DOES THE ROTTEN CHILD SPOIL HIS COMPANION?
SPATIAL PEER EFFECTS AMONG CHILDREN IN RURAL INDIA[*]

Christian HELMERS[a,b,c]   Manasa PATNAM[d]

[a] *Universidad Carlos III de Madrid*

[b] *CSAE, University of Oxford*

[c] *SERC LSE*

[d] *University of Cambridge*

August 2011

ABSTRACT

This paper identifies the effect of neighborhood peer groups on childhood skill acquisition using observational data. We incorporate spatial peer interaction, defined as a child's nearest geographical neighbors, into a production function of child cognitive development in Andhra Pradesh, India. Our peer group definition takes the form of networks, whose structure allows us to separately identify endogenous peer effects and contextual effects. We exploit variation over time to avoid confounding correlated with social effects. Our results suggest that spatial peer and neighborhood effects are strongly positively associated with a child's cognitive skill formation. Further, we find that the presence of peer groups helps provide insurance against the negative impact of idiosyncratic shocks to child learning. We show that peer effects are robust to different specifications of peer interactions and investigate the sensitivity of our estimates to potential mis-specification of the network structure using Monte Carlo experiments.

KEYWORDS: Children, peer effects, cognitive skills, India

JEL Classification: C21, O15, R23

---

1 Introduction

There is a sizeable economic literature on cognitive and non-cognitive skill formation of children (Todd and Wolpin, 2003, 2007; Cunha and Heckman, 2007, 2008). Using a production function framework, this literature investigates the determinants of the creation of a child's cognitive and non-cognitive skills. The most recent advances in this literature attribute important roles to self-productivity and cross-productivity of cognitive and non-cognitive skills as well as to parental investment (Cunha and Heckman, 2008).[1]

However, so far, the literature on skill formation has treated children in isolation, assuming that they are not directly influenced by their peers.[2] At the same time, there is an important literature on peer and neighborhood effects (Borjas, 1995; Becker, 1996; Hoxby, 2000; Sacerdote, 2001; Gaviria and Raphael, 2001; Hanushek et al., 2003; Durlauf, 2004; Lin, 2010). In this literature, individual outcomes are influenced by spatial instead of market interaction of individuals, i.e., the probability of observing an individual behaving in a certain way is a function of either some characteristics of the individual's environment (neighborhood effects) or directly the prevalence of this type of behavior among her peers (peer effects).[3]

Durlauf (2004) lists three specific channels through which such neighborhood and peer effects are mediated. First, psychological factors can stir a child's desire to behave like others (e.g., purely imitative behavior), second, interdependencies in the constraints children face motivate similar behavior because the costs associated with a given behavior depend on whether others behave in the same way (e.g., reduction of stigma arising from deviant behavior), and third, behavior of other children may change the information on the effects of such behavior available to a child (e.g., expected income from an additional year of schooling). Intuitively, all channels depend on the existence of contact between individuals. The probability of contact and its intensity may be a function of geographical distance between individuals, family or friendship ties etc. Independently of the channel, in the presence of peer effects, children are directly influenced by actions and characteristics of their peers. Therefore, peer-effects may be an important determinant of a child's development of cognitive and non-cognitive skills.

The identification of peer effects encounters well known problems laid out in Manski (1993). Manski lists three effects that need to be distinguished in the analysis of peer effects. The first type are endogenous effects which arise from an individual's propensity to behave in some way as a function of the behavior of the group. The second are so-called contextual effects which represent the propensity of an individual to behave in some way as a function of the exogenous characteristics of his peer group. The third type are so-called correlated effects which describe circumstances in which individuals in the same group tend to behave similarly because they have similar individual characteristics or face similar institutional arrangements, i.e., children within the same village may behave similarly. This means that there are unobservables in a group which may have a direct effect on observed outcomes, i.e., disturbances may be correlated across individuals in a group. The main

---

[1]Self-productivity refers to any effect of past periods' cognitive/noncognitive skills on the current period's cognitive/non-cognitive skills respectively, while cross-productivity refers to any effect of past periods' cognitive/non-cognitive skills on current period non-cognitive/cognitive skills.

[2]Although some of this peer influence is absorbed by school- and household-level controls commonly included in the specification of the production function.

[3]Our definitional distinction between neighborhood and peer effects is somewhat arbitrary as Ioannides (2008) notes in his article for the *New Palgrave Dictionary of Economics* that '[t]erms like social interactions, neighbourhood effects, social capital and peer effects are often used as synonyms'.

empirical challenges, therefore, consist in (1) disentangling *contextual* effects, i.e., the influence of exogenous peer characteristics on a child's observed outcome, and *endogenous* effects, i.e., the influence of peer outcomes on a child's outcome, and (2) distinguishing between *social* effects, i.e., exogenous and endogenous effects, and *correlated* effects, i.e., children in the same peer group may behave similarly because they are alike or share a common environment. Such correlated effects can also include sorting of households, i.e., the endogenous location choice by households. The identification problem explains why existing work looking at children's and teenagers' cognitive outcomes incorporating neighborhood effects, such as Brooks-Gunn et al. (1993), McCulloch and Joshi (2001), and Ainsworth (2002), only accounts for contextual effects and assumes the absence of endogenous effects.

In this paper, we use observational data from the Young Lives (YL) project for Andhra Pradesh, India, to examine neighborhood-level peer influences on child cognitive development by estimating a production function of a child's cognitive ability accounting for endogenous and contextual peer effects. We regard our empirical specification, which explicitly allows children to be influenced by and learn from their peers, as a step forward towards a more realistic model of skill formation.

A common justification for neglecting peer effects in the analysis of child skill formation in the existing literature is the lack of appropriate data. The most commonly used data set in this line of research, the US National Longitudinal Survey of Youth (NLSY), is the result of stratified sampling which justifies the assumption of independence of children within the data set. Even if information revealing the identity of a child's peers were available in the survey data, these peers would most likely not have been included in the sample. The nature of the available data, therefore, severely limits the ability to investigate the potential impact of peer effects on skill formation.

In principle, the same applies to the YL data used in our analysis, which justified treating children as independent units in earlier work. However, we show that in our data, surveyed households are located in close geographical proximity within villages due to the small overall size of the surveyed rural villages in Andhra Pradesh. The presence of this spatial pattern, i.e., close geographical proximity of surveyed households within villages, allows us to employ geographical proximity between children to identify spatial peer effects on child cognitive outcomes. Our main identifying assumption is therefore that peer effects arise through geographical proximity between children: children that live next door to each other are more likely to interact and influence each other than children at the other ends of the village. We thus construct a child's peer group based on geographical proximity of other similar aged children *within* the same village using GIS location data.[4] The resulting structure of peer groups enables us to disentangle contextual and endogenous effects despite the lack of experimental data. We use variation in our data over time to avoid confounding social with correlated effects. For estimation, we rely on recent advances in the spatial econometric literature using the spatial nonparametric heteroscedasticity and autocorrelation consistent (SHAC) instrumental variable estimator proposed by Kelejian and Prucha (2007). The SHAC approach allows us to remain agnostic regarding the structure of the spatial dependence in the residuals and to allow for heteroscedasticity of any arbitrary form.

Our results suggest that a child's geographical neighbors are positively associated in a statistically significant and economically important way with a child's own production of cognitive skills between

---

[4]Children in our sample are of the same age, which means that peer influences are reciprocal and contemporaneous which is distinct from role-model influences which could emerge if younger children imitate behavior of older individuals (Durlauf, 2004).

age eight and twelve. Furthermore, we find that contextual effects appear to have little influence on cognitive achievement gains.

The fact that we observe only a sample of children within villages, in principle, does not undermine identification of peer effects because children have not been selected into the sample based on their location within villages. However, sampling means that we only observe a fraction of a child's true peer network, which implies that our assumed network structure might be mis-specified. Regardless of sampling, even if the population of children were observed within villages, mis-specification of peer links between children might still occur given the lack of information on actual peer interaction. We investigate the implications of the potential mis-specification of our proximity-based peer network structure and demonstrate the robustness of spatial peer estimates to different forms of mis-specification both due to sampling and different assumptions about the true population peer network structure. Nevertheless, we only have observational data for our analysis, which means that despite our identification strategy and robustness checks, we are cautious in attributing our results a causal interpretation.

As an additional contribution, we use the augmented skill production function to examine the relevance of peer groups in assisting children recover from shocks. We investigate whether the presence of a peer group helps insure children of shock-affected households against an adverse impact to their cognitive achievement gain. Many studies have found that economic, health or climatic shocks to a household have a negative impact on child schooling and health. This is because in such circumstances households will typically tend to under-invest in education or health related expenditure of their children. While a considerable amount of the literature is devoted to examining risk-sharing and informal insurance arrangements of households (Townsend, 1994; Gertler and Gruber, 2002; De Weerdt and Dercon, 2006), there is little evidence to show how children from such households find support to cope with adversities that significantly compromise investment in their education and health (Tominey, 2009; Ginja, 2010). By utilizing detailed data on idiosyncratic household related shocks, we find that the negative effect of a shock on child cognitive achievement becomes insignificant after incorporating peer effects, which we interpret as evidence that peer groups provide partial insurance. Moreover, our analysis suggests that this peer insurance effect applies only to boys.

Our results contribute to the empirical literature on childhood skill formation by providing evidence for the presence and importance of peer and neighborhood effects in the formation of children's cognitive skills in a rural developing country context. Moreover, we contribute to the existing peer effects literature by providing an example of how to identify endogenous and contextual peer effects without the need for data from a controlled randomized experiment by using GPS location data which are routinely collected in household surveys. Our research design may therefore be applicable in any context in which peer effects are mediated through spatial proximity and location data are available. It can be applied to study peer effects in other contexts both on other outcome variables and populations of interest.

From a policy perspective, understanding the role of social interactions and peer effects in shaping childhood skill formation is important as policy interventions targeting only a subset of children of a population may influence outcomes of other children not directly included in the intervention.[5] Because of the bi-directional nature of peer effects, their presence also implies social multiplier effects which magnify the impact of policy interventions (Manski, 1993; Bobonis and Finan, 2008).

---

[5] See Fafchamps and Vicente (2009) for evidence of such 'diffusion' effects in the context of political awareness campaigns.

As noted by Durlauf (2004), peer effects can also lead to persistence in poverty as neighborhoods can get locked in bad equilibria which are enforced over time by the mutually reinforcing character of peer effects. Therefore, improved understanding of the role of peer interaction, in particular in a developing country context, may contribute to the design of novel interventions aimed at improving children's cognitive skill production and thus success in later life.

The paper is organized as follows: Section 2 discusses our identification strategy which is translated into the specification of our empirical model presented in the same Section. The SHAC estimator used in our analysis is presented in Section 3. The data used is described in Section 4. Section 5 discusses our results and reports our robustness checks; Section 6 concludes.

2   Identification of Peer Effects

There are two main challenges to the identification of peer effects: (1) the separate identification of endogenous and contextual effects and (2) separating social effects, i.e., endogenous and contextual effects, from correlated effects. Such correlated effects subsume a range of potential unobservables common to individuals within peer groups that are correlated with the dependent and independent variables included in the analysis. This may also include the potentially endogenous sorting of households into geographical locations, i.e., the endogenous formation of peer groups. We show in this section how we address these two *conceptually distinct* identification issues with the available data which consists of observations on individual children at two points in time and a social network structure that remains unchanged over time.

2.1   Identifying Endogenous Peer Interaction Effects

Our analysis of the YL data for Andhra Pradesh, India, reveals close geographical proximity of the surveyed households within towns and villages. As an example, Figure 1 shows the map of a sample village in Andhra Pradesh. The figure suggests that groups of households are located close to each other within the village[6]. In fact, the median distance between households within the networks used to define peer interaction, which we will discuss further below, is 126 meters.[7] This short distance is striking in light of the fact that most households are located in rural areas.

The spatial proximity of households allows us to identify surveyed households' geographical neighbors which we use to define each child's peers. The clustering is important, because the close geographical proximity of households allows us to reasonably argue that households interact as neighbors. From the YL child-level questionnaire at age twelve, we have some information on how children spend their time. We know for example, that the median amount of time that children spend playing with their peers is four hours a day, which leaves ample room for neighborhood-based peer interaction.

Note that the spatial clustering emerges randomly. Our data is a random sample of households *within* villages, i.e., households have not been selected into the sample based on their location within villages (see also Section 4). Hence, we can assume that the observed spatial distribution of observed households is representative of the true underlying spatial distribution of households in the

---

[6]Due to data confidentiality agreements, we are unable to disclose the exact locations of sample households or village. Instead we report detailed statistics on the proximity of sample households throughout the paper.

[7]The average distance within networks is 447 meters. These figures refer to networks defined as a child's five nearest neighbors.

5

population (more discussion is provided in Section 2.3). This allows us to use a neighborhood-based definition of social interactions. Adding the fact that only children of the same age are included in the sample, we are able to construct our measure of peer effects based on nearest neighbor networks which represent a child's peer group. These networks are used to analyze how neighborhood related spatial peer group effects affect child related outcomes. We assume that interactions between children occur exclusively through social interactions and are thus unrelated to market interactions, which appears to be a reasonable assumption in our setting.

The identification of peer effects is notoriously difficult as explained by Manski (1993) and Moffitt (2001) (for a summary of the literature see also Blume and Durlauf, 2006). Manksi noted that within a linear framework without additional information, it is impossible to infer from the observed mean distribution of a sample whether average behavior within a group affects the individual behavior of members of that group. In other words, the expected mean outcome of a peer group and its mean characteristics are perfectly collinear due to the simultaneity induced by social interaction. This fundamental identification problem, termed *reflection problem* by Manski, makes it clear that within a linear-in-means model, identification of peer effects depends on the functional relationship in the population between the variables characterizing peer groups and those directly affecting group outcomes.

Lee (2007) was first to show formally that the spatial autoregressive model specification (SAR), widely used in the spatial econometrics literature, can be used to disentangle endogenous and exogenous effects. In a SAR model, identification of endogenous and contextual effects is possible if there is sufficient variation in the size of peer groups within the sample. As stressed by Davezies et al. (2006), Lee's identification strategy crucially requires knowledge of peer group sizes and at least three groups of different size. Bramoullé, Djebbari, and Fortin (2009) [BDF henceforth] propose an encompassing framework in which Manski's mean regression function and Lee's SAR specification arise as special cases. BDF show that endogenous and exogenous effects can be distinguished through a specific network structure, for example the presence of intransitive triads within a network. Intransitive triads describe a structure in which individual i interacts with individual j but not with individual k whereas j and k interact.[8]

We denote the set of children as i ($i = 1, ..., n$) and y_{it} denotes the cognitive achievement of child i at period t, $\mathbf{x}_{it}$ is a $1 \times K$ vector of child and household characteristics. Each child has a peer group P_i of size n_i. By assumption child i is excluded from P_i. Denoting each network as l, we assume that our sample of size l is i.i.d. and from a population of networks with a fixed and known structure.[9] The assumption of a fixed network structure is made on the basis that networks are defined according to the location of the households in which children live. Since households in our sample do not move during the two observed time periods, the network structure is fixed (see also Section 4).[10] We distinguish between three types of effects: a child's outcome y_i is affected by (i) the mean outcome of her peer group (endogenous effects), (ii) her own characteristics, and (iii) the mean characteristics of her peer group (contextual effects):

---

[8]This particular network structure produces exclusion restrictions which achieve identification in the same way as exclusion restrictions achieve identification in a system of simultaneous equations.

[9]In our case, each village is representative of one inter-connected network, l.

[10]Generally, sample attrition in the YL data for Andhra Pradesh is very low with 1.29% between the two survey rounds for the 'older' cohort of children used in our analysis (Outes-Leon and Dercon, 2008).

$$y_{lit} = \beta \frac{\sum_{j \in P_i} y_{ljt}}{n_i} + \gamma x_{lit} + \delta \frac{\sum_{j \in P_i} x_{ljt}}{n_i} + \varepsilon_{li} + \varepsilon_{lt} + u_{lit} \qquad (1)$$

Hence, β captures endogenous effects and δ contextual effects. Correlated effects are represented by ε_{li} and ε_{lt}. We require strict exogeneity of $\mathbf{x}_{lt}$ with respect to u_{lit}. Note that we do not require the residuals u_{lit} to be homoscedastic or normally distributed.

To estimate Equation (1), we construct a neighbor matrix (alternatively interpreted as a peer interaction matrix), W, which is interacted with the outcome variable and exogenous peer characteristics to form spatial lags. We define W using a 'K Nearest Neighbors' (KNN) characterization. KNN is a distance-based definition of neighbors where 'K' refers to the number of neighbors of a location. Distances are computed by the Euclidean distance between GPS locations of households. Therefore, under this approach, the set of 'neighbors' for child/household i includes the K children/households characterized by the shortest distance to child/household i within each village. We set $K = 5$.[11] The choice of K poses a problem similar to the well-known modifiable areal unit problem in spatial econometrics (Openshaw, 1984). We are thus careful to check the robustness of our results to modifications in the definition of K (see Section 5.3).[12] However, using this method, we drop households that are not a nearest neighbor to any other household in the sample.[13] Depending on the number of nearest neighbors used in our definition of W, this leads us to drop a small number of households which causes slight variations in the sample size across specifications (see Section 5). Given that the households are a random sample of the underlying population, dropping such 'island' households should not bias our results.

Alternatively, we can construct the peer reference group as all children of the same age belonging to the same community. The definition of a community in the Young Lives survey pertains to a geographically well defined administrative area, such as a zone/neighbourhood in (semi-)urban and a village in rural areas. As noted by Lee (2007), peer effects are still identified since children interact in community based groups of different sizes. The peer/neighbourhood interaction matrix, W, has block diagonal elements of varying sizes. This brings about variation in reduced-form coefficients across communities of different size that ensures identification. This alternative definition captures both peer and neighborhood effects and offers a less restrictive way of specifying the structure of underlying peer interaction as it avoids any assumptions on the number and direction of peer links.

Manski noted that the 'informed specification of reference groups is a necessary prelude to analysis of social effects' (Manski, 1993: 536). In our analysis, the specification of a child's peer group arises naturally as we assume that children are limited to interaction within a geographically con-

---

[11] The child-level questionnaire at age twelve asks children to indicate the number of friends they have. On average, children report to have 7 friends, with a standard deviation of 4.8. During the first sampling round, the prevalence of 1-year-old children in the population was estimated to be 2%. This meant only villages with at least 5,000 inhabitants were selected among the sample villages to ensure the target sample size of 100 children for the 'younger' cohort was met (Himaz et al., 2008). Considering child (1-5 years) mortality of 21% (Indian National Family Health Surveys 1998-1999), the prevalence of 8-year old children in the population was most likely lower at around 1.6%. This means that the 50 sampled children of the 'older' cohort represent up to 60% of the population of 8-year old children in each sampled sentinel site (see also Section 4). This implies that a five nearest neighbor specification appears appropriate to capture the share of each child's friends included in the sample. Nevertheless, for robustness, we also estimate our model using a network definition based on three and seven nearest neighbors (see Section 5.3).

[12] In unreported results, we also allow K to vary by including all neighbors in the set of KNN that reside within a specific distance band from a child's household. Using this specification, in which links are un-directed, we find similar results. These tables are available from the authors on request.

[13] In fact, for such 'island' households, column sums of the spatial weight matrix W are zero. This occurs as a result of specifying a directed network structure, i.e., while any node has five nearest neighbors, this does not automatically imply the node itself represents one out of five nearest neighbors to any other node in the network.

fined area, their village and more specifically their nearest neighbors within that village. The assumption appears defendable on the grounds that the sample consists of children aged 8-12 who are arguably limited in their independent movements beyond their close environment. The construction of the weights matrix allows us thus to rewrite Equation (1) in structural form as (omitting time subscripts and subsuming correlated effects ε in u):

$$y_l = \beta W y_l + \gamma x_l + \delta W x_l + u_l \tag{2}$$

This implies, that the reduced form is given by;

$$y_l = (I - \beta W)^{-1}(\gamma I + \delta W)x_l + (I - \beta W)^{-1}u_l \tag{3}$$

If we omitted the endogenous effects ($W y_l$) from Equation (2), the model could be estimated using OLS under the assumption that all covariates are independent of the error term, i.e., strictly exogenous. However, OLS is biased and inconsistent in the presence of a spatial autoregressive lag (Anselin, 1988). Denoting the variance-covariance matrix of u_l as ψ_{u_l}, it is easy to see that,

$$E[(W y_l)u_l^{'}] = W(I - \beta W)^{-1}\psi_{u_l} \neq 0 \tag{4}$$

Anselin (1988) suggested an approach based on a Maximum Likelihood (ML) estimator to address the endogeneity problem. To avoid computation accuracy problems in the ML approach noted by Prucha and Kelejian (1999), Kelejian and Prucha (1997, 1998) suggested a spatial two-stage least squares estimator (S2SLS). They suggest using a set of instrument matrices to instrument for $(x_l, W x_l, W y_l)$. From Equation (4), we can see that, ideally the set of instruments contains linearly independent columns of $[x_l, W x_l, W^2 x_l]$. Hence, identification of endogenous and contextual effects is possible if I, W and W^2 are linearly independent.

We use the network structure of our peer reference group to ensure this condition is met. This is the case when the network is characterized by (a small degree of) intransitivity e.g., child i and child j are nearest neighbors, child j and child k are nearest neighbors, but child i and child k are *not* nearest neighbors. This produces a network topology which achieves identification of peer effects as shown by BDF. The networks-based intuition of this strategy is straightforward: $W^2 x_l$ is an identifying instrument for $W y_l$, since x_{kl} affects y_{jl} (since they are connected and interact with each other) but x_{kl} can only affect y_{il} indirectly, through its effect on y_{jl}. Therefore, given our peer network structure, $[x_l, W x_l, W^2 x_l]$ are valid and informative instruments for endogenous peer effects $W y_l$.

To illustrate how our peer network structure achieves identification, Figure 2 shows a fragment of the networks spanned by children in our data set. For example, while child pairs 8,9 and 11,9 are nearest neighbors to each other, child 8 and 11 are not nearest neighbors. Hence, child 11 influences child 8 only through child 9. Similar reasoning applies to the other networks displayed in Figure 2.

2.2   Correlated and Selection Effects

Correlated effects occur when individuals within a peer group behave similarly due to the common environment that they face. This problem may arise in our setting, for example, if children attend the same school in the village or are subject to the same macro-shocks. Selection effects, which can be subsumed under correlated effects, arise when an individual chooses his own peer/reference group, i.e., individuals have not been assigned randomly into peer groups; this causes a bias in the

peer interaction effect due to the presence of unobservables that both influence the choice of peer group and outcome. Group formation is endogenous, for example, when popular students interact primarily with other popular students or when households sort themselves into a locality of their choice. In our case, a possible concern is that households sort into neighborhoods based on factors influencing their children's cognitive development. While in the deprived rural setting of Andhra Pradesh, it seems rather unlikely that parents choose the location of a household based on school characteristics, there may still be other location-specific characteristics that attract households that for example attach greater importance to the education of their children.

Several solutions to the problem of correlated effects have been proposed in the literature. Katz, Kling and Liebman (2001) analyze the influence of a household's neighborhood on individual and household outcomes by looking at the impact of moving from urban public housing to suburban wealthier neighborhoods in Boston. The problem of endogenous sorting among neighborhoods is addressed by the random assignment among families of public support to move between neighborhoods to infer the average causal relationship between moving and outcomes. Evans, Oates, and Schwab (1992) take a more structural approach by using a system of simultaneous equations to account for endogenous sorting because their data is not generated from a randomized intervention. In particular, the comparison of the results obtained from a 'naive' and the structural regressions lends substantial support to the importance of individuals' endogenously sorting into neighborhoods. Gaviria and Raphael (2001) look at the importance of peer interaction in determining teenagers' propensity to engage in deviant behavior such as drug use and alcohol drinking in US schools. In order to address endogenous sorting of families into peer groups, Gaviria and Raphael (2001) differentiate between families that have recently moved to a neighborhood from families that have resided in the neighborhood for a long time. The authors argue that the endogeneity problem is more severe for families that have recently moved into the neighborhood. Sacerdote (2001) addresses the problem of endogenous sorting by using the random assignment of roommates at Dartmouth College to identify peer effects.[14] Yet, while random assignment into peer groups avoids the selection effect, members of a peer group may still be exposed to the same environment and (time-variant) common shocks (e.g. roommates may be exposed to the same noisy street). This means that even in the presence of random peer group formation, that is data from a randomized experiment, correlated effects represent a challenge to the identification of peer effects.

In this paper, following Blume and Durlauf (2006), we employ primarily a first-differenced specification to address the issue of correlated and selection effects. We employ differences between the two available rounds of data to account for unobservables that are constant over time. This means, we explain the change in cognitive skill levels achieved by children between t and $t-1$. We write the change in a child's cognitive skills as a function of the change in a child's own characteristics, parental investment and household characteristics. We allow for peer effects by incorporating spatial lag terms of the dependent variable as well as of a child's peers' characteristics, parental investment and household characteristics. Hence, we rewrite Equation (1) as

$$\triangle y_{li} = \beta \frac{\sum_{j \in P_i} \triangle y_{lj}}{n_i} + \gamma \triangle x_{li} + \delta \frac{\sum_{j \in P_i} \triangle x_{lj}}{n_i} + \triangle u_{li} \tag{5}$$

where $\triangle y_{li} = y_{li,t} - y_{li,t-1}$ denotes the difference in cognitive skill levels between periods t and $t-1$ for child i in network l. $\frac{\sum_{j \in P_i} \triangle y_{lj}}{n_i}$ denotes the spatial autoregressive term and $\triangle x_{li} =$

---

[14]Our brief review covers only some selected papers. For a more general overview we direct the reader to a comprehensive recent review of the social interactions literature by Epple and Romano (2011).

$x_{li,t} - x_{li,t-1}$ denotes the change in child i's own characteristics including parental investment and household characteristics while $\frac{\sum_{j \in P_i} \triangle x_{lj,t}}{n_i}$ denotes the change in child i's peers' characteristics between t and $t-1$. This can be easily seen in terms of the network specification,

$$\triangle y_l = \beta W \triangle y_l + \gamma \triangle x_l + \delta W \triangle x_l + \triangle u_l \tag{6}$$

While we are able to difference out all the child, household and village level fixed effects that are constant over time, correlated effects will still continue to persist if there are common environment related time-varying unobservables that affect both the child's as well as her peer group's outcome. For instance, it is possible that more schools were constructed between the two time periods in a particular village causing growth in education achievement for all children in that village. We address this issue in three ways.

First we explicitly account for village-level changes to education by utilizing information on the older siblings of each child. We are agnostic with regard to whether such time-varying village-level unobservables are due to endogenous sorting or unobserved effects, such as the construction of a new school. We construct a quasi-cohort data set by pooling information on all the older siblings of each child in each village.[15] We restrict the sample to those children who are up to two years older than the reference child in both years. To the extent that school enrollment and continuation is a proxy for student achievement, we calculate the average (highest) grade reached by this subset of older children for each village for both time periods. In order to capture any time-varying village-level effects that could have a direct impact on the education/schooling of children we include this variable (in first difference) in our specification. Our objective is to see whether peer effects still continue to hold even after conditioning on these time-varying, village-level effects. This is incorporated as,

$$\triangle y_l = \beta W \triangle y_l + \gamma \triangle x_l + \delta W \triangle x_l + \theta \triangle \bar{e}_l + \triangle u_l \tag{7}$$

Since we treat each village as one interlinked network (our peer interaction matrix, W, is a block diagonal matrix of the large network of all villages), $\bar{e}_l$ represents the change in average educational attainment of a sample of older children in the village where the assumption is that $\theta \triangle \bar{e}_l = \varepsilon_{lt}$ (see Equation (1)).

Secondly, following BDF, we apply a within transformation at the network-level defined by a child's nearest neighbors ('local' transformation) to account for correlated effects.[16] This involves averaging Equation (6) over all of child i's K nearest neighbors. Using average village-level schooling of older siblings to control for time-varying unobservables that are correlated with the peer effects only accounts for differences across villages. By using the local transformation, we can also account for unobservables that vary by nearest neighbor networks. The structural specification (for expositional simplicity omitting contextual effects) is thus,

$$(I - W) \triangle y_l = \beta (I - W) W \triangle y_l + \gamma (I - W) \triangle x_l + (I - W) \triangle u_l \tag{8}$$

---

[15] Our choice of aggregating information at the village-level is supported by findings such as Bayer et al. (2008) who suggest that endogenous sorting is more likely at broader levels, e.g., the neighborhood-level, than for finer levels of aggregation, in their case the block-level in the Boston metropolitan area.

[16] BDF suggest two ways of accounting for correlated effects: 'local' and 'global' within transformation. In a local transformation the model is written as a deviation from the mean equation of the individual's peers and in a global transformation it is written as a deviation from an individual's network.

Hence, using first differences, we not only wipe out unobserved individual, household and village characteristics that remain constant over time, but also nearest neighbor network-specific correlated effects.

Third, we implement an instrumental variable (IV) approach. We have data on household-specific idiosyncratic shocks (see Section 2.4), which we use as an instrument to predict a child's own cognitive achievement gain. This provides a valid instrument because child i's peers' cognitive achievement gains are not directly affected by the household-specific shock affecting child i, but indirectly through the impact the shock has on child i's cognitive development. We show that the instrument is also highly informative as idiosyncratic shocks are negatively correlated in a statistically significant way with a child's own achievement gains. This IV approach provides an alternative and intuitive way of gauging the effect of correlated effects on our peer effect estimates.

2.3   Mis-specification of the Peer Network Structure

Our identification approach relies on the assumption that social interaction is mediated by geographical proximity: children that are located geographically close to each other within villages are more likely to interact than children geographically afar. Our empirical approach requires us to fully specify the structure of the social interaction between children. An obvious concern with this approach, which we discuss in this section, is the potential mis-specification of the peer network structure, i.e., the spatial weight matrix. We distinguish in our discussion between two related issues that may lead to mis-specification: (1) data missing at random due to random sampling and (2) the unknown population network structure.

2.3.1   Sampling

A widespread problem associated with the empirics of social network data is that of mis-measurement due to sampling. Sampled data on networks are obtained by enumerating links amongst the sample of individuals who are selected from the population either (i) based on the realization of the dependent variable or a variable correlated with the dependent variable or (ii) through some random mechanism. This would lead to measurement error if there are certain units that exist in the population, not represented in the sample, but who are connected with some units included in the sample.

Marsden and Hurlbert (1987) discuss the bias that arises when individuals select into the sample based on their realization of the dependent variable or a variable correlated with the dependent variable. The resulting bias is well-known in the literature on sample selection (Heckman, 1976, 1979). In contrast, when nodes are randomly assigned into the sample, a common assumption (e.g., Hoff, 2007; Taskar et al., 2004) in the estimation of network properties on sampled data is that the missing information on link presence/absence is missing at random. We consider spatial networks where the underlying network structure is based on nearest (spatial) neighbors. In our data, children were sampled randomly and children have not been selected into the sample based on the geographical location of the household they belong to within villages. Hence, it is reasonable to assume that the network data on missing links is missing completely at random because the units in the analysis were sampled at random from the population. In other words the probability of observing a missing link depends only on the probability of two units i and j being observed which is assumed to be random due to sampling. We nevertheless use Monte Carlo experiments to explore more rigorously the potential estimation bias that arises from random sampling in Section 5.3.1.

11

2.3.2 Unknown population network structure

The definition of interactions between children is based on geographical proximity between children. This is an assumption required in the absence of actual information on interaction patterns between children within villages.[17] This implies that our network structure might only imperfectly approximate the true underlying social network structure.

Despite the importance of this issue for identification, the spatial econometrics literature provides hardly any guidance on the magnitude and direction of the bias due to mis-specification of the spatial weights matrix and the circumstances under which it is most likely to affect the estimates. Páez et al. (2008) use Monte Carlo exercises where they vary the level of spatial autocorrelation and network topology to analyze bias in a SAR model from under-specification of the adjacency matrix, i.e., assuming that a given node's degree is smaller than in the true model, and over-specification, i.e., assuming that a given node's degree is larger than in the true model. For the SAR specification, Páez et al. find that bias from under-specification is particularly severe when average degree and/or clustering in a network is low, i.e., the spatial weight matrix is sparse, and true underlying spatial autocorrelation high. Over-specification results in pronounced bias when average degree is low, but clustering high, which means that in networks where there are connected components with few links between components, adding false links results in particularly severe a bias. Lee (2009) derives theoretically the bias arising from mis-specification of the adjacency matrix in a model with spatial lags in the independent variables and provides Monte Carlo results for the SAR model showing that a misspecified spatial weight matrix causes bias in both maximum likelihood and two stage least squares based estimation. His results suggest that the bias from under-specification points downward whereas in the case of over-specification estimates are upward biased. Generally, Lee finds bias from over-specification to be lower than bias from under-specification. In order to investigate potential implications for our results, we report results from similar Monte Carlo experiments in Section 5.3.2 and allow the degree, i.e., the number of nearest neighbors, to vary (see Section 5.3.3). In our Monte Carlo exercise, we allow for both over- and under-specification of the network as well as a mixture of both. Specifically, we simulate random deviations from the true underlying network and do not restrict these simulated networks to be above or below the hypothesized population network. This implies that any random draw from the set of simulated networks could pick up either an over-specification or under-specification.

2.4 Peer Effects and Insurance against Shocks

Having established identification of peer effects, we are also interested in testing whether peer effects can provide insurance against adverse shocks to skill acquisition. In the given context, we have a rather informal way of insurance in mind. Children may intuitively rely more on their peers when their own household is affected by an adverse shock, rather than benefit from support mechanisms involving direct transfers.

To test for insurance, we rewrite our specification in Equation (5) to account for idiosyncratic shocks in period t as

---

[17]Actual information on individuals' networks are rare. Some exceptions are Conley and Udry (2010) who have detailed data on self-reported communication networks of farmers in Ghana and the Add Health database (see Lin, 2010 for a description) that incorporates information on friendship links. However, even when data on self-reported networks are available, the resulting network structure might still be mis-specified due perception bias and other potential reasons for individuals (un)intentionally mis-reporting their social networks.

$$\triangle y_{li} = \phi + \beta \frac{\sum_{j \in P_i} \triangle y_{lj}}{n_i} + \gamma \triangle x_{li} + \delta \frac{\sum_{j \in P_i} \triangle x_{lj}}{n_i} + \eta s_{li} + u_{li} \qquad (9)$$

where s_{li} is a dummy variable indicating whether the child's household has experienced an idiosyncratic shock between the two time periods. A large body of work has attempted to apply a test for full/partial insurance for consumption using various methods. The main challenge has been to identify the coefficient of village/network average consumption. Akin to the literature on peer effects, identification is not straightforward due to the reflection problem as the OLS estimator would mechanically fit the mean. Various methods have been proposed and implemented (Townsend, 1994; Gertler and Gruber, 2002; De Weerdt and Dercon, 2006). A standard IV approach is usually regarded as an appropriate solution (Brock and Durlauf, 2001; Boozer and Cacciola, 2001). Once identified, the test for full/partial insurance is applied by introducing the idiosyncratic shock as an overspecification of the model, i.e., $\hat{\eta} = 0$ is interpreted as evidence in favor of peer group insurance. In our case, *having established identification* on our peer effects variable, the test is to see whether conditional on peer effects, the idiosyncratic shock is orthogonal to our measure of cognitive skill formation. The intuition for this is that shocks affecting a child's skill formation through their impact on household resources should not affect skill growth once skill growth across the child's peers is accounted for. As stressed in the relevant literature, the validity of the test for peer group insurance formulated in Equation (9) rests on the assumption that an individual's utility function is separable in the shock and the outcome variable of interest. This means that utility derived from cognitive skill formation must not depend directly on the household's endowment affected by a given shock, i.e., parents' preferences for a child's skill formation must not change as a function of the shock received.

3   Estimation

While the S2SLS estimator proposed by Kelejian and Prucha (1997, 1998) allows consistent estimation of the coefficient associated with the spatial autoregressive term, it requires the residuals to be i.i.d. and homoscedastic. More recently, Kelejian and Prucha (2007) suggested a spatial nonparametric heteroscedasticity and autocorrelation consistent (SHAC) estimator which accommodates heteroscedasticity of unknown form and spatial autocorrelation in the residuals.[18] Contrary to Conley (1999), Kelejian and Prucha's estimator does not require the spatial process to be stationary, which is essential in the context of SAR models as specified in Equation (5) above.[19]

To see how the S2SLS estimator is implemented, rewrite Equation (5) in a drastically simplified form, omitting spatial lags in covariates X for expositional simplicity, as

$$y = Z\zeta + u \qquad (10)$$

where $Z = [Wy, X]$ and $\zeta = \begin{pmatrix} \beta \\ \gamma \end{pmatrix}$. We use the matrix $H = [X, WX, W^2X]$ to instrument for Wy. The estimator is then given by

---

[18]Empirical applications of the estimator are still rare. Exceptions are the papers by Anselin and Lozano-Gracia (2008) and Ben Arfa et al. (2009). Anselin and Lozano-Gracia use the estimator in a hedonic model of house prices in the US whereas Ben Arfa et al. study the geographical distribution of dairy farms in France.

[19]As noted by Kelejian and Prucha (2007), it suffices that cross-sectional units have different numbers of neighbors to obtain a nonstationary spatial process. Considering the discussion in Lee (2007), this may even be a desired feature in the data to be able to disentangle endogenous and contextual effects.

$$\hat{\zeta} = (\hat{Z}Z)^{-1}\hat{Z}'y \qquad (11)$$

where $\hat{Z} = H(H'H)^{-1}H'Z$. Denote the residuals obtained from this S2SLS estimator as u_n. Kelejian and Prucha (2007) assume that the disturbance process can be described as

$$u_n = R_n\epsilon_n \qquad (12)$$

where ϵ_n is a $n \times 1$ vector of innovations and R_n is an $n \times n$ non-stochastic matrix with unknown elements. Note that vectors and matrices are denoted by n as they may depend on the sample size. Kelejian and Prucha (2007) assume that R_n is non-singular and the row and column sums of R_n and R_n^{-1} are bounded uniformly in absolute value by some constant c_R where $0 < c_R < \infty$. The corresponding variance-covariance (VC) matrix is defined as

$$\psi_{ij,n} = n^{-1}H_n'\Sigma_n H_n \qquad (13)$$

where H_n is a $n \times p_h$ non-stochastic matrix of instruments defined above and $\Sigma_n = R_n R_n'$ denotes the VC matrix of u_n where row and column sums of Σ_n are also uniformly bounded. Spatial dependence is introduced through a kernel function which is a real continuous and symmetric function that defines weights for covariances as $K(d^*_{ij,n}/d_n)$ with $d_{ij,n} \geq 0$ and bandwidth $d_n > 0$. Whenever $d_{ij,n} \geq d(n)$, the Kernel is equal to zero. We choose a plug-in bandwidth based on the distance to child i's K nearest neighbors as discussed above.[20] Since the choice of the kernel is usually of little importance in the implementation of nonparametric estimators, we choose the standard Epanechnikov kernel. Then the (r,s)th element of the true VC matrix Ψ_n of the SHAC estimator is given by

$$\hat{\psi}_{rs,n} = n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}h_{ir,n}h_{js,n}\hat{u}_{i,n}\hat{u}_{j,n}K(d^*_{ij,n}/d_n) \qquad (14)$$

Our theoretical justification for the specification of the model in Equation (5) allows us to treat any potential spatial autocorrelation in the error term as nuisance. Hence, we do not need to impose any particular functional form on the spatial autocorrelation in the error term which corroborates our choice of the SHAC estimator.[21]

4 Data

We use data from the India part of the YL project. YL is a long-term study of childhood poverty being carried out in Ethiopia, India (in the state of Andhra Pradesh), Peru and Vietnam. The survey consists of tracking two cohorts of children over a 15-year period. Currently data from two rounds of data collection are available. In Round 1, 2,000 children aged around one (the 'younger' cohort) and 1,000 children aged around eight (the 'older' cohort) were surveyed in 2002. Following up, Round 2 involved tracking the same children and surveying them in 2006 at age five and twelve respectively.

---

[20]Our choice of the plug-in bandwidth seems appropriate given our modeling choice of peer interaction based on a child's K nearest neighbors. As a potential alternative, see Lambert et al. (2008) for a discussion of data-driven bandwidth selection in the context of the SHAC estimator.

[21]To implement the SHAC estimator, we use the *spdep* (Bivand et al., 2008) and *sphet* (Piras, 2010) packages in R.

The sample of children is representative of the three regions of Andhra Pradesh: Rayalseema, Coastal Andhra and Telangana. The sampling process was fourfold. First, six districts were selected based on the classification of poor/non-poor given by their relative levels of development. In the second stage, twenty sentinel sites (mandals) within these districts were identified based on the same classification. Subsequently, one village was randomly selected from approximately four to five villages that comprised a sentinel site and households within the selected village were also selected randomly. Finally the questionnaires were administered to around 100 one-year-old and 50 eight-year-old children in these villages. Data was collected through household questionnaires, child questionnaires and a community questionnaire.

In Helmers and Patnam (2011), we analyzed the formation of both cognitive and non-cognitive skills paying particular attention to self-productivity as well as cross-productivity effects. We have found statistically and economically significant evidence for self-productivity for cognitive skills and cross-productivity effects of cognitive skills on non-cognitive skills. However, we have not found any evidence of self-productivity for non-cognitive skills nor of non-cognitive skills affecting cognitive skills. We therefore focus our analysis of spatial peer effects on the formation of cognitive skills. Moreover we omit the use of many non-cognitive inputs in the production function that the data allow us to use precisely for this reason. Since we are interested in the determinants of the evolution of cognitive skills over time, we can only use the 'older' cohort of children because for the 'younger cohort' there is no information on children's cognitive skill levels at age one. This means we only use information for the 'older' cohort of children to analyze the determinants of their cognitive skill formation between age eight and twelve. For a more detailed description of the data set see Helmers and Patnam (2011). Table 1 shows some summary statistics for the variables used in our analysis.

4.1   Location and Peer Effects

In order to construct geographical distances between households/children, we collated various geography variables from two GIS files: the Taluk map of Andhra Pradesh which provides digitized Taluk (administrative boundary) polygons, and a household location map that contains, as a point feature class, detailed GPS locations of every household/child in the YL data set. The latter was overlaid with the Taluk map to identify village level clusters for households. This gives us longitude and latitude information on the location of households and children and thus allows us to compute the Euclidean distance between households. The distance is used to determine a household's nearest neighbors which are used to measure spatial peer effects. Note that we have GPS locations of only 750 out of the 1,000 sampled households. To assert that there are no systematic differences in characteristics of households/children for which GPS information is and is not available, we implement a number of tests. First, a Kolmogorov-Smirnov test does not reject equality of the outcome distributions with a p-value of 0.322. We also conduct t-tests for differences in means between the sample children and those for whom we have missing GPS information over a range of observable covariates that are included in our empirical specification. The results are reported in Table A-I. Barring a few variables, we find no significant differences in both child and household demographic characteristics.

4.2   Cognitive Skills

In principle, cognitive skills are unobserved. In order to proxy them, we use observed measures. Since we use a specification in first-differences, we need the same measures at age eight and twelve.

This restricts our choice of possible measures because the survey questionnaires differed between Round 1 and Round 2. The only measures for cognitive skills that are available for children at age eight and twelve are reading and writing test scores.[22] These tests assess mostly a child's general intelligence and her ability to apply acquired knowledge and skills. These skills are distinct from non-cognitive skills which aim to measure a child's personality traits (Borghans et al., 2008). Our specification in first-differences accounts for unobserved initial conditions. Such initial conditions due to a child's unobserved endowment are assumed to exert a constant effect over time on the formation of cognitive skills which means that by taking first-differences we eliminate them from the specification.[23] We focus on the change in reading and writing scores as an indicator of child's cognitive development. There are two reasons for doing so. Firstly, Cueto et. al. (2009) find evidence that the change in skill development of mathematics, for children in India is quite negligible. The authors assess the technical validity of many of the academic and psychometric tests administered in the YL data and find that by Round 2 of data collection, most children in India could already do maths and improved only in the other areas (writing and reading). Those children who had picked up the mathematical skill early on (a high 87%) continued to do well even after four years. Secondly, we find a huge improvement in writing skills of children between the two rounds. The percentage of children who were able to write without difficulty improved from 51% to 69% in four years. Table 1 shows the summary statistics of reading and writing scores at age eight and twelve as well as in first-differences.

4.3 Inputs

Since we implement a specification in first-differences, only inputs into the production of cognitive skills that change over time can be included in the conditioning set of variables. This rules out the use of variables such as a child's gender, caste, birth order etc. Yet we also split the sample by gender to investigate differences in the role peers may play to insure children against adverse idiosyncratic shocks.

The set of potential input variables is further restricted because of the differences in the design of the questionnaires in the two surveys. We use child anthropometries, that is child weight and height as proxies for child health and unobserved initial conditions whose expression varies over time. In addition, we include the change in the number of a child's siblings. This variable captures changes in parental input as well as potential effects arising from within-family interaction and household size. In addition, we include a measure of how much time the child spends working. Given that a child divides her time between school, work, and leisure, the change in time spent working captures changes in the time spent at school and spare-time activities. We therefore do not account separately for a change in schooling. As a measure for the quality of a child's schooling, we include a variable indicating whether the child moved from a public to private school between age eight and twelve. As a direct measure of the resources available to a child, we include the change in

---

[22]We use standardized scores. The reading item required children to read three letters, one word, and one sentence. The reading item was scored as follows: 1 point if children could read the sentence, 0.66 point if they could read the word, 0.33 point if they could read the letters and 0 points if they could not read anything or did not respond. The writing item asked the children to write a simple sentence which was spoken out loud by the examiner. This item was scored as follows: 1 point if children could write the sentence without difficulty or errors, 0.5 point if they could write with difficulty or errors, and 0 points if they could not write anything or did not respond (Cueto et. al, 2009).

[23]We recognize that this is a strong assumption as the expression of a child's initial endowment over time may vary as a function of a child's environment. To the degree that changes in child anthropometries, specifically changes in child height and weight, proxy for time-varying effects of unobserved initial endowments (see Weedon et al., 2007), this problem is mitigated as we include these variables in the conditioning set (see Section 4.3).

household assets.[24] Moreover, we include a dummy variable indicating whether a child's household is located in a (semi-)urban area and a dummy variable for whether the household is located in the coastal area of Andhra Pradesh. These indicator variables capture time-varying location-specific effects.

The survey captured detailed information on various shocks faced by the child's household including economic, climatic, health and other miscellaneous shocks. We were able to divide these shocks into (a) idiosyncratic shocks and (b) covariate shocks. Idiosyncratic shocks include shortfall of food, loss of livestock, death or serious illness of family members, and job loss in the household, and whether the household was subjected to crime. Here, we take care to include only those shocks that are specific to any given household and are not correlated with the occurrence of the same in other household (this is validated in the results section later on). Covariate shocks include natural disasters and calamities and crop failures. We create two shock variables which are dummy variables indicating whether the child's household had experienced an idiosyncratic or covariate shock respectively between age eight and twelve. The covariate shock is part of the basic model specification, whereas the idiosyncratic shock variable is used to test for insurance.

5 Results

We first provide descriptive evidence for spatial peer effects on child cognitive achievement gains. Figures 3 and 4 present nonparametric plots of a child's first-differenced reading and writing scores against the average first-differenced scores of her five nearest neighbors. Both graphs provide descriptive evidence that peer effects matter as they show that a child's own score gain is increasing in her peer's performance.

5.1 Peer Effects

Table 2 shows the results for writing skills from estimating Equation (5) using OLS and the estimation procedure described in Section 3 above. Columns (1) and (2) show OLS results not accounting for potential heteroscedasticity and spatial autocorrelation in the error term. Column (1) does not allow for peer effects; the results show that a number of variables is statistically significant: the change in child height and weight, the change in the time worked, the public school variable, and the indicator for whether the household is located in an urban area. Since the public school variable is equal to one if the child switched from private school to public school between age eight and twelve, zero if no switch occurred and minus one if the child switched in the other direction, the positive coefficient suggests that children benefitted on average from switching from private to public school between age eight and twelve. Since we use first-differenced variables, the location-specific variables, urban and coastal area, capture time-variant location specific unobservables. The fact that the urban indicator variable is statistically significant suggests that this is a relevant concern despite using first-differences, i.e., that different locations are on different trajectories in terms of child skill development. When we account for peer effects in Column (2), we note a positive and statistically significant coefficient associated with endogenous peer effects. An increase of one standard deviation of the endogenous effects leads to an increase of 0.2 standard deviations in the growth of writing skills. The coefficient on the endogenous peer effects as estimated by the OLS

---

[24]We use a wealth index which consists of three components: housing quality, consumer durables such as a refrigerator or a telephone, and services such as electricity or toilets available to the household.

specification is not close to one in the typical linear-in-means sense, because of the non-linearities induced by the network structure. However the OLS estimates still remain biased due to simultaneity issues and other unobservable factors. The other covariates that were statistically significant in Column (1) remain so.

In order to test for the presence of spatial dependence in the error term, we apply a Lagrange Multiplier (LM) test to the OLS residuals in Columns (1) and (2). The null hypothesis is the absence of spatial autocorrelation, which is tested against the presence of autocorrelation captured by a spatial error component (Anselin and Hudak, 1992). The p-value reported in Columns (1) and (2) suggests the presence of spatial dependence in the residuals even when accounting for a spatial lag of the dependent variable. We apply the studentized Breusch-Pagan test to check for the presence of heteroscedasticity. The p-values reported in Column (1) and (2) show strong evidence for the presence of heteroscedasticity.

Column (3) accounts for heteroscedasticity by estimating White's (1980) robust standard errors. Obviously, the point estimates remain unchanged. Also, the peer effect term remains statistically highly significant. Overall, we find that the standard errors are only marginally changed when computing White's (1980) robust standard errors which is surprising given the strong evidence for the presence of heteroscedasticity shown in Columns (1) and (2).

Columns (4) and (5) report the results when using the S2SLS estimator which accounts for endogenous effects through its instrumental variable approach. The estimates in Column (4) assume homoscedasticity and absence of spatial autocorrelation in the residuals of Equation (5), but account for the endogeneity of the spatial lag variable. Hence, the results in Column (4) show the bias when ignoring this endogeneity in Columns (2) and (3). Column (5), in contrast, reports the results for the SHAC estimator, which is least restrictive in terms of assumptions imposed on the residuals of Equation (5) allowing for both heteroscedasticity and spatial autocorrelation and is therefore our preferred estimator. The results in Columns (4) and (5) show a large increase in the coefficient of peer effects. Now, an increase of one standard deviation in a child's peers, has twice the effect on the change in writing skills it had when using OLS. Among the conditioning variables, an increase in the height and weight of a child is negatively associated with cognitive achievement. The coefficient is positive and statistically significant for the public school variable. The urban dummy variable is no longer statistically significant.

Columns (6) and (7) report the results when also accounting for contextual effects. Column (6) reports OLS results, whereas Column (7) reports the SHAC estimates.[25] Again, we note a substantial difference in the coefficients associated with endogenous effects between OLS and the spatial two-step estimator. The coefficient obtained for the S2SLS estimator suggests that a one standard deviation increase in peer skill growth increases writing skill formation by nearly half a standard deviation, which is a sizeable effect. Among contextual effects, only the change in height of a child's peers as well as whether her peers have switched between public and private school are statistically significant. Cooley (2010) provides a detailed discussion on the specification and interpretation of contextual effects in the clasroom/child learning context. She argues that when the child is able to choose her own effort with a view to increase outcomes, it is unclear whether increasing peer exogenous characteristics will have a positive or a negative effect. This is because higher values of peer exogenous characteristics might reduce own outcome values if there are positive spillovers from endogenous peer effects, which we condition on. For instance, consider

---

[25]We computed the corresponding variance inflation factors to investigate the potential presence of multicollinearity, but there is no evidence for this.

a child whose writing scores are increasing in its peers' writing scores as well her own school choice (switching to public school). If all her peers also decided to switch to public school and increased their effort then, controlling for the child's own school switch and her peers' achievement levels, we would expect to see a decrease in own achievement levels because the child is inclined to reduce effort to reap the positive peer spillovers. Overall, the precise interpretation of contextual effects having accounted for endogenous effects is at best ambiguous.

Table 3 reports the corresponding results for the change in reading scores. In Column (1), we report again OLS results ignoring both endogenous and contextual effects. Among the covariates, as for writing scores, the public school dummy variable, and the indicator for whether the household is located in an urban area are statistically significant. In addition, also the variable indicating whether a household is located in the coastal area is statistically significant. In Column (2), we add endogenous peer effects. The coefficient is positive and statistically significant: a one standard deviation increase in a child's peers' skill growth is associated with an increase of a fifth of a standard deviation of the child's own skills. The LM test for spatial autocorrelation in the residuals is flatly rejected. Similarly, the Breusch-Pagan test strongly suggests a non-constant variance of the residuals. Column (3) reports OLS results with robust standard errors where again the standard errors do not differ dramatically. In Columns (4) and (5), we report the results from using the S2SLS estimator. As for writing scores, the coefficient associated with peer effects increased markedly. In Column (5), for our preferred estimation method, the SHAC, a standard deviation of a child's peers' skill change, increases her own skill growth by slightly more than half a standard deviation. Columns (6) and (7) show the coefficients when including contextual effects in the model. The coefficients for peer effects drop slightly for both estimators, OLS and SHAC, and we find only a child's peers' change in weight and the public school variable to be statistically significant among the contextual effects.

In the remainder of this section, we discuss a range of additional results that employ different modifications of the basic specification to investigate the importance of potentially omitted unobservables in driving our results. However, considering the great similarity in the results found in Tables 2 and 3, we limit this discussion to the results obtained for writing test scores.[26]

Table 4 reports results when accounting for potentially omitted unobservables by augmenting the specification with the village-level schooling variable constructed using children's older siblings or the local within-network transformation as described in Section 2.2 above. In addition, the table contains results for an IV estimation, in which we instrument endogenous peer effects with idiosyncratic shocks.

Column (1) shows OLS results when allowing for peer effects and controlling for time-varying unobservables that are correlated with peer effects, such as unobserved changes in the availability or quality of schools. We note that the peer effect coefficient remains nearly unchanged compared to the results reported in Column (2) of Table 2. The schooling of older siblings averaged at the village-level is statistically not significantly different from zero, providing no evidence for a bias of the coefficient associated with endogenous effects due to time-varying schooling-related unobservables. When we look at the second column in which the SHAC results are reported, we note that the added control variable is also not statistically significant and the magnitude of the peer effect estimate very close to the one reported in Table 2.

Column (3) reports the results when using a local within-transformation that accounts for unob-

---

[26]The corresponding results when using reading test scores as the dependent variable are broadly similar to those shown for writing test scores and are available upon request from the authors.

servables at the nearest neighbor network-level.[27] The magnitude of the peer effect coefficient falls slightly relative to Column (2). However, interpretation of the effect is difficult because under the local transformation peer effects represent the deviation of a child's gain from the average gain of her peers. Finally, Columns (4) and (5) report results for an IV approach exploiting the exogeneity of a child's idiosyncratic shock with respect to her peers' cognitive achievement gain. Hence, in a first stage, we use the idiosyncratic shocks as an instrument for change in peer skills. The exclusion restriction is that an idiosyncratic shock hitting child i affects her peers only through its direct effect on the cognitive achievement gain of child i. This is a credible assumption given the idiosyncratic nature of the shocks. To validate this identification strategy, we include average peer idiosyncratic shocks in our contextual effects specification shown in Column (7) of Table 3, while controlling for own idiosyncratic shocks and find a statistically insignificant effect associated with it.[28] The results in Column (4) show that idiosyncratic shocks affect cognitive skill growth adversely in a statistically significant way, which suggests that the instrument is also informative. When we look at the results in Column (5), we note the similarity of the magnitude of the peer effects coefficient with respect to the coefficients obtained using the S2SLS estimator. This is not that surprising given that the S2SLS uses all variables included in the first stage in Column (4) in its instrument set – with the exception of the shock variable. Nevertheless, this finding lends further credibility to our choice of the spatial two-step estimator as our preferred estimator.

We now use our second peer group classification which is based on children belonging to the same community, i.e., a child's peer reference group consists of all other children in the sample who belong to the same community. Table 5 reports results from the community based classification. Since we rely only on variation in group size for identification, we only include endogenous effects in the specification at the community level, i.e., we assume that $\delta = 0$ in Equation (2), to avoid a potential problem of weak instruments. The results are similar to those obtained using the neighborhood-based peer group classification. Peer effects are positive and significant across all different specifications.

Overall, these results provide strong evidence for peer effects to matter for the development of cognitive skills. Moreover, endogenous effects appear to be much more important economically than contextual effects and we find no evidence for time varying correlated effects to bias our results.

5.2 Insurance Test

Next, we provide evidence to show that peer groups can help children cope with idiosyncratic, adverse shocks. Both OLS and spatial estimators with and without the inclusion of peer effects are presented in Table 6. Columns (1) and (2) reports results using the five nearest neighbor network definition whereas Columns (3) and (4) report community-level results. For both, nearest-neighbor and community-level network definitions, we find that the idiosyncratic shock variable has a negative and significant effect on child cognitive achievement measured as the change in writing test scores. This effect, however, becomes insignificant after accounting for peer effects. This result holds

---

[27]Since the local transformation eliminates a considerable amount of variation in the data, we choose to estimate the transformed model assuming that the error process follows a known SAR(1) process using a GMM estimator proposed by Kelejian and Prucha (2009).

[28]The coefficient of $W\Delta HH$ Shock (Idiosyncratic) is -0.004 with a standard error of 0.179. In unreported results, we also regress the peer idiosyncratic shock directly on the outcome variable without including endogenous effects and also find a statistically insignificant effect.

for both neighborhood and community based classifications of peer groups. This means that conditional on peer effects, idiosyncratic shocks have no effect on cognitive skill formation. At the same time, the coefficient associated with peer effects is positive and statistically significant as shown in the preceding section. This provides evidence in favor of peer effects helping children to cope with adverse idiosyncratic shocks.

To investigate potential underlying heterogeneity in peer effects, we split the sample by gender. In rural India, children are often treated differently depending on their gender. In particular, girls maybe be more restricted in the ability to move around freely outside of the household and therefore may be exposed to less interaction with their peers. Another reason to split the sample is the possibility that boys tend to interact more with boys and girls with girls. Table 7 contains the corresponding results. The most interesting finding is that the idiosyncratic shock negatively affects skill formation of girls whereas the effect is insignificant for boys. This supports previous findings in the literature suggesting the presence of widespread discrimination against girls in developing countries. Bjorkman (2009) for instance finds that negative income shocks have large negative and highly significant effects on female enrollment in primary schools whereas the effect on boys' enrollment is smaller and only marginally significant. A large body of evidence exists that shows that girls receive much less schooling in rural India and that intra-household expenditure towards children is skewed in favour of boys (Alderman et. al., 1994; Rosensweig and Schultz, 1982; Behrman, 1998). Further, our results show that the adverse effects of negative shocks to the household on girls' educational achievement continues to persist even after the inclusion of peer effects in the model (however the coefficient on shocks falls marginally). Hence, we find that peer effects tend not to insure girls against adverse shocks. It would be interesting to see whether this effect plays out differentially in rural as compared to urban areas. Given the limited sample size, it is not possible to carry out this exercise with the present data, but we highlight this as a matter for future research.

5.3 Robustness

In this section, we conduct various robustness checks to assess the robustness of our estimates to different sorts of bias. Our key identifying assumption centers around the specification of the peer network structure. This section, therefore, investigates the sensitivity of our results to different assumptions regarding the peer network structure.

5.3.1 Monte Carlo Experiment: Repeated Sampling

First, we provide Monte Carlo evidence to investigate the sensitivity of our results to mis-specification of the social network structure due to the fact that we have data only on a random sample of children. We simulate data representing, say, a fictitious village. We generate 260 data points and assign them random locations in space. This represents our 'population'. Based on the locations of the units in our population, we build our peer reference group based on the envisaged spatial structure, i.e., five nearest neighbors. We then estimate the parameters of a simple SAR model for the population (β denotes the peer effect and subscript l a network):

$$y_l = \alpha + \beta W y_l + \delta x_l + u_l \tag{15}$$

Next, we draw random samples, of size 30%, 40%, 50% and 60% from our population. For each sample size, we draw 200 bootstrap samples and then estimate Equation (15) to obtain sample

estimates. We compare the distribution of these sample estimates, i.e., peer effect estimates, to the true population estimate. Figure 5 plots the histograms of the sample point estimates as a deviation from the population estimate for each of the four sample sizes. The figure shows that on average, the sample estimates converge in density to the population value. This provides some evidence that estimates based on a sample structure of networks are likely to be robust to sampling error if sampling is undertaken at random and the network structure is built around the assumption that the probability of observing missing links depends only on observing the units that constitute the links. We note here that this is a feature of the network structure that we postulate, i.e., a spatially driven network structure. This assumption may be invalid for other types of network structures, such as kinship, friendship and more generally structures which incorporate links that are strategically formed.

5.3.2 Monte Carlo Experiment: Assessing Parameter Bias using Networks with Controlled Topology

Next we explore bias that arises from empirical mis-specification of the true albeit unknown network structure. To do this, we use a controlled network topology driven Monte Carlo experiment similar to Páez et al. (2008) discussed in Section 2.3. As in Section 5.3.1 above, we estimate a population model and compare it to the distribution of sample estimates, but this time we do not impose any specific network structure on the population. Instead, we allow the network structure to be chosen at random, controlling only two topological network parameters: network density and spatial dependence.

The simulation is carried out in two steps. First, we simulate several random networks in the population given a specific combination of network topology parameter values. We do this for all combinations within a specified range of parameter values. For each random network we estimate the population model and compare it to the distribution of sample estimates obtained from 200 sampling draws from the population. The sample estimates are based on a model whose network structure is given by the five nearest neighbors definition. Hence, in essence, we compare the peer effect estimate obtained from the population with a random network structure to the estimate obtained from a sample with a five nearest neighbor network structure over a range of two network topology parameters.

We calculate the Mean Squared Error (MSE) of the sampling estimates from its true population value to benchmark the parameter configurations. The MSE combines both, the estimation variance as well as the bias of an estimate into a measure of goodness of fit. The mean square error of any given parameter, β, is given by (Florax and Rey, 1995):

$$MSE(\beta) = \Sigma_r \frac{(\hat{\beta}_r - \bar{\hat{\beta}})^2}{R} + \left[\frac{\Sigma_r (\beta_s - \hat{\beta}_r)}{R} \right]^2 \qquad (16)$$

where $\hat{\beta}_r$ is the estimate in replication r, $\bar{\hat{\beta}}$ is the mean of the estimate for all replications, β_s is the true value of the parameter, and R is the number of replications in the simulation experiment. This measure adjusts for both over- and under-specification by using squares of the total deviations thus accounting for both negative and positive deviations which ensures that the bias will not average out. Figure 6 shows a surface plot of the MSE along with the relevant network parameters associated with each MSE value. The figure shows that the MSE is lowest when the spatial dependence in the true social network is high and network density is low. The quality of the

five nearest neighbor sample estimator suffers with decreasing spatial dependence as captured in the true weighting scheme.

To place these results in the context of our data, we map the spatial topology of a random sample village in our data. To do this we analyze the spatial pattern of the sample village incident point data. We make use of a variant of Ripley's K-function statistic which evaluates a given spatial distribution in relation to complete spatial randomness.[29] When the observed K-value is larger than the expected K-value for a particular distance, the distribution is more clustered than a random distribution at that distance. When the observed K-value is smaller than the expected K, the distribution is more dispersed than a random distribution at that distance (for further details see Boots and Getis, 1988). Figure 7 plots the distribution of both observed and expected K-values along with the confidence intervals. The figure shows that the observed K-value is larger than the higher confidence envelope implying that spatial clustering for *smaller distances* is statistically significant. At larger distances (exceeding 2 kilometers), the observed K-value is smaller than the lower confidence envelope, therefore, spatial dispersion for that distance is statistically significant. This shows that the optimal spatial network involves specifying relatively few neighbors so as to capture high spatial clustering.[30]

Thus, since we find the presence of substantial spatial autocorrelation in our estimates and specify a sparse network structure, these Monte Carlo results suggest that the empirical bias resulting from potential mis-specification of the sample network structure might be relatively minor.

5.3.3 Network Size

In order to check that our peer effects are not driven solely by choosing a five nearest-neighbor peer group structure, we provide additional results from varying the size of neighbor groups. In the network data, the size of a child's peer group is restricted to some arbitrary number as administered in the survey. Therefore, it is difficult to see how the results would change if the survey had recorded more or less peers for the same sample of respondents. In our analysis, we rely on the construction of peer groups *after* data collection. As described earlier we use the method of K-nearest neighbors to construct peer groups for each child. Initially, we restricted this set to a number of five. We now consider nearest-neighbor groups of three and seven. Tables A-II, and A-III in the Appendix show results for these different neighbor sizes limiting ourselves to the change in writing test scores as the results for reading test scores are qualitatively very similar. Our results remain largely unchanged; the coefficient on peer effects is large, positive and statistically significant for nearly all specifications. The magnitude of the coefficient however falls when we use three nearest neighbors. When using seven nearest neighbors, the magnitude increases slightly relative to the results obtained when using five nearest neighbors. This might reflect the findings by Lee (2009) discussed above. Lee (2009) suggests that the bias from under-specification of the true network structure points downward whereas that from over-specification upward relative to the estimate obtained from the true network structure. The results reported in Tables A-II and A-III, therefore, provide additional support for our choice to construct peer groups based on a five nearest neighbor

---

[29]A variant of the K-function, termed L(D) is given by $L(d) = \sqrt{\frac{A\sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n} k(i,j)}{\pi n(n-1)}}$ where d is the distance, n is equal to the total number of features. A represents the total area of the features and $k(i,j)$ is the weight. The weight will be equal to one when the distance between i and j less than d and will equate zero otherwise. We specify 10 distance bands to capture the correct level of clustering.

[30]The maximum nearest neighbor distance in our 5 nearest neighbor specification is 1.5 km which incorporates all levels of spatial clustering.

structure.

5.3.4 Network Scramble

Finally, we provide a falsification test for our networks based identification strategy. All our results indicate the presence of positive and significant peer effects based on a very specific distance-based peer interaction network of a child. We now show that such a result is not obtained from considering *just any random* peer group network. In essence, we validate the strength and significance of the actual observed network by ruling out the presence of peer effects within randomly generated networks. Our objective is to demonstrate that no statistically significant peer interaction is found among children that have been assigned randomly to a peer network. This is a test for our identifying assumption that geographical proximity mediates peer effects. To test this, we randomly assign each child in the sample five nearest neighbors and estimate the model in Equation (6) using the S2SLS estimator employing the randomly generated spatial weight matrix. Again, we limit ourselves to cognitive skills measured as writing skills as the results carry over to reading skills. We repeat this exercise 300 times, each time generating random five-nearest neighbor networks. The histogram in Figure 8 shows the empirical distribution of the point estimates obtained from the 300 replications. The mean estimate is 0.019 with a standard error of 0.296, which means that we cannot reject the null hypothesis that peer effects are equal to zero. Moreover, the majority of point estimates are statistically insignificant. For peer effects associated with each of the 300 iterations, we find that we are unable to reject the null that the coefficient (point estimate) is different from zero for 275 coefficients out of 300. Figure 9 plots the joint distribution of coefficients alongwith their standard errors obtained from randomizing the network 300 times. The grey dotted elipse in the figure encloses the area under which coefficient estimates are statistically insignificant. We observe that most point estimates lie within this area. This shows that repeated experiments with different randomized networks produce statistical insignificant peer effects on average. Hence, this exercise corroborates our approach to constructing nearest neighbor peer networks based on geographical proximity.

6 Conclusion

In this paper, we analyze the formation of cognitive skills of children in Andhra Pradesh, India, allowing for spatial peer effects. Making use of the specific nature of our data set, i.e., available data on spatial proximity of households, we define a child's peers as her nearest neighbors in terms of geographical distance. Exploiting intransitivity within the networks formed by nearest neighbors, we are able to address Manski's reflection problem. Using first-differences allows us to identify contextual and endogenous effects separately and to avoid confounding social effects with unobserved heterogeneity. We use a number of additional model specifications to rule out that time-varying unobservables that are correlated with peer effects drive these results. For our preferred estimation method, which is also the least restrictive one, we find that an increase of one standard deviation in the growth of the cognitive achievement of a child's peers, increases cognitive achievement of the child by half a standard deviation. This is a sizeable effect and suggests that peer effects are an important determinant of skill formation thus far neglected in the literature on child skill formation. We also find evidence for peer effects to provide insurance for children of shock-affected affected households against adverse effects to their cognitive skill acquisition. Interestingly, this result does not hold when we estimate the model for a sample containing only girls; we interpret this as ev-

idence suggestive of girls being less able to cope with negative idiosyncratic shocks through their peer support.

We regard this research project as an exciting step forward towards accounting for peer inter-action in the literature on the formation of childhood skill acquisition with large potential for novel policy-relevant insights.

References

[1] Ainsworth James (2002): 'Why Does it Take a Village? The Mediation of Neighborhood Effects on Educational Achievement', *Social Forces*, Vol. 81(1), pp. 117-152.

[2] Alderman H., Behrman R., Lavy V., R. Menon (2001): 'Child health and school enrollment: A longitudinal analysis', *Journal of Human Resources*, Vol. 36(1), pp. 185-205.

[3] Anselin Luc (1988): 'Spatial Econometrics: Methods and Models', Kluwer Academic Publishers.

[4] Anselin Luc and Sheri Hudak (1992): 'Spatial econometrics in practice: A review of software options', *Regional Science and Urban Economics*, Vol. 22, Issue 3, pp. 509-536.

[5] Anselin Luc and Nancy Lozano-Gracia (2008): 'Errors in variables and spatial effects in hedonic house price models of ambient air quality', *Empirical Economics*, Vol. 34, pp. 5-34.

[6] Bayer Patrick, Stephen Ross, and Giorgio Topa (2008): 'Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes,' *Journal of Political Economy*, Vol. 116, No. 6, pp. 1150-1196.

[7] Behrman Jere (1988): 'Intrahousehold allocation of nutrients in rural India: Are boys favored? Do parents exhibit inequality aversion?', *Oxford Economic Papers*, 40(1), pp. 32-54.

[8] Ben Arfa Nejla, Carmen Rodriguez, Karine Daniel and J.S. Shonkwiler (2009): 'Spatial Structure of the French dairy sector: a spatial HAC estimation', mimeo.

[9] Bivand Roger with contributions by Luc Anselin, Olaf Berke, Andrew Bernat, Marilia Carvalho, Yongwan Chun, Carsten Dormann, Stéphane Dray, Rein Halbersma, Nicholas Lewin-Koh, Jielai Ma, Giovanni Millo, Werner Mueller, Hisaji Ono, Pedro Peres-Neto, Markus Reder, Michael Tiefelsdorf and Danlin Yu (2008): 'spdep: Spatial dependence: weighting schemes, statistics and models,' R package version 0.4-24.

[10] Bjorkman Martina (2009) 'Income Shocks and Gender Gaps in Education: Evidence from Uganda', *mimeo, IIES, Stockholm.*

[11] Blume Lawrence and Steven Durlauf (2006): 'Identifying Social Interactions: A Review', Michael Oakes and Jay Kaufman (eds.) Methods in Social Epidemiology, Wiley.

[12] Bobonis Gustavo and Frerderico Finan (2008): 'Neighborhood Peer Effects in Secondary School Enrollment Decisions,' mimeo.

[13] Boots, B. and A. Getis.(1988): 'Point Pattern Analysis', *Sage University Paper Series on Quantitative Applications in the Social Sciences*, series no. 07-001. Sage Publications.

[14] Borjas George (1995): 'Ethnicity, Neighborhoods, and Human-Capital Externalities', *American Economic Review*, Vol. 85, No. 3, pp. 365-390.

[15] Borghans Lex, Angela Lee Duckworth, James Heckman and Bas ter Weel (2008) 'The Economics and Psychology of Personality Traits', *Journal of Human Resources* 43.4: 972-1059.

[16] Boozer M.A. and S.E. Cacciola (2001): 'Inside the'Black Box'of Project STAR: Estimation of peer effects using experimental data', *Economic Growth Center*, Discussion Paper 832, Yale.

[17] Bramoullé Yann, Habiba Djebbari, and Bernard Fortin (2009): 'Identification of peer effects through social networks', Journal of Econometrics, Vol. 150, pp. 41-55.

[18] Brock W. and S. Durlauf (2001) *Handbook of Econometrics*, Elsevier Science B.V, Ch. Interactions-Based Models, pp. 3299-3380.

[19] Brooks-Gunn Jeanne, Greg Duncan, Pamela Klebanov, and Naomi Sealand (1993): 'Do Neighborhoods Affect Child and Adolescent Development?', *American Journal of Sociology*, Vol. 99, No. 2, 353-395.

[20] Cliff Andrew and John Keith Ord (1973): 'Spatial Autocorrelation', Pion, London.

[21] Cooley Jane (2010): 'Can achievement peer effect estimates inform policy? A view from inside the black box', University of Wisconsin mimeo.

[22] Conley Timothy (1999): 'GMM estimation with cross sectional dependence', *Journal of Econometrics*, Vol. 92, pp. 1Ű45.

[23] Conley Timothy and Christopher Udry (2010): 'Learning about a New Technology: Pineapple in Ghana,' *American Economic Review*, Vol. 100(1), pp. 35-69.

[24] Cunha Flavio, James Heckman and Susanne Schennach (2006): 'Estimating the Technology of Cognitive and Noncognitive Skill Formation', mimeo, Univerity of Chicago, Department of Economics.

[25] Cunha Flavio and James Heckman (2007): 'The Technology of Skill Formation', *American Economic Review* 97.2: 31-47.

[26] Cunha Flavio and James Heckman (2008): 'Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation', *Journal of Human Resources*, 43(4):738-782 (2008);

[27] Cueto Santiago, Leon Juan, Guerrero Gabriela and Ismael Munoz (2009): 'Psychometric characteristics of cognitive development and achievement instruments in Round 2 of Young Lives', Young Lives Technical Note No. 15.

[28] De Weerdt Joachim and Stefan Dercon (2006): 'Risk-Sharing Networks and Insurance against Illness', *Journal of Development Economics*, Vol. 81, pp. 337-356.

[29] De Giorgi Giacomo, Pellizzari Michele, and Silvia Redaelli (2010): 'Identification of Social Interaction through Partially Overlapping Peer Groups,' *American Economic Journal: Applied Economics*, Vol. 2, pp. 241-275.

[30] Durlauf Steven (2004): 'Neighborhood Effects', in J.V. Henderson and J.F. Thisse (eds.) Handbook of Regional and Urban Economics, Vol. 4.

[31] Epple, Dennis and Richard E. Romano, (2011): 'Peer Effects in Education: A Survey of the Theory and Evidence,' In: Jess Benhabib, Alberto Bisin and Matthew O. Jackson, Editor(s), *Handbook of Social Economics*, North-Holland, Vol. 1, pp. 1053-1163.

[32] Evans William, Wallace Oates and Robert Schwab (1992): 'Measuring Peer Group Effects: A Study of Teenage Behavior', *Journal of Political Economy*, Vol. 100, pp. 966-991.

[33] Fafchamps Marcel and Pedro Vicente (2009): 'Political Violence and Social Networks: Experimental Evidence from a Nigerian Election', CSAE Working Paper 2009-14.

[34] Florax R.J.G.M. and S. Rey, (1995): ' The impact of misspecified spatial structure in linear regression models'. In: Anselin, L., Florax, R.J.G.M. (Eds.), New Directions in Spatial Econometrics. Springer-Verlag, Berlin, pp. 111Ű135.

[35] Gertler Paul and Jonathan Gruber (2002): 'Insuring Consumption against Illness', *American Economic Review*, Vol. 92, No. 1, pp. 51-70.

[36] Ginja Rita (2010): ŚIncome Shocks and Investments in Human Capital', Working Paper University College London.

[37] Hanushek Eric, John Kain, Jacob Markman, and Steven Rivkin (2003): 'Does peer ability affect student achievement?', *Journal of Applied Econometrics*, Vol. 18, Issue 5, pp. 527 - 544.

[38] Heckman James (1976): 'The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models, *Annals of Economic and Social Measurement*, Vol. 5, pp. 475Ű492.

[39] Heckman James (1979): 'Sample Selection Bias as a Specification Error', *Econometrica*, Vol. 47, pp. 153Ű161.

[40] Helmers Christian and Manasa Patnam (2011): 'The Formation and Evolution of Childhood Skill Acquisition: Evidence from India', *Journal of Development Economics*, Vol. 95, pp. 252-266.

[41] Himaz Rozana, Sheikh Galab and Prudhvikar Reddy (2008): 'Young Lives Round 2 Survey Report Initial Findings: Andhra Pradesh, India,' Young Lives Country Report.

[42] Hoff Peter (2007): 'Multiplicative latent factor models for description and prediction of social networks', *Computational and Mathematical Organization Theory,* 2007.

[43] Hoxby Caroline M. (2000): 'Peer Effects in the Classroom: Learning from Gender and Race Variation,' NBER Working Paper No. 7867.

[44] Ioannides Yannis (2008): 'Social interactions', in *The New Palgrave Dictionary of Economics*, Second Edition, eds. Steven N. Durlauf and Lawrence E. Blume, Palgrave MacMillan.

[45] Katz Lawrence, Jeffrey Kling and Jeffrey Liebman (2001): 'Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment', *The Quarterly Journal of Economics,* Vol. 116, No. 2, pp. 607-654.

[46] Kelejian Harry and Ingmar Prucha (1997): 'Estimation of spatial regression models with autoregressive errors by twostage least squares procedures: a serious problem', *International Regional Science Review,* Vol. 20, pp. 103Ű111.

[47] Kelejian Harry and Ingmar Prucha (1998): 'A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances', *Journal of Real Estate Finance and Economics*, Vol. 17, pp. 99Ű121.

[48] Kelejian Harry and Ingmar Prucha (1999): 'A generalized moments estimator for the autoregressive parameter in a spatial model', *International Economic Review*, Vol. 40, pp 509Ű533.

[49] Kelejian Harry and Ingmar Prucha (2007): 'HAC Estimation in a Spatial Framework', *Journal of Econometrics,* Vol. 140, pp. 131-154.

[50] Kelejian Harry and Ingmar Prucha (2009): 'Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances,' *Journal of Econometrics*, forthcoming.

[51] Lambert Dayton, Raymond Florax and Cho Seong-Hoon (2008): 'Bandwidth selection for spatial HAC and other robust covariance estimators', Deptartment of Agricultural Economics Purdue University Working Paper No. 08-10.

[52] Lee Lung-fei (2007): 'Identification and estimation of econometric models with group interactions, contextual factors and fixed effects', *Journal of Econometrics*, Vol. 140, pp. 333-374.

[53] Lee S.-Y. (2009): 'Bias from Misspecified Spatial Weight Matrices in SAR Models: Theory and Simulation Studies,' mimeo Department of Economics,San Francisco State University.

[54] Lin, Xu. (2010): 'Identifying Peer Effects in Student Achievement by Spatial Autoregressive Models with Group Unobservables', *Journal of Labor Economics*, 28(4).

[55] Manski Charles (1993): 'Identification of Endogenous Social Effects: The Reflection Problem', *Review of Economic Studies*, Vol. 60, No. 3, pp. 531-542.

[56] Marsden Peter and Jeanne Hurlbert (1987): 'Small Networks and Selectivity Bias in the Analysis of Survey Network Data,' *Social Networks*, Vol. 9, pp. 333-349.

[57] McCulloch Andrew and Joshi Heather (2001): 'Neighbourhood and family influences on the cognitive ability of children in the British National Child Development Study', *Social Science & Medicine*, Vol. 53(5), pp. 579-591.

[58] Moffitt (2001): 'Policy interventions, low-level equilibria, and social interactions', in Social Dynamics, eds. Steven Durlauf and Peyton Young, MIT Press.

[59] Openshaw Stan (1984): 'The Modifiable Areal Unit Problem,' Norwich Geo Books.

[60] Outes-Leon and Dercon (2008): 'Survey Attrition and Attrition Bias in Young Lives,' Young Lives Technical Note No. 5.

[61] Paez Antonio, Darren Scott and Erik Volz (2008): ' Weight matrices for social influence analysis: An investigation of measurement errors and their effect on model identification and estimation quality', *Social Networks*, 30, pp. 309-317.

[62] Piras Gianfranco (2010): 'sphet: Spatial Models with heteroskedastic innovations,' R package version 0.1-22.

[63] Rosenzweig Mark R. and Paul Schultz (1982), 'Market opportunities, genetic endowments, and intrafamily resource distribution: Child survival in rural India', *American Economic Review*, 72(4), pp. 803-815.

[64] Sacerdote Bruce (2001): 'Peer Effects with Random Assignment: Results for Dartmouth Roommates', *Quarterly Journal of Economics*, Vol. 116, No. 2, pp. 681-704.

[65] Taskar B., Wong M., Abbeel P. and Koller D. (2004): 'Link prediction in relational data', *Advances in Neural Information Processing Systems*, , vol. 16, 2004.

[66] Tominey Emma (2009): ŚThe Timing of Parental Income on Child Outcomes: The Role of Permanent and Transitory Shocks', Working Paper Centre for Economic Performance, LSE.

[67] Townsend Robert (1994): 'Risk and Insurance in Village India', *Econometrica*, Vol. 62(3), pp. 539-591.

[68] Weedon M.N., Lettre G., Freathy R.M., Lindgren C.M., Voight B.F., Perry J.R., Elliott K.S., Hackett R., Guiducci C. and Shields B. (2007): 'A Common Variant of Hmga2 is Associated with Adult and Childhood Height in the General Population', *Nature Genetics*, Vol. 39(10), pp. 1245-50.

[69] White Halbert (1980): 'A heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity', *Econometrica*, Vol. 48, pp. 817Ű838.

Figure 1: GIS Data Map - Example of a Sample Village

Figure 2: Network Structure - Example of Intransitive Network Structure
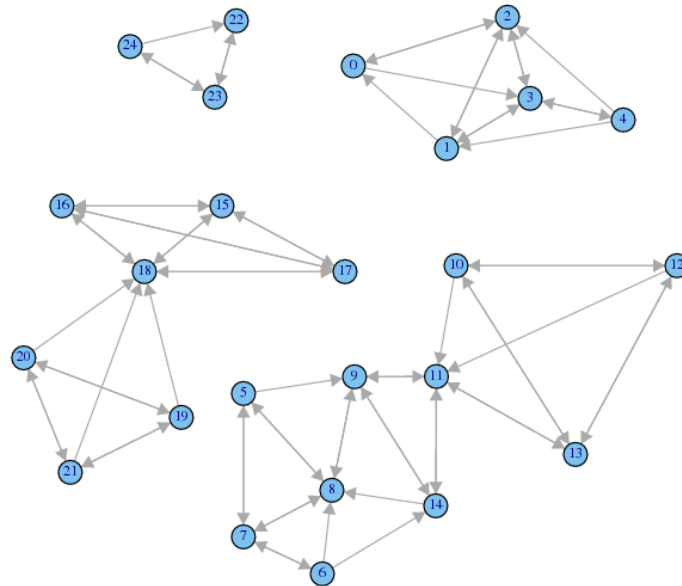
Table 1: Summary Statistics

Variable	No. Obs.	Median	Mean	St. Dev.
CH Writing Score (Age 8)	731	2	2.098	0.691
CH Writing Score (Age 12)	731	3	2.640	0.580
CH Reading Score (Age 8)	731	3	3.046	1.047
CH Reading Score (Age 12)	731	4	3.660	0.780
$\triangle$ CH Writing Score	731	1	0.541	0.873
$\triangle$ CH Reading Score	731	0	0.621	1.114
$\triangle$ CH Weight[2]	731	10.85	12.938	10.95
$\triangle$ CH Height[3]	731	23.1	22.602	11.23
$\triangle$ CH Siblings	731	0	0	0.546
$\triangle$ CH Years of Schooling	731	4	3.652	1.095
$\triangle$ CH Work	731	0	-0.136	0.583
$\triangle$ CH Public School[4]	731	0	-0.123	0.511
$\triangle$ HH Size	731	0	-0.350	1.551
$\triangle$ HH Wealth	731	0.429	0.053	0.012
$\triangle$ HH Covariate Shock	731	0	0.440	0.496
$\triangle$ HH Idiosyncratic Shock	731	0	0.426	0.494
$\triangle$ VIL Years of Schooling	731	3.13	2.970	0.582
Urban Area	731	0	0.265	0.441
Coastal Andhra	731	0	0.184	0.388

Notes:
1. CH: Child; HH: Household; VIL: Village.
2. Child weight: Weight for age z-score.
3. Child height: Height for age z-score.
4. Public School: Defined as [(yes=1) at age 12 - (yes=1) at age 8].

Figure 3: Nonparametric Plots of △ Own Writing Score vs. △ Peer Writing Score
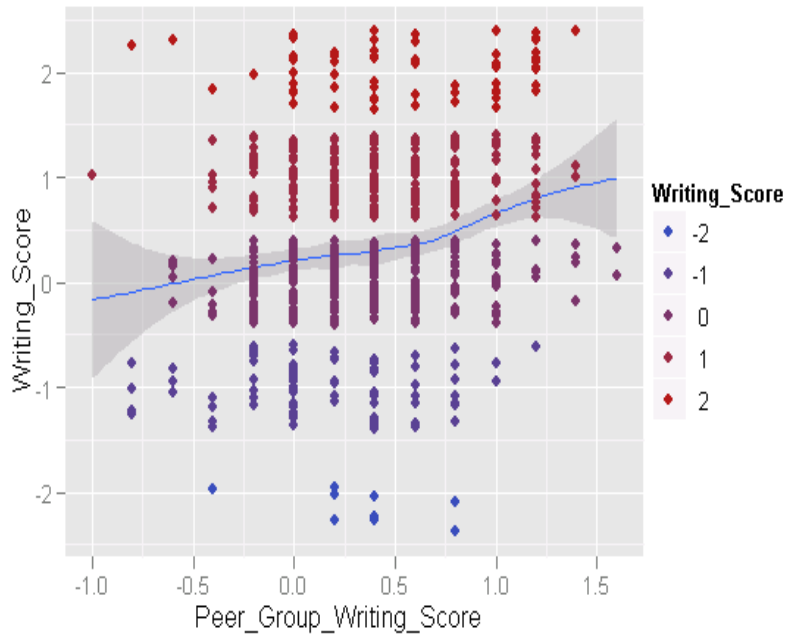
Figure 4: Nonparametric Plots of △ Own Reading Score vs. △ Peer Reading Score
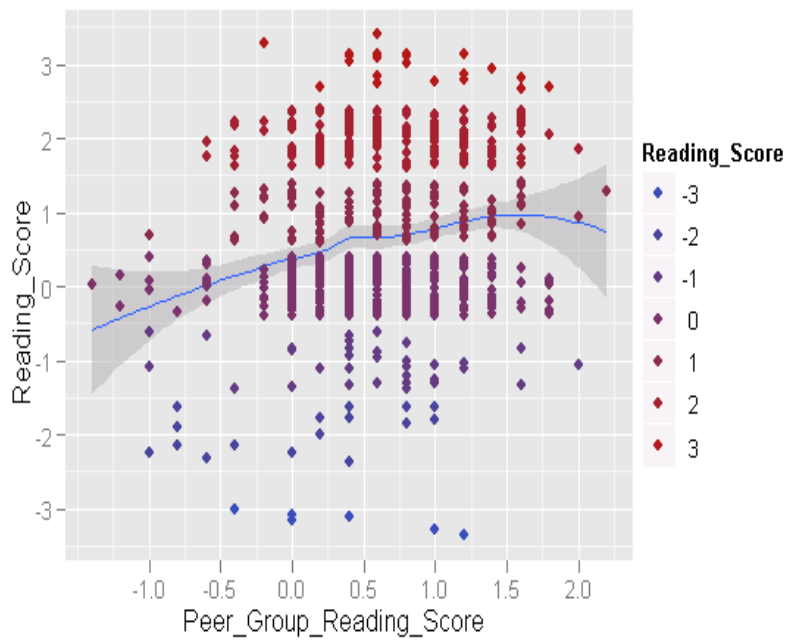
Figure 5: Monte Carlo - Sampling
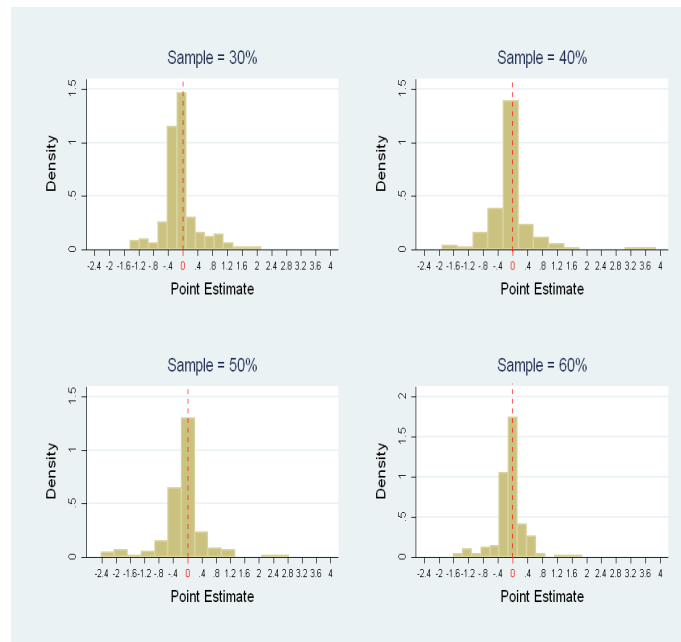
Note: The histograms show the distributions of the difference between the point estimate obtained from using the population network and estimates obtained from 200 bootstrap samples for each different sample size.

Figure 6: Monte Carlo - Mis-specification
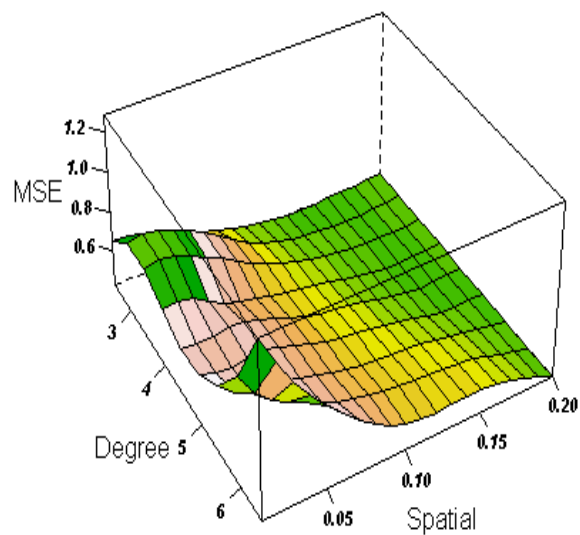
Note: The surface plot maps the distribution of the mean square error over two network topology parameters - network density parameter (degree) and spatial dependence parameter (spatial).

Figure 7: Spatial Topology of a Sample Village
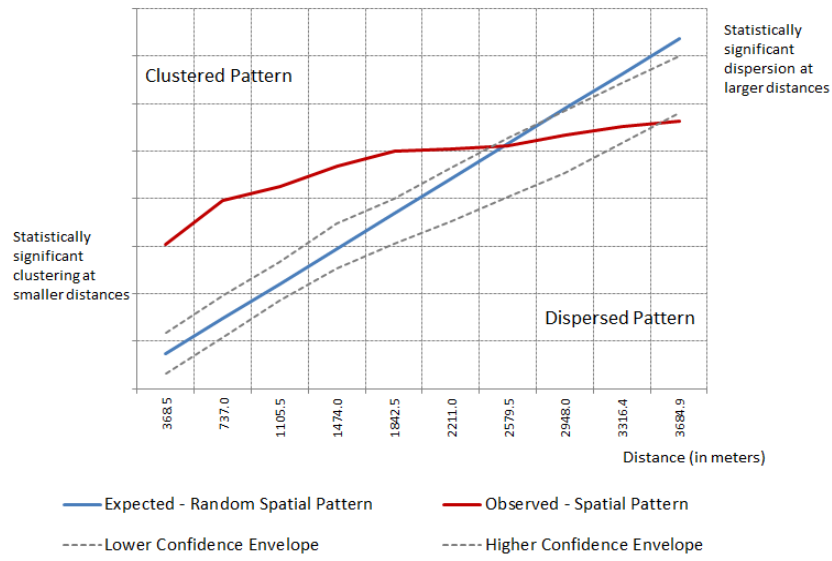
Note: The graph plots the spatial pattern of a sample village. The line plots the distribution of K-values calculated based on Ripley's K-function which evaluates a given spatial distribution in relation to complete spatial randomness. The x-axis represents the distances at which the observed K-value is larger or smaller than the expected K-value. For more details see Boots and Getis (1988).

Figure 8: Network scramble - Histogram of point estimates of endogenous peer effects
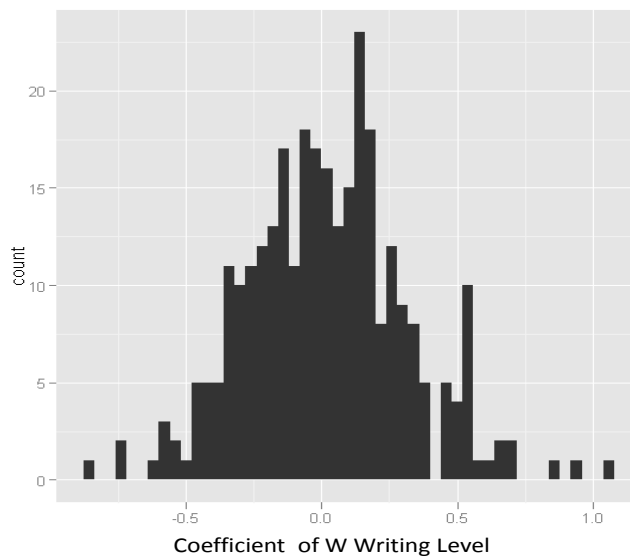
Figure 9: Network scramble - Lattice map of peer effect coefficients and standard errors
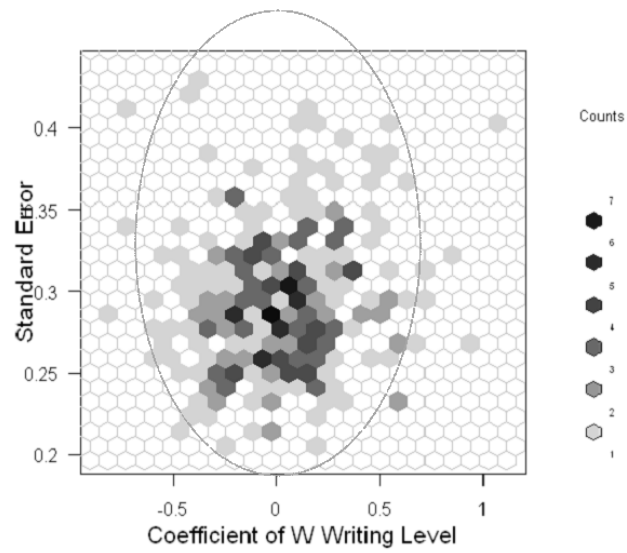
Note: The figure shows the joint distribution of coefficients along with their standard errors obtained from randomizing the network 300 times. Coefficients estimates that fall within the area enclosed by the grey dotted ellipse are statistically insignificant, i.e., not different from zero.

Table 2: Results for **5** Nearest Neighbors: Writing

	OLS		OLS Robust	SAR	SHAC	Contextual Effects	
						OLS Robust	SHAC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
W △ Writing Level	–	0.389**	0.389**	0.871**	0.871**	0.316**	0.963**
		(0.071)	(0.076)	(0.075)	(0.143)	(0.081)	(0.360)
△ CH Weight[2]	-0.011**	-0.010**	-0.010**	-0.009*	-0.009*	-0.008*	-0.007+
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)
△ CH Height[3]	-0.007*	-0.008*	-0.008*	-0.009*	-0.009*	-0.010**	-0.010**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
△ CH Siblings	0.079	0.068	0.068	0.054	0.054	0.071	0.047
	(0.056)	(0.055)	(0.054)	(0.057)	(0.061)	(0.054)	(0.063)
△ CH Work	-0.152**	-0.096+	-0.096+	-0.026	-0.026	-0.072	-0.020
	(0.054)	(0.054)	(0.051)	(0.060)	(0.062)	(0.059)	(0.072)
△ Public School	0.159**	0.163**	0.163**	0.168**	0.168**	0.189**	0.201**
	(0.060)	(0.059)	(0.059)	(0.061)	(0.066)	(0.064)	(0.076)
△ HH Assets	-0.001	0.007	0.007	0.017	0.017	-0.045	-0.036
	(0.238)	(0.234)	(0.247)	(0.241)	(0.263)	(0.252)	(0.280)
△ HH Shock (Covariate)	0.089	0.054	0.054	0.011	0.011	0.034	-0.003
	(0.072)	(0.071)	(0.073)	(0.074)	(0.085)	(0.078)	(0.093)
Urban Area	-0.194*	-0.150+	-0.150+	-0.095	-0.095	-0.052	-0.035
	(0.087)	(0.085)	(0.073)	(0.090)	(0.064)	(0.090)	(0.072)
Coastal Andhra	0.020	0.018	0.018	0.016	0.016	0.102	-0.015
	(0.075)	(0.073)	(0.074)	(0.076)	(0.058)	(0.084)	(0.100)
W △ CH Weight[2]						-0.004	0.004
						(0.007)	(0.007)
W △ CH Height[3]						0.012+	0.015**
						(0.007)	(0.005)
W △ CH Siblings						-0.018	-0.143
						(0.118)	(0.120)
W △ CH Work						-0.218*	0.001
						(0.096)	(0.156)
W △ Public School						-0.210*	-0.223*
						(0.107)	(0.103)
W △ HH Assets						0.134	0.122
						(0.463)	(0.411)
W △ HH Shock (Covariate)						0.295*	0.055
						(0.139)	(0.202)
Breusch-Pagan Test[4]	0.001	0.015					
LM Test[5]	0.000	0.009					
Observations	**731**	**731**	**731**	**731**	**731**	**731**	**731**

Notes:
1. CH: Child; HH: Household.
2. Child weight: Weight for age z-score.
3. Child height: Height for age z-score.
4. P-values (H_0: constant error variance).
5. P-values (H_0: no spatial autocorrelation).
6. Standard Errors in parentheses (White Standard Errors for OLS Robust).
7. + indicates significance at 10%; * at 5%; ** at 1%.

Table 3: Results for **5** Nearest Neighbors: Reading

	OLS	OLS	OLS Robust	SAR	SHAC	Contextual Effects OLS Robust	Contextual Effects SHAC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
W $\triangle$ Writing Level	–	0.394**	0.394**	1.127**	1.127**	0.372**	0.783[+]
		(0.071)	(0.071)	(0.334)	(0.318)	(0.070)	(0.420)
$\triangle$ CH Weight[2]	-0.008[+]	-0.007	-0.007	-0.004	-0.003	-0.005	-0.005
	(0.005)	(0.005)	(0.006)	(0.005)	(0.007)	(0.006)	(0.006)
$\triangle$ CH Height[3]	-0.003	-0.004	-0.004	-0.005	-0.005	-0.005	-0.007
	(0.004)	(0.004)	(0.006)	(0.005)	(0.007)	(0.006)	(0.007)
$\triangle$ CH Siblings	-0.069	-0.076	-0.076	-0.091	-0.091	-0.073	-0.083
	(0.075)	(0.074)	(0.071)	(0.079)	(0.076)	(0.070)	(0.069)
$\triangle$ CH Work	-0.062	-0.039	-0.039	0.002	0.002	-0.011	0.007
	(0.072)	(0.071)	(0.075)	(0.078)	(0.081)	(0.085)	(0.093)
$\triangle$ Public School	0.130[+]	0.145[+]	0.145[+]	0.172*	0.172[+]	0.182*	0.202*
	(0.081)	(0.079)	(0.085)	(0.086)	(0.093)	(0.087)	(0.099)
$\triangle$ HH Assets	-0.108	-0.010	-0.010	0.171	0.171	-0.019	0.048
	(0.318)	(0.312)	(0.313)	(0.344)	(0.345)	(0.321)	(0.333)
$\triangle$ HH Shock (Covariate)	0.015	-0.016	-0.016	-0.077	-0.077	-0.045	-0.083
	(0.096)	(0.095)	(0.100)	(0.105)	(0.123)	(0.109)	(0.128)
Urban Area	-0.232*	-0.199[+]	-0.199[+]	-0.138	-0.138	-0.050	-0.071
	(0.116)	(0.114)	(0.109)	(0.125)	(0.103)	(0.129)	(0.119)
Coastal Andhra	-0.245*	-0.170[+]	-0.170[+]	-0.031	-0.031	-0.057	-0.052
	(0.100)	(0.099)	(0.096)	(0.123)	(0.095)	(0.110)	(0.081)
W $\triangle$ CH Weight[2]						-0.007	-0.013[+]
						(0.010)	(0.008)
W $\triangle$ CH Height[3]						0.001	-0.005
						(0.010)	(0.007)
W $\triangle$ CH Siblings						0.093	0.088
						(0.164)	(0.117)
W $\triangle$ CH Work						-0.241*	-0.166
						(0.126)	(0.143)
W $\triangle$ Public School						-0.323*	-0.318*
						(0.136)	(0.132)
W $\triangle$ HH Assets						-0.458	-0.410
						(0.601)	(0.515)
W $\triangle$ HH Shock (Covariate)						0.361[+]	0.186
						(0.191)	(0.251)
Breusch-Pagan Test[4]	0.000	0.000					
LM Test[5]	0.000	0.011					
Observations	**731**	**731**	**731**	**731**	**731**	**731**	**731**

Notes:

1. CH: Child; HH: Household.
2. Child weight: Weight for age z-score.
3. Child height: Height for age z-score.
4. P-values (H_0: constant error variance).
5. P-values (H_0: no spatial autocorrelation).
6. Standard Errors in parentheses (White Standard Errors for OLS Robust).
7. [+] indicates significance at 10%; * at 5%; ** at 1%.

Table 4: Results for **5** Nearest Neighbors: Alternative Specifications (Writing)

| | | | | IV | |
	OLS	SHAC	SAR-LD[4]	First Stage	Second Stage
	(1)	(2)	(3)	(4)	(5)
W $\triangle$ Writing Level	**0.387****	**0.882****	**0.822**$^{+}$	–	**0.844****
	(0.072)	**(0.146)**	**(0.466)**		**(0.311)**
$\triangle$ CH Shock (Idsync.)	–	–	–	-0.164*	–
				(0.068)	
$\triangle$ CH Weight[2]	-0.010**	-0.008*	-0.009^{+}	-0.012**	-0.010**
	(0.003)	(0.003)	(0.005)	(0.003)	(0.003)
$\triangle$ CH Height[3]	-0.008*	-0.009*	-0.013**	-0.008*	-0.007*
	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)
$\triangle$ CH Siblings	0.067	0.055	0.067	0.074	0.081
	(0.055)	(0.061)	(0.067)	(0.056)	(0.056)
$\triangle$ CH Work	-0.095^{+}	-0.029	-0.017	-0.136*	-0.118
	(0.054)	(0.062)	(0.083)	(0.054)	(0.054)
$\triangle$ Public School	0.162**	0.171**	0.227**	0.165**	0.140**
	(0.059)	(0.066)	(0.087)	(0.060)	(0.060)
$\triangle$ HH Assets	0.006	0.020	0.014	0.004	0.003
	(0.234)	(0.263)	(0.313)	(0.237)	(0.256)
$\triangle$ HH Shock (Covariate)	0.054	0.011	–	0.132^{+}	0.081
	(0.071)	(0.085)		(0.074)	(0.075)
$\triangle$ **VIL Years of Schooling**	0.013	-0.040	–	0.045	-0.016
	(0.064)	(0.054)		(0.065)	(0.068)
Urban Area	-0.153^{+}	-0.087	–	-0.223*	-0.045
	(0.086)	(0.065)		(0.088)	(0.073)
Coastal Andhra	0.014	0.028	–	0.000	0.078
	(0.076)	(0.058)		(0.077)	(0.075)
Observations	**731**	**731**	**731**	**731**	**731**

Notes:
1. CH: Child; HH: Household; VIL: Village.
2. Child weight: Weight for age z-score.
3. Child height: Height for age z-score.
4. Refers to the local-differenced specification using the SHAC estimator (Kelejian and Prucha, 2009).
5. Standard Errors in parentheses; White Standard Errors for OLS Robust; Bootstapped, bias-corrected Standard errors for IV results.
6. $^{+}$ indicates significance at 10%; * at 5%; ** at 1%.

Table 5: Results for Community: Writing

	OLS	OLS Robust	SAR	SHAC
	(1)	(2)	(3)	(4)
W △ Writing Level	–	**0.378****	**0.728****	**0.728****
		(0.076)	**(0.171)**	**(0.149)**
△ CH Weight[2]	-0.011**	-0.010**	-0.010**	-0.010*
	(0.003)	(0.003)	(0.003)	(0.004)
△ CH Height[3]	-0.008*	-0.008*	-0.009**	-0.009*
	(0.003)	(0.003)	(0.003)	(0.003)
△ CH Siblings	0.085	0.082	0.079	0.079
	(0.054)	(0.050)	(0.054)	(0.052)
△ CH Work	-0.151**	-0.105*	-0.064	-0.064
	(0.053)	(0.050)	(0.057)	(0.055)
△ Public School	0.158**	0.163**	0.168**	0.168*
	(0.059)	(0.058)	(0.059)	(0.070)
△ HH Assets	0.002	-0.049	-0.096	-0.096
	(0.230)	(0.236)	(0.231)	(0.240)
△ HH Shock (Covariate)	0.094	0.068	0.043	0.043
	(0.071)	(0.072)	(0.072)	(0.080)
Urban Area	-0.168*	-0.117^{+}	-0.069	-0.069
	(0.084)	(0.071)	(0.087)	(0.064)
Coastal Andhra	0.040	0.036	0.032	0.032
	(0.073)	(0.074)	(0.073)	(0.056)
Observations	**756**	**756**	**756**	**756**

Notes:
1. CH: Child; HH: Household.
2. Child weight: Weight for age z-score.
3. Child height: Height for age z-score.
4. Standard Errors in parentheses (White Standard Errors for OLS Robust).
5. $^{+}$ indicates significance at 10%; * at 5%; ** at 1%.

Table 6: Testing for Insurance - Results for Peer Groups and Idiosyncratic Shocks: Writing

	Neighbours Classification		Community Classification	
	OLS	SHAC	OLS	SHAC
	(1)	(2)	(3)	(4)
W $\triangle$ Writing Level	–	**0.894****	–	**0.537****
		(0.135)		**(0.191)**
$\triangle$ HH Shock (Idsync.)	**-0.164***	**-0.116**	**-0.152***	**-0.098**
	(0.066)	**(0.068)**	**(0.080)**	**(0.072)**
$\triangle$ CH Weight[2]	-0.012**	-0.009*	-0.012**	-0.010*
	(0.003)	(0.004)	(0.003)	(0.004)
$\triangle$ CH Height[3]	-0.008*	-0.009*	-0.008*	-0.010*
	(0.003)	(0.003)	(0.003)	(0.003)
$\triangle$ CH Siblings	0.074	0.053	0.080	0.076
	(0.056)	(0.062)	(0.054)	(0.052)
$\triangle$ CH Work	-0.136*	-0.020	-0.135*	-0.052
	(0.054)	(0.060)	(0.053)	(0.053)
$\triangle$ Public School	0.165**	0.179**	0.164**	0.175**
	(0.060)	(0.067)	(0.060)	(0.071)
$\triangle$ HH Assets	0.004	0.027	0.005	-0.096
	(0.237)	(0.265)	(0.229)	(0.242)
$\triangle$ HH Shock (Covariate)	0.132^{+}	0.041	0.134^{+}	0.067
	(0.074)	(0.082)	(0.073)	(0.078)
$\triangle$ VIL Years of Schooling	0.045	-0.049	0.042	-0.007
	(0.065)	(0.053)	(0.064)	(0.047)
Urban Area	-0.223*	-0.097	-0.196*	-0.076
	(0.088)	(0.068)	(0.085)	(0.069)
Coastal Andhra	0.000	0.027	0.023	0.031
	(0.077)	(0.058)	(0.076)	(0.058)
Observations	**731**	**731**	**756**	**756**

Notes:
1. CH: Child; HH: Household; VIL: Village.
2. Child weight: Weight for age z-score.
3. Child height: Height for age z-score.
4. Standard Errors in parentheses (White Standard Errors for OLS Robust).
5. $^{+}$ indicates significance at 10%; * at 5%; ** at 1%.

Table 7: Results for Boys Vs. Girls: Writing

	Boys		Girls	
	OLS	SHAC	OLS	SHAC
	(1)	(2)	(3)	(4)
W $\triangle$ Writing Level	–	**0.574****	–	**0.818****
		(0.207)		**(0.199)**
$\triangle$ **HH Shock (Idsync.)**	**-0.051**	**-0.050**	**-0.265****	**-0.231***
	(0.099)	**(0.108)**	**(0.087)**	**(0.093)**
$\triangle$ CH Weight^2	-0.019**	-0.017**	-0.001	-0.002
	(0.005)	(0.006)	(0.006)	(0.005)
$\triangle$ CH Height^3	-0.013*	-0.012^{+}	-0.007	-0.010*
	(0.005)	(0.005)	(0.005)	(0.005)
$\triangle$ CH Siblings	0.020	-0.000	0.108	0.111
	(0.088)	(0.107)	(0.068)	(0.058)
$\triangle$ CH Work	-0.068	-0.055	-0.257**	-0.140
	(0.084)	(0.078)	(0.074)	(0.073)
$\triangle$ Public School	0.193*	0.192*	0.185*	0.180*
	(0.088)	(0.091)	(0.084)	(0.092)
$\triangle$ HH Assets	0.014	-0.033	-0.114	-0.177
	(0.325)	(0.376)	(0.338)	(0.321)
$\triangle$ HH Shock (Covariate)	0.126	0.110	0.165^{+}	0.085
	(0.110)	(0.117)	(0.100)	(0.102)
$\triangle$ VIL Years of Schooling	0.092	0.028	-0.079	-0.053
	(0.099)	(0.083)	(0.087)	(0.063)
Urban Area	-0.172	-0.070	-0.262*	-0.101
	(0.126)	(0.099)	(0.118)	(0.083)
Coastal Andhra	0.043	0.107	0.007	0.001
	(0.115)	(0.088)	(0.103)	(0.084)
Observations	**374**	**374**	**365**	**365**

Notes:
1. CH: Child; HH: Household; VIL: Village.
2. Child weight: Weight for age z-score.
3. Child height: Height for age z-score.
4. Standard Errors in parentheses (White Standard Errors for OLS Robust).
5. $^{+}$ indicates significance at 10%; * at 5%; ** at 1%.

APPENDIX

Table A-I: Missing Data Characteristics

Variable	Mean Sample	Mean Missing	Difference	Std. Error (of Difference)
△ CH Weight[3]	12.866	12.922	-0.056	0.715
△ CH Height[4]	22.568	23.640	-1.071	0.731
△ CH Siblings	0.0026	0.078	-0.075	0.039
△ CH Years of Schooling	3.654	3.654	0.0004	0.079
△ CH Work	-0.1370	-0.320	**0.183**	0.041
△ CH Public School[5]	-0.123	-0.160	0.036	0.036
△ HH Size	-0.362	-0.304	-0.057	0.115
△ HH Wealth	0.053	0.057	-0.004	0.009
△ HH Covariate Shock	0.436	0.176	**0.259**	0.034
△ HH Idiosyncratic Shock	0.421	0.267	**0.154**	0.035
Urban Area	0.192	0.423	**-0.231**	0.031
CH Male	0.501	0.526	-0.024	0.036

Notes:
1. Differences that are significant at 5% are indicated in bold.
2. CH: Child; HH: Household; VIL: Village.
3. Child weight: Weight for age z-score.
4. Child height: Height for age z-score.
5. Public School: Defined as [(yes=1) at age 12 - (yes=1) at age 8].

Table A-II: Results for **3** Nearest Neighbors: Writing

	OLS	OLS	OLS Robust	SAR	SHAC	Contextual Effects OLS Robust	SHAC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
W $\triangle$ Writing Level	–	0.355**	0.355**	0.468*	0.468*	0.342**	0.061
		(0.057)	(0.076)	(0.062)	(0.196)	(0.064)	(0.437)
$\triangle$ CH Weight[2]	-0.012**	-0.010**	-0.010**	-0.010**	-0.010*	-0.008*	-0.009**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)
$\triangle$ CH Height[3]	-0.008*	-0.008*	-0.008*	-0.008*	-0.008*	-0.010**	-0.011**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
$\triangle$ CH Siblings	0.084	0.062	0.062	0.055	0.055	0.064	0.082
	(0.057)	(0.055)	(0.055)	(0.057)	(0.058)	(0.054)	(0.062)
$\triangle$ CH Work	-0.159**	-0.115^{+}	-0.115^{+}	-0.101^{+}	-0.101^{+}	-0.098^{+}	-0.124^{+}
	(0.054)	(0.053)	(0.051)	(0.057)	(0.058)	(0.056)	(0.069)
$\triangle$ Public School	0.165**	0.172**	0.172**	0.174**	0.174**	0.197**	0.186**
	(0.061)	(0.060)	(0.060)	(0.061)	(0.063)	(0.064)	(0.071)
$\triangle$ HH Assets	-0.030	-0.036	-0.036	-0.038	-0.038	-0.038	-0.020
	(0.241)	(0.235)	(0.248)	(0.236)	(0.253)	(0.250)	(0.258)
$\triangle$ HH Shock (Covariate)	0.091	0.047	0.047	0.034	0.034	0.012	0.039
	(0.073)	(0.072)	(0.074)	(0.076)	(0.083)	(0.078)	(0.097)
Urban Area	-0.171^{+}	-0.133	-0.133^{+}	-0.120	-0.120	-0.068	-0.085
	(0.090)	(0.088)	(0.075)	(0.091)	(0.073)	(0.086)	(0.092)
Coastal Andhra	0.031	0.038	0.038	0.040	0.040	0.076	0.104
	(0.076)	(0.075)	(0.076)	(0.075)	(0.064)	(0.084)	(0.101)
W $\triangle$ CH Weight[2]						-0.002	-0.005
						(0.005)	(0.005)
W $\triangle$ CH Height[3]						0.012*	0.009*
						(0.005)	(0.004)
W $\triangle$ CH Siblings						0.010	0.058
						(0.112)	(0.120)
W $\triangle$ CH Work						-0.127*	0.001
						(0.081)	(0.139)
W $\triangle$ Public School						-0.178*	-0.143
						(0.090)	(0.105)
W $\triangle$ HH Assets						0.068	0.162
						(0.374)	(0.404)
W $\triangle$ HH Shock (Covariate)						0.200*	0.265
						(0.116)	(0.179)
Breusch-Pagan Test[4]	0.000	0.009					
LM Test[5]	0.000	0.005					
Observations	**731**	**731**	**731**	**731**	**731**	**731**	**731**

Notes:
1. CH: Child; HH: Household.
2. Child weight: Weight for age z-score.
3. Child height: Height for age z-score.
4. P-values (H_0: constant error variance).
5. P-values (H_0: no spatial autocorrelation).
6. Standard Errors in parentheses (White Standard Errors for OLS Robust).
7. $^{+}$ indicates significance at 10%; * at 5%; ** at 1%.

Table A-III: Results for **7** Nearest Neighbors: Writing

	OLS		OLS Robust	SAR	SHAC	Contextual Effects	
						OLS Robust	SHAC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
W $\triangle$ Writing Level	–	0.423**	0.423**	0.948**	0.948**	0.291**	0.793*
		(0.084)	(0.076)	(0.086)	(0.143)	(0.096)	(0.414)
$\triangle$ CH Weight[2]	-0.011**	-0.010**	-0.010**	-0.009*	-0.009*	-0.009**	-0.008*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)
$\triangle$ CH Height[3]	-0.007*	-0.008*	-0.008*	-0.009*	-0.009*	-0.009**	-0.009**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)
$\triangle$ CH Siblings	0.070	0.049	0.049	0.023	0.023	0.061	0.035
	(0.055)	(0.054)	(0.053)	(0.056)	(0.058)	(0.052)	(0.059)
$\triangle$ CH Work	-0.159**	-0.115*	-0.115+	-0.062	-0.062	-0.079	-0.060
	(0.053)	(0.053)	(0.051)	(0.057)	(0.055)	(0.058)	(0.068)
$\triangle$ Public School	0.161**	0.166**	0.166**	0.172**	0.172**	0.185**	0.197**
	(0.060)	(0.059)	(0.059)	(0.061)	(0.068)	(0.063)	(0.075)
$\triangle$ HH Assets	-0.030	-0.037	-0.037	-0.045	-0.045	-0.091	-0.074
	(0.236)	(0.232)	(0.245)	(0.238)	(0.260)	(0.249)	(0.268)
$\triangle$ HH Shock (Covariate)	0.091	0.061	0.061	0.024	0.024	0.053	-0.037
	(0.071)	(0.071)	(0.073)	(0.073)	(0.082)	(0.078)	(0.090)
Urban Area	-0.181*	-0.137	-0.137+	-0.081	-0.081	-0.078	-0.057
	(0.086)	(0.085)	(0.073)	(0.089)	(0.064)	(0.095)	(0.073)
Coastal Andhra	0.036	0.043	0.043	0.052	0.052	0.133	0.061
	(0.074)	(0.073)	(0.074)	(0.075)	(0.054)	(0.088)	(0.092)
W $\triangle$ CH Weight[2]						-0.000	0.009
						(0.008)	(0.009)
W $\triangle$ CH Height[3]						0.022**	0.024**
						(0.008)	(0.007)
W $\triangle$ CH Siblings						0.162	0.008
						(0.142)	(0.176)
W $\triangle$ CH Work						-0.257*	-0.098
						(0.106)	(0.159)
W $\triangle$ Public School						-0.197+	-0.208+
						(0.116)	(0.115)
W $\triangle$ HH Assets						0.402	0.272
						(0.496)	(0.419)
W $\triangle$ HH Shock (Covariate)						0.285+	0.130
						(0.158)	(0.192)
Breusch-Pagan Test[4]	0.009	0.016					
LM Test[5]	0.000	0.016					
Observations	**731**	**731**	**731**	**731**	**731**	**731**	**731**

Notes:

1. CH: Child; HH: Household.
2. Child weight: Weight for age z-score.
3. Child height: Height for age z-score.
4. P-values (H_0: constant error variance).
5. P-values (H_0: no spatial autocorrelation).
6. Standard Errors in parentheses (White Standard Errors for OLS Robust).
7. + indicates significance at 10%; * at 5%; ** at 1%.