternal grandfather died between age 65 and 74	Maternal grandfather died age 75 or older	Maternal grandfather is still alive
Share married to a relative	0.579	0.634	0.655	0.679
Number of observations	413	339	252	857

Table A.4. IV-LIML estimates for the effect of consanguineous unions on children's outcomes, among parents that are aware of the detrimental effects of consanguinity compared to those that are not

	Dependent variable					
	Raven's test score		Height for age		Severely stunted	
	Parents are aware	Parents are unaware	Parents are aware	Parents are unaware	Parents are aware	Parents are unaware
	(1)	(2)	(3)	(4)	(5)	(6)
Parents are related	-0.025	-0.784	-0.016	-0.967	0.087	0.981*
	(0.283)	(0.691)	(1.53)	(2.64)	(0.201)	(0.544)
N. of observations	182	593	182	593	182	593

Note. The table report the IV-LIML coefficients of having related parents, instrumented by grand-parental land ownership and wealth, on the corresponding child outcomes. All regressions control for district fixed-effects and a set of covariates. Standard errors are clustered at the village level.