Peer Effects in Consumption*

Punarjit Roychowdhury†
Southern Methodist University

February 10, 2017

Abstract

The recent literature has shown the existence of significant peer effects in consumption with much of the evidence coming from the developed economies. Whether peers play a significant role in influencing consumption of individuals in low income countries remains unclear. In this paper, using newly available household level data from India, I examine whether consumption of individuals in a low income country is influenced by that of their peers' as in the more prosperous countries. Using an instrumental variables/fixed effects approach, I find that a 1 Indian Rupee increase in average peer consumption expenditure causes households to increase their own consumption expenditure by 0.7 Indian Rupee. Falsification tests and robustness checks support the validity of my results. My findings suggest that policies that influence a household's consumption will also affect the consumption of the household's peers through social interactions. This implies traditional analyses of consumption intervention programs that do not take into account such spillover effects will understate the total social impact of the programs, and hence lead to inaccurate evaluation of cost-effectiveness of such programs.

Keywords: Consumption, India, Instrumental Variables, Peer Effects, Risk Sharing.
JEL Classification: C21, D12, O12, Z13.

*I am deeply indebted to Daniel Millimet for his advice and continued feedback. I am also grateful to Tim Salmon, Elira Kuka, session participants at the Southern Economic Association Annual Conference 2016, North American Summer Meeting of the Econometric Society 2016 and the Stata Texas Econometrics Camp 2016, and seminar participants at SMU and CSSSC for suggestions and comments. All remaining errors are my own.

†Address for Correspondence: Department of Economics, Southern Methodist Universty, 3300 Dyer Street, Umphrey Lee Center, Suite 301, Dallas, TX 75275-0496. Email: proychowdhur@smu.edu.

1 Introduction

The starting point of this research is the observation that peers play a significant role in influencing consumption behavior of individuals. This evidence comes from a small but growing body of literature on neighborhood and peer effects in consumption, mostly in developed countries (e.g., Grinblatt et al., 2008; Angelucci and De Giorgi, 2009; Kuhn et al., 2011; Moretti, 2011; Roth, 2015; De Giorgi et al., 2016).[1] What remains unclear is whether this is a universal social phenomenon or an artifact of affluent market oriented lifestyle.[2] Also, the question of why peers matter in shaping consumption choices has received little attention in the existing literature.

This paper fills these gaps in the literature on peer effects in consumption by addressing two (connected) questions, namely: (1) Does consumption of individuals' peers affect their own consumption in low income countries as in more prosperous countries; and (2) if it does, what is the mechanism driving the peer effects. I examine these questions using newly available household level data from India – which is home to one-third of the world's poor.[3]

There are at least two reasons why a careful analysis of peer effects in consumption in context of low income countries is crucial. First, understanding the magnitude and nature of peer effects is imperative for accurate evaluation of consumption intervention programs (e.g., consumption tax policies, conditional cash transfer programs, etc.) that are used massively in low income countries as developmental policies. This is because, if there are non-negligible peer effects, such policies, in addition to having a direct effect, would have an indirect spillover effect.[4] If this is not taken into account, the total 'social effect' of the policies would be underestimated.

Secondly, if consumption of individuals' peers affect their own consumption in a positive way, ceteris paribus, this would imply that the individuals must be lowering their savings or taking up loans to finance their increase in consumption when peer consumption rises. This is likely to magnify the risk of poor and middle income individuals (who make up the majority of the population in any low income country) of getting trapped in poverty (or severely hamper their ability to move out of poverty traps) and increase their economic vulnerability in the long run owing to 'under-saving' or 'over-borrowing' (Moav and Neeman, 2012).

[1] For early theoretical works on social infuences in consumption see Duesenberry (1949), Leibenstein (1950), Becker (1974) and Pollak (1976).

[2] According to Fafchamps and Shilpi (2008), the source of this suspicion perhaps goes back to the writings of Rousseau, who proposed the 'good savage' parable in response to Hobbes' depiction of the state of nature as anarchic and violent. Similar views have also later been expressed by Scott (1976) and also some others, who argue that hedonistic and market-based way of life exacerbates social comparisons.

[3] http://time.com/2999550/india-home-to-most-poor-people/

[4] Also referred to as 'social multiplier effect' (Akerlof, 1997; Glaeser et al., 2003).

Thus finding evidence of positive peer effects would highlight the importance of formulating innovative social policies that represses social pressure and using such policies in conjunction with traditional anti-poverty and redistributive policies in order to augment social welfare.

I begin by constructing a simple model of social interactions in consumption that allows for both endogenous peer effects (peer effects due to peer consumption) as well as exogenous peer effects (peer effects due to peer characteristics). In my model, peers' consumption expenditure is viewed as a *social norm* and households[5] pay a cost for deviating from this norm. In this model, each household wants to *conform* as much as possible to the social norm of its peer group, which is defined as the average consumption expenditure of its peers. Peer characteristics, on the other hand, are assumed to impact the marginal returns of own consumption. Based on this model, I derive the basic econometric specification which predicts a positive monotonic relationship between own and average peer consumption expenditure. I empirically test this prediction using data from the 2012 Indian Human Development Survey (Desai et al., 2015). I define a household's peer group as other households living in its village (for rural areas) or neighborhood (for urban areas) since "people almost certainly compare themselves to their immediate geographical neighbors" (Deaton, 2001, p. 21). Fafchamps and Shilpi (2008), in fact, note that since social mobility is very low in low income countries as a result of which people live along people they grew up with, immediate neighbors constitute almost a 'natural' peer group for people in living in these countries.[6]

As noted by Manski (1993), identification of endogenous peer effects is 'notoriously difficult'. More specifically, there are two econometric problems that hamper inferences about peer influences on individual behavior. The first problem that arises is due to a simultaneity bias. This bias is generated by a 'reflection problem' – the simultaneous determination of own and peer outcome (which is consumption in the present case). In general, the problem cannot be resolved by reduced-form modeling because the structural parameters are under-identified.

The second factor complicating identification and estimation of peer influences is an omitted variables bias. As noted by Grinblatt et al. (2008), in the absence of a perfect set of controls, one cannot validate a peer influence on consumption by observing that a group of neighbors purchase similar baskets of goods (or spend similar amounts of resources on consumption). Inferences will be biased whenever there are group level unobservables that are correlated with consumption expenditure of all those belonging to the group (i.e., correlated unobservables). Indeed, these unobservables may potentially be even correlated

[5] In this paper, I use the terms individual and household interchangeably.

[6] In the Supplementary Appendix, I consider an alternative definition of peer group based on caste affiliation and geographic location of households as a sensitivity check of my baseline analysis.

with the characteristics of the households that affect their consumption.[7]

In the present paper, I tackle these problems based on a strategy of instrumental variables (IV)/fixed effects. I create my instruments based on assumptions invoked in my theoretical model. Specifically, my model assumes that a household faces various idiosyncratic shocks (e.g., death of a household member, accident/injury, job loss, incident of crime, etc) that affects its own consumption, and some of these shocks are *observable* (to the econometrician). I further assume that these shocks are 'household specific'. In other words, it is only the own idiosyncratic shocks that affect own consumption and peer idiosyncratic shocks do not have any influence on own consumption. As argued by Helmers and Patnam (2014, p. 95), "this is a credible assumption given the idiosyncratic nature of the shocks".

The above assumptions allow me to instrument average peer consumption (which is the source of the simultaneity bias) by average peer observable idiosyncratic shocks, the data on which is, thankfully, available from the IHDS 2012. The intuition is that since own idiosyncratic shocks affect own consumption and do not contain any information about consumption of other households, to this extent, average peer idiosyncratic shocks should affect average peer consumption and that there should not be any effect of average peer idiosyncratic shocks on the target household's consumption after conditioning on own idiosyncratic shocks (to show that my instruments are plausibly exogenous, I carry out balancing tests (Bifulco et al., 2011; Lavy and Schlosser, 2011) and other standard IV diagnostic tests).

My IV strategy is similar to the 'spatial IV' method used widely in empirical spatial literature (for an overview see Gibbons and Overman, 2012; Gibbons et al. 2015). The strategy requires some/all exogenous characteristics of neighboring spatial units (the spatial units being households in my case) to be used as instruments for spatial interaction term which is endogenous. The major requirements for this strategy, thus, is to carefully justify why such characteristics might potentially affect a spatial unit's own outcome but not the outcome of its neighbors. This approach has recently been used in some papers in the peer effects literature including Gaviria and Raphael (2001), Goux and Maurin (2007), Fletcher (2010, 2012, 2015), Helmers and Patnam (2014) and McVicar and Polanski (2014).

The above discussed IV strategy allows me to obtain consistent estimates of endogenous peer effects if the group level unobservables are correlated only with the consumption of households. However, this strategy may not be sufficient to identify the endogenous peer effect if peer group unobservables are correlated with household characteristics or household specific idiosyncratic shocks. To address this concern, I include a full set of district fixed

[7] Another source of correlated unobservables is non-random sorting of households into peer groups since this would imply that unobservable characteristics of households are correlated with the characteristics of the group. However, as I argue in section 4, this is unlikely to be a cause of concern in the present case since social mobility is very low in India.

effects. In India, districts refer to an administrative division of states and each district is comprised of several 'similar' villages and/or neighborhoods. As long as the villages or neighborhoods within a particular district do not differ along unobservable dimensions that are correlated with household characteristics and household specific idiosyncratic shocks, the district fixed effects should be sufficient to resolve the omitted variable bias.

The IV/fixed effects-based identification strategy used in this paper not only allows one to clearly identify endogenous peer effects, but also has the advantage of being fairly flexible in terms of data requirement. Alternative methods, although novel and unique, are either unable to isolate endogenous peer effects from exogenous ones (e.g., Sacerdote, 2001; Graham, 2008; Ammermueller and Pischke, 2009), or do so at the cost of being extremely restrictive in terms of data requirement. For instance, the method proposed by Lee (2007) and developed later by Bramoulle et al. (2009) allows one to identify endogenous peer effects but requires peer groups to be 'small' on average and that there should be sufficient variation in peer group size. Again, using the empirical strategies proposed by Bramoulle et al. (2009), Calvo-Armengol et al. (2009) and Lee et al. (2010) to isolate endogenous peer effects from exogenous ones requires a researcher to be able to observe *all* social interaction links in the data (that is, the researcher must have very detailed network data). Such requirements are not met by most micro datasets, including the dataset used in the present paper.

My results are striking. In consonance with the prediction of social conformity theory, I find that average peer consumption has a significant positive impact on households' own consumption. More specifically, I find that an increase in average peer consumption expenditure by 1 Indian Rupee causes an average household to increase its own spending by 0.7 Indian Rupee. This translates into a social multiplier of about 3. Further, I find that the increase in own consumption due to a one standard deviation increase in average peer consumption is exactly equal to the increase in own consumption due to a one standard deviation increase in a household's own income. Thus, the estimated endogenous peer effect is not only economically significant in absolute terms, but is also remarkably strong compared to other major determinants of consumption.

To examine the robustness of my results, I conduct several robustness checks. First, I carry out a falsification test in which I create placebo peer groups. Second, I control for covariate (community) shocks that are likely to impact consumption of (almost) all households within a village/neighborhood. Third, I present IV results of additional specifications using different subsets of the vector of peer idiosyncratic shocks as instruments in the first stage and including the remaining peer shocks as covariates in the second stage. Fourth, I run my baseline regressions transforming consumption expenditure and income in logs. Fifth, I instrument household income by literacy status of the father of household' head (or father of

household head's husband). My baseline results remain robust throughout these exercises. Additionally, I carry out a disaggregate analysis and find that the increase in own consumption expenditure due to endogenous peer effect is driven by an increase in own expenditure on temptation goods and food.

An alternate explanation of why one's consumption might co-move with peers' consumption, particularly in a less developed country, is risk-sharing (Townsend, 1994). To show that this is not the mechanism driving my results, I do a couple of things. First, I rule out *full* risk sharing as a plausible mechanism by constructing a stylized model of full risk sharing and showing that my empirical findings do not corroborate the implications of this model. Second, I rule out *partial* risk sharing by analyzing different subsamples and showing that endogenous peer effects do not significantly vary by degree of households' exposure to risk and level of economic vulnerability.

As noted previously, much of the existing research on peer influences in consumption focuses on developed countries. One of the earliest empirical studies in this literature is that by Grinblatt et al. (2008). Using Finnish data, they provide evidence of endogenous peer effects in automobile purchases. Kuhn et al. (2011), using data from the Dutch Postcode Lottery (PCL), also find significant peer effects in car consumption. Moretti (2011), in an interesting paper, estimate peer effects and social multiplier in consumption of movies using high quality box-office data from the United States. De Giorgi et al. (2016) exploit detailed network data in a recent working paper to show the existence of non-negligible peer effects in aggregate consumption in context of Denmark.

The only two studies that I am aware of which focus on understanding neighborhood and peer effects in context of low income countries are those by Angelucci and De Giorgi (2009), and Roth (2015).[8] Angelucci and De Giorgi (2009), using data from a program that targets poor households in small rural communities in Mexico with bimonthly conditional grants to improve living standards, find strong evidence of positive program externalities on non-eligible households. They, however, estimate only a composite social interaction effect, and, unlike me, do not attempt to disentangle the endogenous peer effects from the exogenous ones. Roth (2015), on the other hand, uses a randomized conditional cash transfer program from Indonesia to analyze peer effects in visible consumption of poor households. He finds that the expenditure share of visible goods rises for untreated households in treated sub-districts, whose reference group visible consumption is exogenously increased.

[8] Ofcourse, there is separate body of literature that looks at how relative consumption and income affects subjective welfare of individuals in low income countries (see for e.g. Ravallion and Lokshin, 2005; Fafchamps and Shilpi, 2008). Also, in a recent paper Roychowdhury (2016), looks at the psychosocial impact of inequality on conspicuous consumption using data from India. These papers, however, are not concerned with quantifying or evaluating endogenous peer effects in consumption.

My work complements the existing literature on peer influences in consumption and contributes to it in several ways. First, this paper is among the very few papers that look at how social preferences affect aggregate consumption decision of households using a large cross-section data from a low-income country. As such, my findings are likely to have a high relevance for designing anti-poverty policies and safety nets. Second, the novelty of the dataset[9] used in this paper allows me to construct instruments for peer consumption which are credibly excluded from the regression equation in order to identify endogenous peer effects without having to assume zero exogenous peer effects.[10] This is a substantial improvement over many prior papers in the peer effects literature employing an IV strategy to estimate endogenous peer effects in absence of high dimensional social network data. Generally, these papers simply assume away all exogenous peer effects in the main equation predicting the outcome and use these as instruments for peer outcome variable (e.g., Gaviria and Raphael, 2001; McVicar and Polanski, 2014). However, as noted by Fletcher (2010), it is not very clear if any of their instruments are appropriately excluded from the main equation. Finally, this paper also explores whether the peer effects in consumption arise due to behavioral reasons or due to risk sharing. This is in contrast to most previous studies on this topic which focus mainly on quantifying the magnitude of peer effects. However, in a low income environment, given that risk sharing might potentially be extremely important, it is necessary to understand whether this is actually what is driving the results or not for effective policy design.[11]

The paper unfolds as follows. In section 2, I set out the theoretical model and derive my empirical specification. Section 3 discusses the data. Section 4 discusses the identification issues and presents the empirical strategy. The main results and the results of robustness checks are presented in Section 5. Section 6 explores whether my results could be explained by a model of risk sharing. The last section concludes.

[9] That is, the availability of information on household specific idiosyncratic expenditure shocks.

[10] In fact, being able to control for a host of exogenous peer effects also means that the concern of omitted variable bias emanating from group level unobservables (which may remain despite the inclusion of district fixed effects) gets considerably reduced.

[11] For instance, if risk sharing is the mechanism at work, then a policy like conditional cash transfer will almost certainly lead to welfare improvement in the sense that after the policy is implemented, everyone in the community (even those who may not have benefited from the policy) can increase their consumption without having to lower their savings or taking more loans. In contrast, if people want to "keep up" for behavioral reasons, then a conditional transfer might hurt those who are poor or belong to the middle class, and who have not benefited from the policy, since they would now have to save less or borrow more in order to spend more in order to keep up with the average consumption of the community which has increased due to the increase in consumption of those who received the cash transfer.

2 A Simple Model of Social Interactions in Consumption

In this section, I develop a theoretical framework of peer effects in consumption to motivate my empirics. I adjust a standard model of consumer choice with a cost imposed on the decision maker when deviating from peer group choice. I argue that this cost represents a social cost from not conforming to the group. Based on my theoretical model, I derive the main estimating equation. My model predicts that a household's consumption is positively related with that of its peers'.

Basic Setup. Suppose a finite set of households $\{i = 1, 2, ..., M\}$ is partitioned into non-overlapping social groups indexed by $r = 1, 2, ..., \overline{r}$. Let M_r be the rth social group of size m_r. I define household i's peer group as all other members in its social group. Let $M_{i,r}$ be individual i's group of peers of size $m_r - 1$. A peer is any other household whose consumption decision and personal characteristics may affect i's consumption decision.

Households in my framework live for one period and have a given endowment of income; let $y_{i,r}$ denote household i's income. Household i uses its income to finance its consumption $c_{i,r}$ and leave bequests $b_{i,r}$ for its future generations which is motivated by joy-of-giving. Even though in a narrow sense $b_{i,r}$ captures financial bequests, it can be interpreted as any investment (e.g., human capital) from current income that enhances the productive capacity of children (e.g., health, education).

Preferences. I assume households have common preferences given by the utility function

$$U_{i,r} = u(c_{i,r}, b_{i,r}; \mathbf{\Psi}_{i,r}) + s\left(|c_{i,r} - \overline{c}_{-i,r}|\right) \tag{1}$$

where $u(.)$ is a 'conventional' utility function which depends on $c_{i,r}$ and $b_{i,r}$ and $v(.)$ represents a social utility function or a deviation payoff function which measures the utility attributable to the deviation of own consumption from average peer consumption given by $\overline{c}_{-i,r} = \frac{1}{\acute{m}_r - 1} \sum_{j \in M_{i,r}} c_{j,r}$.[12] The component $\mathbf{\Psi}_{i,r}$ introduces the exogenous heterogeneity that captures the differences between households. Sources of such heterogeneity are household i's demographic characteristics, $\mathbf{x}_{i,r} = [x^1_{i,r}, x^2_{i,r}, ..., x^K_{i,r}]$, idiosyncratic preference shocks that affect wellbeing of household $i, \epsilon_{i,r}$, as well as average income and non-income demographic characteristics of household i's peers given by $\overline{y}_{-i,r}$ and $\overline{\mathbf{x}}_{-i,r} = [\overline{x}^1_{-i,r}, \overline{x}^2_{-i,r}, ..., \overline{x}^K_{-i,r}]$ respectively, where $\overline{y}_{-i,r} = \frac{1}{\acute{m}_r - 1} \sum_{j \in M_{i,r}} y_{j,r}$ and $\overline{x}^k_{-i,r} = \frac{1}{\acute{m}_r - 1} \sum_{j \in M_{i,r}} x^k_{j,r}$, $k = 1, 2, ..., K$.

[12] I assume there is no utility payoff associated with deviation of agents' bequests from that of their peers' because peers' bequests are generally not visible.

I assume that $u(.)$ satisfies all the standard properties of an utility function, i.e., $\partial u(.)/\partial c_{i,r} > 0$, $\partial u(.)/\partial b_{i,r} > 0$, $\partial^2 u(.)/\partial c_{i,r}^2 \leqslant 0$, $\partial^2 u(.)/\partial b_{i,r}^2 \leqslant 0$. Further, I assume that the deviation pay off is falling in the magnitude of deviation $|c_{i,r} - c_{-i,r}|$, i.e., $\partial s(.)/\partial |c_{i,r} - \overline{c}_{-i,r}| < 0$. This implies that each household wants to minimize the social distance between itself and its peers and that it suffers a loss in utility if it does not conform to the majority of its peers' consumption which can be viewed as a 'social norm'. The loss in utility due to non-conformity could be in terms of social sanctions, social penalty, a loss of popularity and a loss of reputation. This is the standard way economists have been modeling conformity (see, among others, Kandel and Lazear, 1992; Bernheim,1994; Akerlof, 1997; Glaeser and Scheinkman, 2001; Liu et al., 2014).[13]

Characterizing Preference Shocks. I assume that the idiosyncratic preference shocks that affect decisions of household i (and which are a part of $\Psi_{i,r}$) can be decomposed into two components as follows:

$$\epsilon_{i,r} = \mathbf{z}'_{i,r}\boldsymbol{\mu} + e_{i,r} \tag{2}$$

where $\mathbf{z}_{i,r}$ denotes a $J \times 1$ vector of 'observed' (or, reported) idiosyncratic shocks $z_{i,r}^j$, i.e., $\mathbf{z}'_{i,r} = [z_{i,r}^1, z_{i,r}^2, ..., z_{i,r}^J]$, and $e_{i,r}$ denote the 'unobserved' idiosyncratic shock. I will come back to this assumption later.

Maximization Problem. The problem of household i is to choose $c_{i,r}$ and $b_{i,r}$ by maximizing (1) subject to a income budget constraint. To simplify the maximization problem, let $u(.) = \left[\chi + \lambda b_{i,r} + \mathbf{x}'_{i,r}\boldsymbol{\varphi} + \tau\overline{y}_{-i,r} + \overline{\mathbf{x}}'_{-i,r}\boldsymbol{\kappa} + \mathbf{z}'_{i,r}\boldsymbol{\mu} + \varepsilon_{i,r}\right] c_{i,r}$ and $s(.) = -\frac{\phi}{2}\left(c_{i,r} - \overline{c}_{-i,r}\right)^2$, where $\lambda > 0$, $\boldsymbol{\varphi} \neq \mathbf{0}, \boldsymbol{\mu} \neq \mathbf{0}$, $\tau \neq 0$, $\boldsymbol{\kappa} \neq \mathbf{0}$ and $\phi > 0$.[14] Note that ϕ is the parameter describing *taste for conformity* – higher the value of ϕ (i.e., higher weight is assigned to social feeling), higher is households' loss in utility due to non-conformity.[15] Household i's maximization problem, thus, becomes:

$$\max_{c_{i,r}, b_{i,r}} \left[\chi + \lambda b_{i,r} + \mathbf{x}'_{i,r}\boldsymbol{\varphi} + \tau\overline{y}_{-i,r} + \overline{\mathbf{x}}'_{-i,r}\boldsymbol{\kappa} + \mathbf{z}'_{i,r}\boldsymbol{\mu} + e_{i,r}\right] c_{i,r} - \frac{\phi}{2}\left(c_{i,r} - \overline{c}_{-i,r}\right)^2 \tag{3}$$

$$\text{subject to } c_{i,r} + b_{i,r} \leq y_{i,r}; c_{i,r}, b_{i,r} \geq 0 \tag{4}$$

[13] Another interesting paper is that of Clark and Oswald (1998) who propose a choice-theoretical justification for the conformist model.

[14] Note that the specific form of the deviation pay-off function has also been used by Akerlof (1997).

[15] For simplicity I have assumed that everyone has equal taste for conformity. However in reality, taste for conformity parameter could be household specific.

Estimating Equation. The first order condition of the above maximization problem yields the best response function for household i which is my basic estimating equation:

$$c_{i,r} = \alpha + \beta \overline{c}_{-i,r} + \gamma y_{i,r} + \mathbf{x}'_{i,r}\boldsymbol{\varpi} + \overline{\mathbf{x}}'_{-i,r}\boldsymbol{\delta} + \theta \overline{y}_{-i,r} + \mathbf{z}'_{i,r}\boldsymbol{\rho} + \xi_{i,r} \tag{5}$$

where $\alpha = \chi/(\phi + 2\lambda)$, $\beta = \phi/(\phi + 2\lambda)$, $\gamma = \lambda/(\phi + 2\lambda)$, $\boldsymbol{\varpi} = \boldsymbol{\varphi}/(\phi + 2\lambda)$, $\delta = \boldsymbol{\kappa}/(\phi + 2\lambda)$, $\theta = \tau/(\phi + 2\lambda)$, and $\rho = \boldsymbol{\mu}/(\phi + 2\lambda)$. $\xi_{i,r}$ is the unobserved error term.

The coefficient β is called the *endogenous peer effect*. It measures the impact of a household's peer consumption on its own consumption. The *exogenous peer effects* or *contextual effects* are captured by $\boldsymbol{\delta}$ and θ. These measure the impact of peer characteristics and peer income on household consumption. The *individual effects*, or the impact of own income and own characteristics, are captured by γ and $\boldsymbol{\varpi}$. Finally, ρ measures the effect of idiosyncratic shocks that a household faces on its own consumption. Since $0 < \beta < 1$, my model predicts a positive relationship between own consumption and peer consumption. The focus of this paper is to test this prediction.[16]

The estimation of Equation (5) faces several challenges. First, average peer consumption is endogenous due to simultaneity. Second, there might be unobserved aggregate shocks that are correlated with average peer consumption and possibly with other regressors and/or there might be non-random sorting of individuals into social groups – both of which would bias the estimates of endogenous and exogenous peer effects. Nonetheless, I leave the discussion of these issues for section 4 and now I turn to describe my data.

3 Data

3.1 The Indian Human Development Survey 2012

The Data. The Indian Human Development Survey (IHDS) 2012 is a nationally representative multitopic household survey conducted by the National Council for Applied Economic

[16] In this paper, I have modeled social utility as arising from social norms mainly because people in low/middle income countries need to heavily rely on social networks and peer groups due to missing markets and lack of properly functioning formal institutions, and hence must abide by social norms to avoid social penalty. However, it is worth noting that, in addition to conforming with social norms, in principle, doing better than the average might also generate positive social utility. More precisely, the social utility function might be of the form: $s(c_{i,r}, \overline{c}_{-i,r}) = \theta[-\frac{\phi}{2}(c_{i,r} - \overline{c}_{-i,r})^2] + (1-\theta)v(c_{i,r} - \overline{c}_{-i,r})$, where $0 < \theta < 1$, $v' > 0$ and $v'' < 0$ (i.e., diminishing marginal utility from getting ahead of others holds). Even with this type of a social utility function, it can be shown that the final comparative static/prediction remains unchanged. In terms of empirical modeling, identification and policy implications, there would also be no change. The only change in this case would be that the endogenous peer effect would need to be partly attributed to social conformity and partly to getting ahead of others'. In the Supplementary Appendix, I present a generalized version of the theoretical model showing this formally.

Research (NCAER) in New Delhi and University of Maryland (Desai et al., 2015). It was designed to complement existing Indian household surveys by bringing together a wide range of topics in a single survey. This breadth permits analyses of associations across a range of social and economic conditions. The sample was drawn using stratified random sampling.[17]

The IHDS 2012, conducted between November 2011 and October 2012, covers 42,152 households in 1420 villages and 1042 census-defined urban neighborhoods located throughout India.[18] The survey covered all the states and union territories of India except Andaman and Nicobar, and Lakshadweep. These two account for less than 0.05 percent of India's population. The data is publicly available from the Data Sharing for Demographic Research program of the Inter-university Consortium for Political and Social Research (ICPSR).[19]

The IHDS 2012 is based on two one-hour interviews with a knowledgeable informant in each household. The interviews covered health, education, employment, economic status, marriage, fertility, gender relations and social capital of the households. The survey instruments were translated into 13 Indian languages and were administered by local interviewers. The main advantage of using the IHDS is that it includes many questions that are not asked in the larger and more commonly used Indian household survey, the National Sample Survey (NSS). In particular, detailed questions on income, consumption expenditure, and expenditure shocks are asked in the IHDS 2012. It also has a broad array of demographic and family questions which permits me to include a variety of control variables. Moreover, the survey provides the precise geographic location of households (i.e., one can identify a household's village or neighborhood from the survey). All these features make the use of IHDS 2012 attractable for the present research.

Peer Group. Like most other household surveys, IHDS does not have precise social interactions data. Deaton (2001) suggests that, in absence of such data, the most sensible social groups of households are those that live in the immediate geographic location. Fafchamps and Shilpi (2008), in fact, suggest that in context of low income countries immediate neighbors is likely to constitute a 'natural' reference group. This is because social-psychologists have shown that, when making relative consumption assessments, people compare them-

[17] An earlier wave of the data (IHDS 2005) is available. However, it does not contain information on some aspects (e.g., household specific idiosyncratic shocks) that are necessary for my analysis.

[18] According to the Indian National Census (2011a), the definition of urban area is as follows: (1) All places with a municipality, corporation, cantonment board or notified town area committee, etc.; (2) All other places which satisfied the following criteria: (i) A minimum population of 5,000; (ii) At least 75% of the male main working population engaged in non-agricultural pursuits; and (iii) A density of population of at least 400 persons per sq. km. The urban neighborhoods are defined as per the Indian Census. Each decade, the Census draws neighborhoods of approximately equal size (about 200 households) in all Indian towns and cities (urban areas).

[19] http://www.icpsr.umich.edu/icpsrweb/DSDR/studies/36151

selves with a peer group composed of people who started from the same conditions, e.g., those they grew up with. And since in India social mobility is remarkably low (Ravallion and Lokshin, 2005; Munshi and Rosenzweig, 2009) meaning that most people live along people they grew up with, "people almost certainly compare themselves to their immediate geographical neighbors" (Deaton, 2001, p. 21). Also, as noted by Akerlof (1997), social interactions are more likely to take place among people living in the same locality which may in turn affect household decision making. Due to these reasons, I assume that the villages (in rural areas) and neighborhoods (in urban areas), which basically are small geographic units[20] populated by households who are similar in many dimensions and are exposed to similar geographic and institutional conditions, form the relevant social groups,[21] The peer group or the reference group of a household is, therefore, comprised of all other households in its village or neighborhood.[22]

3.2 Variables and Sample Summary Statistics

Outcome Variable. The aim of the paper is to examine social interactions in consumption. The key outcome variable is, therefore, household consumption expenditure. There are fifty-two consumption categories in the IHDS 2012 (see Table A1 in the Supplementary Appendix for a complete list of consumption categories). Thirty of the consumption categories, which are frequently purchased items, use a thirty day time frame while the other twenty-two use a three hundred and sixty five day time frame. I convert all expenditures to the annual time frame. The sum of the expenditure of households on all these consumption categories form the household consumption expenditure.

Demographics. My theoretical model requires me to control for household characteristics that includes household current income and other demographic variables. Fortunately, unlike most other household surveys, IHDS 2012 has very good data on current household income. The current income of households reported in the survey is the sum total (for each household) of wages and salaries, non-farm business income, net agricultural income, remittances, property and other income and public benefits. Each of these incomes are in turn constructed from more than fifty different sources of income queried in the survey.[23]

[20] The average area of villages included in the IHDS is approximately 3.3 sq. miles. While the average area of urban neighborhoods are not available from the survey, these are also likely to be reasonably small since, as mentioned previously, the average number of households residing in a census-defined urban neighborhood is roughly 200.

[21] Implicitly, the assumption here is that neighborhoods are urban counterparts of rural villages.

[22] Cojocaru (2014) summarizes various empirical studies confirming that peer or reference groups that are used by individuals for social comparisons are indeed local.

[23] Although the IHDS takes considerable care in designing a questionnaire that enumerates several sources

The set of demographic variables can be classified into two categories: characteristics of household heads and socioeconomic features of households. Characteristics of household heads include age, gender, marital status, and literacy status (whether literate or not). Also a set of dummy variables indicating caste affiliation of the household head are included: Brahmin, non-Brahmin forward caste, other backward caste (OBC), Scheduled Castes (SC), Scheduled Tribes (ST) and others. The socioeconomic features of households that are used as controls are: household size, number of years they have been living in their current village or neighborhood, proportion of children, adolescents and adults in the household, binary variables indicating the number of married household members and a binary variable indicating whether the geographical location of the household is classified as urban or rural as per the Indian National Census (2011a).

Idiosyncratic Shocks. Additionally, I also need to control for idiosyncratic shocks that potentially might impact household consumption expenditure. As noted before, IHDS 2012 has several questions on expenditure shocks, i.e., events that have impacted households' consumption expenditure in the recent past. Most of these events are likely to be unforeseen and hence might be considered random. This is an unique feature of the IHDS 2012 compared to other similar household surveys. As I will show later, the availability of data on expenditure shocks in the IHDS 2012 is extremely important for the present study not only because my theoretical model requires me to control for idiosyncratic shocks, but also because my identification strategy (which I shall discuss later) crucially hinges on being able to find data on these. My vector of observable idiosyncratic household specific consumption shocks consist of the following events that have been reported by the knowledgeable informant of the households as to have impacted their consumption expenditure: death of a household member, loss of job in the household, incident of theft and incident of break-in. Here, I take care to include only those shocks that are specific to any given household and are not correlated with the occurrence of the same shock in other households.[24]

Analytic Sample. My estimation sample consists of 40,980 households from 1,411 rural villages and 1,039 urban neighborhoods located across 375 districts: these are households in

of income, measurement errors in income cannot be ruled out. I address this issue further in Section 5.2.

[24] By definition idiosyncratic shocks are different from community shocks. While community shocks generally affect most or all members of the households within a particular community, idiosyncratic shock hits only a few of them. In other words, while the proportion of households who are likely to be hit by a community level shock is close to 1, this figure is likely to be much smaller for idiosyncratic shocks. My sample reveals that, on an average, the proportion of households in a village/neighborhood in which there is a death is 0.19, the proportion of households that suffer a job-loss is 0.03, the proportion of households that report an incident of theft is 0.04, and the proportion of household that report an incident of break-in is 0.02. This in line with the notion of idiosyncratic shocks

the IHDS 2012 where I have individual level information for household heads and for which the household head is above 18 years of age, household current income is more than zero but less than Rs. 1,000,000 (equivalent to $15,385), total consumption expenditure is more than zero, information on other household characteristics is non-missing and finally the household lives in a village or neighborhood with not less than three (sampled) members.

Table 1 presents the description and summary statistics of all the variables used in this study.

4 Empirics

4.1 Identification Issues

There are two main challenges in identifying and estimating peer effects based on my basic empirical specification given by Equation (5).

Simultaneity. Firstly, since household i's consumption affects its peers' mean consumption and vice versa, one cannot distinguish if a group member's action is the cause or the effect of peers' influence. This in turn will imply that β is subject to endogeneity bias. Manski (1993) labeled this the 'reflection problem'.

To illustrate this problem, I consider a setting where each group $r = 1, 2, ...\overline{r}$ comprises of only two households, household 1 and 2. The consumption equations are then:

$$c_{1,r} = \alpha + \beta c_{2,r} + \gamma y_{1,r} + \mathbf{x}'_{1,r}\varpi + \mathbf{x}'_{2,r}\boldsymbol{\delta} + \theta y_{2,r} + \mathbf{z}'_{1,r}\rho + \xi_{1,r} \quad (6)$$

$$c_{2,r} = \alpha + \beta c_{1,r} + \gamma y_{2,r} + \mathbf{x}'_{2,r}\varpi + \mathbf{x}'_{1,r}\boldsymbol{\delta} + \theta y_{1,r} + \mathbf{z}'_{2,r}\rho + \xi_{2,r} \quad (7)$$

Observe that household 1 (household 2)'s unobserved trait $\xi_{1,r}$ ($\xi_{2,r}$) is correlated with household 2 (household 1)'s consumption $c_{2,r}$ ($c_{1,r}$) through household 1 (household 2)'s consumption if there is endogenous peer effect (i.e., $\beta \neq 0$). As such, estimation of Equation (5) via OLS will yield biased parameter estimates because unbiased estimation of the coefficients requires that the error term be uncorrelated with the right-hand-side variables.

Correlated Unobservables. Secondly, even though I have a comprehensive vector of group level controls (i.e., peer characteristics), there might still be certain unobservable environmental attributes that are specific to social groups and/or common to all members of a particular group. That is, the unobserved error term might be of the following form:

$$\xi_{i,r} = \upsilon_r + \varepsilon_{i,r}, \tag{8}$$

where υ_r is the vector of group specific unobserved characteristics common to all group members (e.g., same motivation towards consumption or similar credit constraints), and $\varepsilon_{i,r}$'s are innovations. It is evident that υ_r's are correlated with mean peer consumption expenditure (to see this substitute Equation (8) in Equation (6) and Equation (7)). In fact, the village/neighborhood level unobserved effects may, in theory, be correlated with household characteristics, peer characteristics as well as with household specific idiosyncratic shocks. Econometrically, this would imply existence of a non-zero correlation between the group unobservables υ_r and one or more regressors in Equation (5). If there are such unobserved heterogeneity across social groups, then estimates of peer effects will be biased.

Additionally, the problem of correlated unobservables could arise if households self-select into social groups with specific objectives (Falk and Knell, 2004).[25] One way of doing this is typically via migration or residential relocation (Stark and Taylor, 1991). For instance, a poor person living in a prosperous neighborhood, to reduce his/her feeling of relative deprivation, might want to relocate to a less prosperous neighborhood. Frequently there is such positive selection in which 'similar' people join or are assigned to the same group (Sacerdote, 2011). This non-random sorting potentially implies that unobserved individual characteristics are correlated with the characteristics of the group which could cause substantial upward bias in the estimated magnitude of the endogenous and contextual peer effects. However, this is not a cause of concern in my case, given that the spatial mobility is extremely low in India (Ravallion and Lokshin, 2005; Munshi and Rosenzweig, 2009). In fact, the data that I use in this paper also shows that around 97 percent of the sampled households have been living in the same place for more than 10 years.

4.2 Empirical Strategy

To identify endogenous peer effects, I implement an approach based on IV/fixed effects. My IV strategy, in essence, is similar to the method of 'spatial IV' used widely in empirical spatial literature (see Gibbons and Overman (2012) and Gibbons et al. (2015) for an overview). Not only does this strategy offers the advantage of computation ease, IV estimation remains consistent even in the presence of spatially correlated error terms (Kelejian and Prucha, 1998; Brueckner, 2003). In addition, I try to address the potential problem of unobserved heterogeneity across social groups – that is not taken care by my IV strategy – by including

[25] Nesse (2004), for instance, argues that motivated to satisfy particular psychological desires, individuals can create their own social groups.

district fixed effects.

IV Strategy. I begin by describing my IV strategy. To start with, I assume that village/neighborhood unobserved effects (if present) are uncorrelated with household characteristics and observable shocks. I shall relax this assumption later.

To see how my approach works, I express my baseline model given by Equation (5) in matrix notations. For expository purposes, I assume that own income and peer income do not affect consumption (i.e., $\gamma = \theta = 0$). Also assume that households have unique characteristics and face unique observable idiosyncratic shocks (i. e., $K = J = 1$). Recall that in my framework, I have assumed that households are affected by all other households in their social group and by none outside it. This means that the observed social interactions can be modeled as an $M \times M$ block-diagonal matrix $\mathbf{G}$ specified as

$$\begin{aligned} \mathbf{G} &= diag(\mathbf{G}_1, \mathbf{G}_2, ..., \mathbf{G}_{\overline{r}}) && (9) \\ \mathbf{G}_r &= \frac{1}{m_r - 1}\left(\boldsymbol{\iota}_{m_r}\boldsymbol{\iota}'_{m_r} - \mathbf{I}_{m_r}\right), \ r = 1, 2, .., \overline{r} && (10) \end{aligned}$$

where M denotes total number of households, $\overline{r}$ is the number of social groups, m_r is the group size of group r, $\boldsymbol{\iota}_{m_r}$ is the m_r-dimensional vector of ones, and $\mathbf{I}_{m_r}$ is the m_r-dimensional identity matrix. Thus Equation (5) can be rewritten in matrix form as follows:

$$\mathbf{C}_r = \alpha\boldsymbol{\iota}_{m_r} + \beta\mathbf{G}_r\mathbf{C}_r + \mathbf{X}_r\varpi + \delta\mathbf{G}_r\mathbf{X}_r + \mathbf{Z}_r\rho + \boldsymbol{\xi}_r, \qquad E(\boldsymbol{\xi}_r|\mathbf{X}_r, \mathbf{Z}_r, \mathbf{G}_r) = 0 \tag{11}$$

where $\mathbf{C}_r$, $\mathbf{X}_r$ and $\mathbf{Z}_r$ are $m_r \times 1$ vectors of consumption, individual characteristics and idiosyncratic shocks respectively, and $\boldsymbol{\xi}_r$ is the vector of the composite error term $\boldsymbol{\iota}_{m_r}\upsilon_r + \boldsymbol{\varepsilon}_r$. Note that Equation (11) is similar to a spatial autoregressive (SAR) model (e.g., Cliff and Ord, 1981).

Concatenating Equation (11) over all groups yields

$$\mathbf{C} = \alpha\boldsymbol{\iota} + \beta\mathbf{G}\mathbf{C} + \mathbf{X}\varpi + \delta\mathbf{G}\mathbf{X} + \mathbf{Z}\rho + \boldsymbol{\xi} \tag{12}$$

where $\mathbf{C}$ (respectively $\boldsymbol{\iota}$, $\mathbf{X}$ and $\mathbf{Z}$) is obtained by stacking vectors $\mathbf{C}_r$ (respectively $\boldsymbol{\iota}_{m_r}$, $\mathbf{X}_r$ and $\mathbf{Z}_r$) for $r = 1, 2, ..., , \overline{r}$. Note that since my theoretical framework implies $\beta < 1$, $\mathbf{I} - \beta\mathbf{G}$ is invertible. I can, thus, write Equation (12) as:

$$\mathbf{C} = \alpha\left(\mathbf{I} - \beta\mathbf{G}\right)^{-1}\boldsymbol{\iota} + \left(\mathbf{I} - \beta\mathbf{G}\right)^{-1}\left(\mathbf{X}\varpi + \delta\mathbf{G}\mathbf{X}\right) + \left(\mathbf{I} - \beta\mathbf{G}\right)^{-1}\mathbf{Z}\rho + \left(\mathbf{I} - \beta\mathbf{G}\right)^{-1}\boldsymbol{\xi} \tag{13}$$

Note that since $(\mathbf{I}-\beta\mathbf{G})^{-1}=\sum_{k=0}^{\infty}\beta^k\mathbf{G}^k$, I can rewrite Equation (13) as

$$\mathbf{C}=\alpha/(1-\beta)\boldsymbol{\iota}+\varpi\sum_{k=0}^{\infty}\beta^k\mathbf{G}^k\mathbf{X}+\delta\sum_{k=0}^{\infty}\beta^k\mathbf{G}^{k+1}\mathbf{X}+\rho\sum_{k=0}^{\infty}\beta^k\mathbf{G}^k\mathbf{Z}+\sum_{k=0}^{\infty}\beta^k\mathbf{G}^k\boldsymbol{\xi} \tag{14}$$

Further, notice that from Equation (14) the expected mean peer consumption can be written as:

$$\mathbf{E}[\mathbf{GC}\ |\mathbf{X},\mathbf{Z}]=\alpha/(1-\beta)\boldsymbol{\iota}+\rho\mathbf{GZ}+\rho\beta\sum_{k=0}^{\infty}\beta^k\mathbf{G}^{k+2}\mathbf{Z}+\varpi\mathbf{GX}+(\varpi\beta+\delta)\sum_{k=0}^{\infty}\beta^k\mathbf{G}^{k+2}\mathbf{X} \tag{15}$$

This implies that one can use $\mathbf{GZ}$ to instrument for $\mathbf{GC}$.[26] In other words, average peer observable idiosyncratic shocks serve as valid instruments for average peer consumption. Note that my IV strategy exploits the fact that in my theoretical model, the idiosyncratic shocks are assumed to be household specific and that they *do not* contain any information about the consumption of other households, even those located in the same village/neighborhood. As argued by Helmers and Patnam (2014, p. 95), "this is a credible assumption given the idiosyncratic nature of the shocks". Since this assumption implies that average peer idiosyncratic shocks affect average peer consumption and that average peer idiosyncratic shocks have no direct impact on households' own consumption (other than through peer consumption), these could potentially be used as instruments for average peer consumption. My first stage regression, therefore, becomes:

$$\overline{c}_{-i,r}=\Pi_o+\Pi_1 y_{i,r}+\mathbf{x}'_{i,r}\boldsymbol{\Pi}_2+\overline{\mathbf{x}}'_{-i,r}\boldsymbol{\Pi}_3+\Pi_4\overline{y}_{-i,r}+\mathbf{z}'_{i,r}\boldsymbol{\Pi}_5+\overline{\mathbf{z}}'_{-i,r}\Pi_6+u_{i,r} \tag{16}$$

In Equation (16), $\overline{\mathbf{z}}'_{-i,r}$ denotes the vector of average observable peer idiosyncratic shocks where the idiosyncratic shocks are chosen from the events reported by the knowledgeable informant of the households in 2012 IHDS to have had affected their consumption expenditure.

Fixed Effects. The IV strategy described above is sufficient to produce consistent parameter estimates of my baseline econometric model under the null hypothesis of no correlated

[26] In principle, the higher order spatial lags of $\mathbf{Z}$ and $\mathbf{X}$ – $\mathbf{G}^2\mathbf{Z}, \mathbf{G}^3\mathbf{Z}$,...., $\mathbf{G}^2\mathbf{X}, \mathbf{G}^3\mathbf{X}$, ... – can also be used instruments along with $\mathbf{GZ}$. However, as noted by Gibbons and Overman (2012), the higher order lags do not work well as instruments in practice and gives rise to 'weak instruments/identification' problem. This is because there is little independent variation (and hence little additional information) in the higher order spatial lags of $\mathbf{Z}$ and $\mathbf{X}$, conditional on $\mathbf{GZ}$ and $\mathbf{GX}$.

effects (i.e., $\upsilon_r = 0$) or when there are group specific unobserved traits correlated with only consumption expenditure of the households (assuming that the IVs are valid). But what happens if village or neighborhood level unobservables are allowed to be correlated with households' own characteristics and/or idiosyncratic shocks? Unfortunately, the above described IV strategy may not work when there are such correlations. This is because if group level unobservables are correlated with households' observed idiosyncratic shocks, then the average peer observed idiosyncratic shocks are correlated with the error term (since the group level unobservables are a part of the error term). This, in turn, would imply that my instruments are no longer exogenous. This might also occur also if the group unobservables are correlated with one or more household characteristics. For instance, suppose that income of every household is correlated with unobserved village/neighborhood level characteristics as well as with household specific idiosyncratic shocks. This means average peer income will be correlated with group level unobservables as well as with average observed peer shocks. This, in turn, is likely to imply that average observed peer idiosyncratic shocks will be correlated with group level unobservable which again violates the exogeneity requirement of the instrument.

To alleviate this concern, I incorporate a full set of district fixed effects. Districts, which represent administrative divisions of an Indian state, are clusters of several villages/neighborhoods located in the same geographical area.[27] The logic behind including the district fixed effects is that these would capture the unobserved heterogeneity at the level of districts. Since villages/neighborhoods within a particular district are 'similar',[28] the incorporation of the district fixed effects along with the fact that I have a comprehensive vector of group level controls (peer characteristics) should be sufficient for the above discussed IV strategy to produce consistent parameter estimates.[29] The assumption is that, given that I have a comprehensive vector of group level attributes (i.e., contextual effects) in my baseline econometric equation, the remaining group-level heterogeneity (if any) does not vary across villages/neighborhoods within a district. In the section on robustness checks, I check the sensitivity of my baseline results to the inclusion of some covariate (community) shocks (i.e., shocks that are likely to affect most households within a village/neighborhood) as controls.

[27] In India, as of 2014, there are 29 states and, on an average, there are 23 districts in each state. The average area of a district is 1775 sq. miles.

[28] In India, districts are divided on the basis of ethical, cultural and social interaction rather on the basis of easiness or prosperity (Indian National Census, 2011b). As such, villages/neighborhoods within a particular district are likely to be similar along observable and unobservable ethical and cultural dimensions.

[29] My identification strategy ensures that it is not the 'price effect' that is driving my result. This is because my IV strategy will produce consistent parameter estimates even when there are unobserved differences in prices of goods across villages/neighborhoods. In fact, if prices are the *only* omitted village/neighborhood level characteristics, I do not even need to use district fixed effects since prices, presumably, are uncorrelated with the demographic characteristics (and shocks) of households.

Validity of Identification Strategy. While in principle the above described IV/fixed effects strategy allows me to consistently estimate the parameters of the baseline econometric model given by Equation (5), in practice the success of it depends on whether the set of instruments, $\overline{\mathbf{z}}_{-i,r}$, constructed using the expenditure shocks chosen from IHDS 2012 satisfy two crucial conditions implied by the theoretical model. First, a testable condition is that average peer idiosyncratic shocks must be correlated with average peer consumption (which should hold as long as the shocks affect household consumption expenditure). Second, the idiosyncratic shocks hitting the peers of household i should affect the household's own consumption *only* indirectly through their impact on its peers' consumption. This is tantamount to saying that average peer idiosyncratic shocks must be uncorrelated with the error term. This second citerion is an untestable maintained assumption.

To examine whether the chosen shocks satisfy the first condition, I carry out the Kleibergen-Paap rk LM test (2006). The LM test seeks to test whether that the excluded instruments are correlated with the endogenous regressors. The null hypothesis of this test is that the minimum canonical correlation between the endogenous variables and the instruments is not statistically different from zero. Rejection of the null hypothesis indicates that the model is identified. Further, since IV estimates based on 'weak' instruments are biased towards OLS estimates (Bound et al., 1995; Staiger and Stock, 1997; Stock et al., 2002) I report the F-statistic from the first stage regressions which is the test to examine strength of instruments. According to Staiger and Stock (1997) and Stock and Yogo (2005), an F-statistic value of 10 (or higher) implies rejection of the null hypothesis of weak instruments.

While, as noted above, it is in general not possible to test whether the instruments are uncorrelated with the error term, I carry out *balancing tests* (Bifulco et al., 2011; Lavy and Schlosser, 2011) to assess the likelihood of my instruments being uncorrelated with the unobserved error term. This test seeks to assess the correlation between the instruments and unobserved error term based on the correlation between the instruments and observed characteristics of households.[30] If the instruments are found to be uncorrelated with observable household characteristics related to consumption (more than what would be expected by chance), they may also be uncorrelated with the unobservable factors related to consumption (that are in the error term), following the logic of Altonji et al. (2005). As argued by Fletcher (2010, 2012, 2015), this is suggestive, but not conclusive, that the instruments can be treated as plausibly exogenous (or random 'shocks').

Further, since my model is over identified, I carry out the Hansen J test (1982), which

[30] I perform this test by regressing households' own background characteristics (income, household size, etc.) on peer shocks (instruments) controlling for district fixed effects, plus own shocks, and peer background characteristics.

is an overidentification test designed to examine the validity of the instruments (in terms of satisfying the exogeneity restriction). This test seeks to obtain multiple estimates of the treatment effect (which in the present case are the endogenous peer effect) based on various subsets of the instrumental variables and tests whether the obtained treatment effects are same. If all instruments are uncorrelated with the error term, all subsets should (asymptotically) return the same estimate of the treatment effect. The joint null hypothesis of this test is that the instruments are valid instruments (i.e., uncorrelated with the error term) and that the excluded instruments are correctly excluded from the estimating equation. A rejection of the null hypothesis casts doubt on the validity of the instruments. However, failure to reject the null does not necessarily mean that the exclusion restriction holds.

Model Estimation. I estimate my baseline model by the technique of Generalized Method of Moments (GMM). Since my model is overidentified, I report the two-step GMM estimates or optimal GMM estimates, which is the most efficient GMM estimator for overidentified models with heteroscedastic errors of unknown form (for a detailed overview of the two-step GMM see Cameron and Trivedi, 2005 and Baum et al., 2007).

5 Results

5.1 Main Results

OLS Results. Table 2 reports the naive OLS results. These estimates provide a useful benchmark with which to compare the results from the IV method. Column (1) regresses household consumption on average peer consumption, household income, household characteristics and household specific idiosyncratic shocks. Column (2) extends the set of regressors to include contextual effects. Column (3) adds district fixed effects. Note that, I report only the point estimates of endogenous peer effect for every specification. Full results are relegated to the Supplementary Appendix due to space constraints (see Table A2).

Across all specifications, I find a strong association between own and peer consumption with the coefficient of average peer consumption varying between 0.2 and 0.6. Moreover all the coefficients are statistically significant at 1% level of significance. These estimates are not, however, what I would wish to take seriously, given the identification issues discussed above, which OLS fails to resolve.

IV Estimates. Before presenting results using instrumental variables, I present the results from the balancing tests (Bifulco et al., 2011; Lavy and Schlosser, 2011) to show that the preferred instrumental variables are plausibly exogenous. In Table 3, I show that the in-

strumental variables are conditionally uncorrelated with all household characteristics. This evidence is suggestive, but not conclusive, that the instruments may also be uncorrelated with unobservable factors related to household consumption (which are in the error term), following the logic of Altonji et al. (2005), which is a maintained untestable assumption for the analysis to be valid. Overall, these results add confidence that the instruments are quasi-random within districts.

The IV results are reported in Table 4. As in the case of OLS, I report only the estimates of endogenous peer effects for each specification. Full IV results (including the first stage) are presented in Table A3 in the Supplementary Appendix.

I start by describing the performance of IV/fixed effect models in terms of the diagnostic tests. First, notice that all the specifications perform remarkably well in terms of the Hansen (1982) overidentification test, Kleibergen-Paap (2006) rk LM test for underidentification as well as the F-test for excluded instruments to assess the strength of the instruments. Specifically, for all the specifications, based on the Hansen J statistic, I am strongly unable to reject the joint null hypothesis that the instruments are uncorrelated with the error term and that the excluded instruments are correctly excluded from the estimated equation. Next, across all the specifications reported, the estimated Kleibergen-Paap rk LM statistic allows me to clearly reject the null hypothesis that the instruments are uncorrelated with the endogenous regressors and that the model is not identified. Finally, the first stage F-statistic for excluded instruments lies well above 10 across all the specifications in each panel, which clearly indicates that none of the specifications suffer from the weak instrument problem.[31]

Next, I turn to the actual two-step GMM estimates of the endogenous peer effects in consumption. Strikingly, across all specifications, I find strong evidence of endogenous peer effects in consumption. This is in line with the OLS results. Specifically, column (1) shows that the estimated impact of peers' consumption on own consumption based on the regression that neither includes contextual effects nor district fixed effects is 0.78 (s.e. $= 0.147$). Column (2) shows that the estimated impact of endogenous peer effects based on the regression that adds contextual effects (but not district fixed effects) to the set of regressors is 0.85 (s.e. =

[31] The first stage IV results reported in Table A3 in the Supplementary Appendix also show that the individual instruments (peer idiosyncratic shocks) are significantly correlated with the endogenous variable (peer consumption expenditure). That, for my preferred specification (column 5), these correlations are positive – conditional on the fact the peer income is held constant in the first stage – is not surprising: a death in the household in India is followed by a huge social ceremony that entails a huge expenditure. Similarly, a theft or breakin may cause households to replace the stolen goods or enhance the security of their homes – both of which is likely to cause expenditure to rise. Finally, a job loss in the household may lead to increase in consumption expenditure due to at least two reasons. First, a job loss might be followed by the person losing job to enroll in a job-training program which might increase total household expenditure. Second, there is ample evidence that suggests that job loss is associated with health compromising behaviors like alcohol abuse, which might also drive up total household expenditure.

0.128). Both these coefficients are also statistically significant at 1% level of significance. These results indicate that under the null hypothesis of no correlated unobservables, there is clear evidence of strong endogenous peer effects in consumption. However, if there are village/neighborhood level unobservables correlated with household characteristics and/or household specific idiosyncratic shocks, then these estimates are not consistent.

To account for the potential problem of correlated unobservables, I report the results of my preferred specification reported in column (3) which regresses own consumption on average peer consumption controlling for own income, own characteristics, own idiosyncratic shocks, average peer income, average peer characteristics and district fixed effects. I find the estimate of endogenous peer effect to be 0.69 (s.e. = 0.223) which is statistically significant at 1% level of significance. This means that an increase in (mean) peer consumption expenditure by 1 Indian Rupee causes households to increase their own consumption expenditure by roughly 0.7 Indian Rupee, on average. In terms of standard deviations, this translates into a 0.34 standard deviations (roughly 37,637 Indian Rupees/year) increase in own consumption in response to a one standard deviation (roughly 54,546 Indian Rupees/year) increase in peer consumption. Note that this effect is not only large in absolute terms but is in fact exactly equal to the impact of an equal increase (in terms of standard deviation) in own income. Overall, my results indicate the presence of strong peer effects in consumption.

To my knowledge, the only other study that seeks to estimate peer effects in aggregate consumption is that by De Giorgi et al. (2016). Using network data from Denmark they find that the elasticity of own consumption with respect to peers consumption is 0.3. How do my results compare to the findings of De Giorgi et al. (2016)? To answer this question, I estimate a double log version of my baseline specification (i.e., I use both households own consumption and average peer consumption in logs). My IV strategy delivers an estimate of the elasticity of roughly 0.5 (s.e. = 0.124).[32] This indicates that not only there are substantial peer effects in consumption in a low income country, but these effects are larger than those obtained for high income countries.

Social Multiplier. Peer effects potentially imply that small exogenous shock at the individual level is magnified through the social interactions process to deliver larger aggregate level social effects. Glaeser and Scheinkman (2001) and Glaeser et al. (2003) define the *social multiplier* as the ratio of the individual effect from an *exogenous* shock to the aggregate effect from the same shock. As argued by Sacerdote (2011), social multiplier is useful because it delivers the parameter of direct interest to policy makers, namely if a policy can exogenously induce one additional person to take action A, how many total people will take

[32] Results of this specification are not reported, but are available from the author on request.

action A in *equilibrium*? In the Supplementary Appendix, I show that the magnitude of the social multiplier based on my baseline regression Equation (5) is $1/(1-\beta)$. Thus, based on my preferred specification (i.e., $\beta \approx 0.7$) the value of the social multiplier in the present work is roughly equal to 3. This means that the equilibrium response to a shock that induces an exogenous variation in mean household consumption spending is about 3 times the initial average response.

5.2 Robustness Checks

In this section, I conduct various robustness checks to assess the robustness of my estimates to different sorts of bias.

Placebo Groups. As my first robustness check, I provide a falsification test. All my results indicate the presence of a positive and significant endogenous peer effects in consumption of a household where peer groups are assumed to comprise of all other people living in the same village or neighborhood. I now show that such a result is not obtained from considering *just any random* peer group. In essence, I validate the strength and significance of the actual observed peer group by ruling out the presence of peer effects within randomly generated peer groups. Put differently, this is a test for my identifying assumption that geographical proximity mediates peer effects.

To test this, I *randomly* assign peers to each household from across the entire sample. The total number of peers assigned to each household is equal the average size of peer groups, which is equal to 16. Subsequently, I estimate Equation (5) controlling for district fixed effect. I repeat this exercise 100 times, each time randomly assigning peers to every household.

The histogram in Figure 1 shows the empirical distribution of t-statistics corresponding to the endogenous peer effects obtained from 100 replications. The mean t-statistic is 1.06. This means that I cannot reject the null hypothesis that endogenous peer effect is equal to zero. Moreover, the majority of point estimates of endogenous peer effects are statistically insignificant as indicated by the distribution of the t-statistics. Specifically, for peer effects associated with each of the 100 iterations, I find that I am unable to reject the null that the coefficient (point estimate) is equal to zero for 89 coefficients out of 100. This shows that repeated experiments with different randomized social groups produce statistically insignificant endogenous peer effects on average. Hence, this exercise lends support to my approach of constructing peer groups based on geographical proximity.

(Group-level) Omitted Variable Bias. Despite having a comprehensive set of group level controls, as well as district fixed effects in my preferred specification of the baseline regression

model given by Equation (5), there might still be some omitted factors that vary between groups within districts. If this is the case, the estimates of the endogenous peer effects may not be consistent. To examine this, I augment my set of regressors by including three important covariate (community) shocks, namely, crop failure, droughts/floods and conflict. By definition, covariate shocks affect almost all households within a neighborhood/village. As such, it would be interesting to check the sensitivity of my baseline results to inclusion of these group level shocks.

Table 5 reports the results. I find that inclusion of the covariate shocks makes almost no difference to the estimates of the endogenous peer effects. Moreover, the coefficient of the covariate shocks themselves turn out to be statistically insignificant. This means that the covariate shocks, conditional on the other individual and group level regressors and district fixed effects, have no impact on households' own consumption expenditure. This suggests that group level shocks are unlikely to be responsible for the co-movement in household consumption within peer groups reported in my baseline estimation.

Exclusion of IVs. One potential critique of my IV strategy is that peer idiosyncratic shocks may have direct impact on own consumption. If this is true, the exclusion restriction for the instruments to work well is violated. To examine this, I report results of additional specifications in Table 6.

In each of the specification, I use two out of four instruments to instrument the endogenous peer effect and include the remaining two instruments as additional regressors in the second stage. I do this for all possible combinations of my instruments (hence, I report results from six specifications). I then examine whether the coefficients of instrument included as regressors in the second stage are statistically significant or not. I find that, across all the specifications, the coefficients of instruments in question are statistically indistinguishable from zero. This implies that, conditional on *any* two instruments being valid, the other instruments are excluded from the second stage. In other words, if one is willing to believe that only two instruments are valid (and is not sure about the validity of the remaining ones), this exercise shows that conditional on that belief, the remaining instruments are likely to be valid as well. This increases my confidence in the overall set of instruments used to identify my baseline model.

Using Consumption and Income in Logs. As noted by Suri (2004), a question that arises in the empirical work relating to consumption is whether to look at logs or levels of consumption, since in theory, consumption is assumed to follow a log normal distribution (Battistin et al., 2009). Similar question also arises in context of income. Although, given that I have a large enough sample, using consumption and income in logs is not necessary, it

is worthwhile to check whether there is any qualitative change in my baseline results when I take a log-transformation of the variables in question.

I present the results in column (1) of Table 7. Reassuringly, I find that the results remain qualitatively unaltered.

Measurement Error in Income. Measurement error in household income is always a concern in survey data. To show that this issue does not influence my main results, I re-estimate my baseline econometric model now instrumenting household's own income by literacy status of household head's father (or household head's husband's father, in case the household head is a woman). I also instrument average peer income by average literacy status of peer household head's father (or peer household head's husband's father). The assumption here is that whether a household head's father was literate or not affects earnings of the household, but whether a peer household head's father was literate or not has no direct influence on the household's own earnings. Column (2) of Table 7 reports the results from this exercise. As it turns out, the point estimate of endogenous peer effect remains almost unchanged compared to the baseline results.

5.3 Extension: Disaggregate Analysis

My data distinguishes between several different categories of consumption expenditure. This allows me to investigate more specifically the composition of the increase in own consumption expenditure in response to an increase in average peer consumption. I report the estimation results in Table 8.

I find that an increase in average peer consumption has a huge impact on own expenditure on temptation goods,[33] and a moderate impact on own food expenditure. More specifically, a 1 Indian Rupee increase in average peer consumption leads to a 0.48 Indian Rupee increase in own expenditure on temptation goods and 0.13 Indian Rupee increase in own expenditure on food items. These effects are statistically as well as economically significant. The impacts of an increase in average peer consumption on own health expenditure and own education expenditure, although positive, are much smaller compared to what I have found for the cases of temptation goods and food. Moreover, these effects are also imprecisely estimated. Overall, the results of the disaggregate analysis suggests that the increase in own consumption due to endogenous peer effect is driven by increase in own expenditure on temptation goods and food.[34]

[33] Expenditure on temptation goods include spending on entertainment, vacation, jewelry, clothing/bedding, footwear, etc. See Table A1 in the Supplementary Appendix for the full list of temptation goods.

[34] The disaggregate analysis is subject to a caveat. Note that my baseline model given by Equation (5)

6 Alternative Mechanism: Can the Results be Explained by Risk Sharing?

A plausible cause of co-movement in consumption of individuals' is risk-sharing or mutual insurance against idiosyncratic income shocks (Cochrane, 1991; Townsend, 1994). As such, could my results be explained by this mechanism? To examine this issue, I construct a stylized model of optimal risk allocation following Bardhan and Udry (1999) and Helmers and Patnam (2015) in the Supplementary Appendix and discuss whether the empirical results that I have obtained corroborate the predictions of this model.

My model predicts that under full risk sharing, household consumption perfectly co-moves with peer consumption. Moreover, household level income plays no role in determining household consumption, once average peer consumption has been controlled for. The intuition is as follows. In a full risk sharing environment where people have no access to formal saving mechanism, their consumption is determined as though all income (output) over all agents in the relevant risk sharing group (peer group) were pooled together and then redistributed such that everyone gets an equal share of the resources pooled. This implies that every agent's consumption is equal to the average of its peers' consumption. Thus, if resources pooled increases, the consumption of a particular agent as well as that of its peers' increases, implying a one-to-one relation between agents' and their peers' consumption. On the other hand, if total amount of resources pooled remains unchanged (which implies that average peer consumption remains unchanged), individual income should have no impact on individual consumption whatsoever.

Thus as noted by Suri (2004) and Chiappori et al. (2014), a test for full risk sharing would be to estimate Equation (5) and testing whether the coefficient of endogenous peer effect is equal to one and also whether the coefficient of own household income is equal to zero. If indeed the coefficient of endogenous peer effect turns out to be equal to one and that of own household income is zero, then full risk sharing may be causing the co-movement in consumption. This, however, is not what I have found in my baseline analysis. In all my specifications, I have found the magnitude of the coefficient of endogenous peer effect to be less than one and the coefficient of household income to be significantly different from zero at conventional levels of significance (see Table A3). Thus, I can rule out full risk sharing as

implies that if I use a expenditure on a certain type of goods as my dependent variable, expenditure on all other consumption goods ends up in the error term on the right hand side (unless they are explicitly controlled for). This means that the error term becomes correlated with all the regressors in the model. Now, if my instruments are correlated with some of these regressors in addition to being correlated with average peer consumption, then the estimate of the coefficient of average peer consumption will no longer be consistent.

a plausible explanation of my results.

Even though my empirical findings are not consistent with the full-risk sharing model, they could in principle be explained by a model of *partial* risk-sharing. According to the partial risk sharing hypothesis, if individuals are 'imperfectly' coinsuring income fluctuations (i.e., they do not entirely pool their income but save a part of it), then it is neither necessary for own consumption to perfectly to co-move with average peer consumption, nor is it necessary for own income to have zero impact on own consumption. This means that the endogenous peer effects obtained in this paper may be due to partial risk sharing rather than due to social conformity. So, in order to establish the validity of social conformity as the plausible mechanism driving my findings, it becomes critical to rule out partial risk sharing (although it is apparent that distinguishing social conformity from partial risk sharing is extremely difficult in practice since the predictions of the two models are virtually identical).

To assess whether the model of partial risk sharing can explain my results, I carry out a heterogeneity/subsample analysis. The central idea of the analysis is as follows. Note that, according to Townsend (1994), mutual coinsurance is likely to be more relevant for people living in a high risk environment with (almost) no formal risk reduction mechanism. As such, it is likely to be the case that, under partial risk sharing, the co-movement in consumption is likely to be greater for the people who are economically more vulnerable and who are exposed to high degrees of risk compared to the people who are economically less vulnerable and who are exposed to low degrees of risk (in fact, the risk sharing model may actually be irrelevant for the latter group of people). I assume people who are economically more vulnerable and live in a high risk environment are more likely to be living in rural areas than in urban areas, are members of low caste than of high caste, are not likely to be literate and/or live below the poverty line. So in order to rule out the partial risk sharing mechanism, I estimate my baseline regression model for each of these subsamples and then test whether the estimated endogenous peer effects differ significantly from that estimated using the respective sample-counterparts (e.g., I estimate the endogenous peer effect for the rural sample and test whether that differs significantly from the endogenous peer effects estimated using the urban sample). As hypothesized, if partial risk sharing indeed plays a role, the estimated endogenous peer effects must be statistically distinguishable from one-another with that being stronger for the regressions based on rural sample, low-caste sample, sample consisting of people who are not literate and sample consisting of households whose consumption/income are below the poverty line. If however, the coefficients of endogenous peer effects are not statistically distinguishable or if the endogenous peer effects are weaker for the economically less vulnerable subsamples, then it is unlikely that the partial risk sharing is the mechanism driving my results.

Table 9 reports the results from the subsample analysis. To test whether the difference between the estimated endogenous peer effects across two subsamples is significant, I report the results of the z-test (Paternoster et al., 1998). Strikingly, for all the cases I either find that the value of z-statistic is much lower than 2 (implying that the difference between the estimated endogenous peer effects is not statistically significant) or that the estimate of the endogenous effect is significantly lower for the subsample which is economically more vulnerable. This suggests that my baseline results are unlikely to be explained by the mechanism of partial risk sharing.[35]

7 Conclusion

A small but growing body of literature documents that peers play a significant role in influencing consumption of individuals in the developed world. However whether this is a universal social phenomenon or an artifact of affluent market oriented lifestyle is not entirely clear. In this paper, I explore this issue using household level data from India. Specifically, I attempt to understand whether consumption of individuals' peers affect their own consumption in a low income country as in the more prosperous economies. I define a household's peer group as other households living in its village/neighborhood. In assessing the influences of peers in this context, there are two key empirical challenges including shared group-level unobservables, and simultaneity of peer influences. I address these issues by using an instrumental variables/fixed effects approach that compares households in the same district but different villages/neighborhoods who are thus exposed to different sets of peers. In particular, I use plausibly exogenous variation in idiosyncratic expenditure shocks faced by peers as instruments for peers' consumption expenditure. Preferred specification suggests that a 1 Indian Rupee increase in consumption expenditure of a household's peers causes the household's own consumption expenditure to increase by 0.7 Indian Rupee which translates into a social multiplier of about 3. This means that the equilibrium response to a shock that induces an exogenous variation in mean household consumption spending is about 3 times the initial average response. Falsification tests and robustness checks support the validity of the results. Overall, thus my results indicate that Indian households do not differ from their counterparts in more affluent economies: peers play a crucial role in determining their

[35] In the Supplementary Appendix (Tables A5 and A6), I provide additional evidence that partial risk sharing is unlikely to be driving my baseline results. Specifically, I show that the estimated endogenous peer effect is not significantly higher for those households who belong to relatively more vulnerable social groups compared to those who belong to relatively less vulnerable social groups. I also show that the estimated endogenous peer effect is higher for those households belonging to more heterogenous social groups compared to those who belong to less heterogenous social groups. These findings are exactly opposite to what one would expect to find if the mechanism driving my baseline results is partial risk sharing.

consumption behavior.

While this paper is able to address many of the relevant econometric issues in estimating the importance of peer influences on consumption, there may be a few limitations. For example, an important drawback is that the social interaction structure considered in this paper may not be very realistic. That is, instead of being equally affected by all the others in the same group, a household may be influenced more significantly by some people in the peer group, especially by its 'friends'.

Even with such limitations of the study, I find robust evidence that peer consumption expenditure provide a strong influence on households' own consumption expenditure. These findings suggest that policies that influence a household's consumption expenditure will also affect the consumption decisions of the household's peers through social interactions. This implies that traditional analyses of consumption intervention programs that do not take into account such spillover effects likely understate the total social impact of the programs. Moreover, the findings also suggest that it might be worthwhile to use innovative social policies that represses social pressure together with traditional anti-poverty policies for raising social welfare in a low/middle income environment.

References

Akerlof, G. A. (1997). 'Social Distance and Social Decisions', *Econometrica*, 65, 1005–1027.

Altonji, J. G., Elder, T. E. and Tabor, C. R. (2005). 'Selection on observed and unobserved variables: assessing the effectiveness of catholic schools', *Journal of Political Economy*, 113, 151–184.

Ammermueller, A. and Pischke, J-S. (2009). 'Peer effects in European primary schools: evidence from the Progress in International Reading Literacy Study', *Journal of Labor Economics*, 27(3), 315–348.

Angelucci, M. and De Giorgi, G. (2009). 'Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?', *American Economic Review*, 99, 486–508.

Bardhan, P. and Udry, C. (1999). *Development Microeconomics*, New York: Oxford University Press.

Battistin, E., Blundell, R. and Lewbel, A. (2009). 'Why Is Consumption More Log Normal than Income? Gibrat's Law Revisited', *Journal of Political Economy*, 117, 1140–1154.

Baum, C. F., Schaffer, M. E. and Stillman, S. (2007). 'Enhanced routines for instrumental variables/generalized method of moments estimation and testing', *Stata Journal*, 7, 465–506.

Becker, G. S. (1974). 'A Theory of Social Interactions', *Journal of Political Economy*, 82, 1063–1093.

Bernheim, B. D. (1994). 'A Theory of Conformity', *Journal of Political Economy*, 102, 841–877.

Bifulco, R., Fletcher, J. and Ross, S. (2011). 'The effect of classmate characteristics on individual outcomes: evidence from the add health', *American Economic Journal: Economic Policy*, 3, 25–53.

Bound, J., Jaeger, D. and Baker, R. (1995). 'Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak', *Journal of American Statistical Association*, 90, 443–450.

Bramoullé, Y., Djebbari, H. and Fortin, B. (2009). 'Identification of peer effects through social networks', *Journal of Econometrics*, 150, 41–55.

Brueckner, J. K. (2003). 'Strategic interaction among governments: an overview of empirical studies', *International Regional Science Review,* 26, 175–188.

Calvo-Armengol, A., Patacchini, E. and Zenou, Y. (2009). 'Peer effects and social networks in education', *Review of Economic Studies*, 76, 1239–1267.

Cameron, A. C. and Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*, New York: Cambridge University Press.

Chiappori, P., Samphantharak, K., Schulhofer-Wohl, S. and Townsend, R. (2014). 'Heterogeneity and risk sharing in village economies', *Quantitative Economics*, 5, 1–27.

Clark, A. E. and A. J. Oswald (1998). 'Comparison-concave utility and following behaviour in social and economic settings', *Journal of Public Economics*, 70, 133–155.

Cliff, A. D. and Ord, J. K. (1981). *Spatial Processes: Models and Applications*, London: Pion.

Cochrane, J. (1991). 'A Simple Test of Consumption Insurance', *Journal of Political Economy*, 99, 957–976.

Cojocaru, A. (2014). 'Fairness and inequality tolerance: Evidence from the Life in Transition Survey', *Journal of Comparative Economics,* 42, 590–608.

De Giorgi, G., Frederiksen, A. and Pistaferri, L. (2016). 'Consumption Network Effects', NBER Working Paper 22357.

Deaton, A. (2001). 'Relative deprivation, inequality and mortality', NBER Working Paper 8099.

Desai, S., Dubey, A. and Vanneman, R. (2015). India Human Development Survey-II (IHDS-II) (Computer file), University of Maryland and National Council of Applied Economic Research, New Delhi (producers), 2015. Ann Arbor, MI: Inter-university Consortium for Political and Social Research (distributor).

Duesenberry, J. S. (1949). *Income, Saving and the Theory of Consumer Behavior*, Cambridge: Harvard University Press.

Fafchamps, M. and Shilpi, F. (2008). 'Subjective welfare, isolation, and relative consumption', *Journal of Development Economics*, 86, 43–60.

Falk, A. and Knell, M. (2004). 'Choosing the joneses: endogenous goals and reference standards', *Scandinavian Journal of Economics*, 106, 417–435.

Fletcher, J. (2010). 'Social Interactions and Smoking: Evidence using Multiple Student Cohorts, Instrumental Variables, and School Fixed Effects', *Health Economics*, 19, 466–484.

Fletcher, J. (2012). 'Peer influences on adolescent alcohol consumption: evidence using an instrumental variables/fixed effect approach', *Journal of Population Economics*, 25,1265–1286.

Fletcher, J. (2015). 'Social interactions and college enrollment: A combined school fixed effects/instrumental variables approach', *Social Science Research*, 52, 494–507.

Gaviria, A. and Raphael, S. (2001). 'School-based Peer Effects and Juvenile Behavior', *Review of Economics and Statistics*, 83, 257–268.

Goux, D. and Maurin, E. (2007). 'Close Neighbours Matter: Neighbourhood Effects on Early Performance at School', *Economic Journal*, 117, 1193–1215.

Gibbons, S. and Overman, H. G. (2012). 'Mostly Pointless Spatial Econometrics', *Journal of Regional Science*, 52, 172–191.

Gibbons, S. Overman, H. G., and Patacchini, E. (2015). 'Spatial Methods', In: G. Duranton, J. V. Henderson and W. C. Strange (Eds.), *Handbook of Regional and Urban Economics*, 5A, 115–168, Amsterdam: North Holland.

Glaeser, E. L. and Scheinkman, J. (2001). 'Measuring social interactions', In: S. Durlauf, S., H.P. Young (Eds.), *Social Dynamics*, 83–131, Cambridge: MIT Press.

Glaeser, E. L., Scheinkman, J. A. and Sacerdote, B. I. (2003). 'The Social Multiplier', *Journal of the European Economic Association*, 1, 345–353.

Graham, B. S. (2008). 'Identifying Social Interactions through Conditional Variance Restrictions', Econometrica, 76, 643–660.

Grinblatt, M., Keloharju, M. and Ikäheimo, S. (2008). 'Social Influence and Consumption: Evidence from the Automobile Purchases of Neighbors', *Review of Economics and Statistics*, 90, 735–753.

Grodner, A. and Kniesner, T. J. (2006). 'Social interactions in labor supply', *Journal of the European Economic Association*, 4, 1226–1248.

Hansen, L. (1982). 'Large sample properties of generalized method of moments estimators', *Econometrica*, 50, 1029–1054.

Helmers, C. and M. Patnam (2015). 'Does the rotten child spoil his companion? Spatial peer effects among children in rural India', *Quantitative Economics*, 5, 67–121.

Indian National Census (2011a). 'Provisional Population Totals: Urban Agglomerations and Cities', Government of India, http://censusindia.gov.in/2011-prov-results/paper2/data_files/India2/1.%20Data%20Highlight.pdf (last accessed September 30, 2016).

Indian National Census (2011b). 'Top 10 Smallest Districts of India by Total Area', Government of India, http://www.census2011.co.in/facts/smalldistricts.html (last accessed September 30, 2016).

Kandel, E. and Lazear, E.P. (1992). 'Peer pressure and partnerships', *Journal of Political Economy*, 100, 801–817.

Kelejian, H. H. and Prucha, I. R. (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*, 17, 99–121.

Kleibergen, F. and Paap, R. (2006). 'Generalized Reduced Rank Tests Using the Singular Value Decomposition', *Journal of Econometrics*, 133, 97–126.

Kuhn, P., Kooreman, P., Soetevent, A. and Kapteyn, A. (2011). 'The Effects of Lottery Prizes on Winners and Their Neighbors: Evidence from the Dutch Postcode Lottery', *American Economic Review*, 101, 2226–2247.

Lavy, V. and Schlosser, A. (2011). 'Mechanisms and Impacts of Gender Peer Effects at School', *American Economic Journal: Applied Economics*, 3, 1–33.

Lee, L. F. (2007). 'Identification and estimation of econometric models with group interactions, contextual factors and fixed effects', *Journal of Econometrics*, 140, 333–374.

Lee, L. F., Xiaodong Liu, X. and Lin, X. (2010), 'Specification and Estimation of Social Interaction Models with Network Structures', *Econometrics Journal*, 13, 145–176.

Leibenstein, H. (1950). 'Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand', *Quarterly Journal of Economics*, 64, 183–207.

Liu, X., Patacchini, E. and Zenou, Y. (2014). 'Endogenous peer effects: local aggregate or local average?' *Journal of Economic Behavior and Organization*, 103, 39–59.

Manski, C. (1993). 'Identification of Endogenous Social Effects: The Reflection Problem', *Review of Economic Studies*, 60, 531–542.

McVicar. D. and Polanski, A. (2014). 'Peer Effects in UK Adolescent Substance Use: Never Mind the Classmates?', *Oxford Bulletin of Economics and Statistics*, 76, 589–604.

Moav, O. and Neeman, Z. (2012). 'Saving Rates and Poverty: The Role of Conspicuous Consumption and Human Capital', *Economic Journal*, 122, 933–956.

Moretti, E. (2011). 'Social Learning and Peer Effects in Consumption: Evidence from Movie Sales', *Review of Economic Studies*, 78, 356–393

Munshi, K. and Rosenzweig, M. (2009). 'Why is Mobility in India so Low? Social Insurance, Inequality, and Growth', NBER Working Paper No. 14850.

Nesse, R. M. (2004). 'Natural selection and elusiveness of happiness', *Philosophical transactions of the Royal Society B (Biological Sciences)*, 359, 1333–1347.

Paternoster, R., Brame, R., Mazerolle, P. and Piquero, A. (1998). 'Using the correct statistical test for the equality of regression coefficients', *Criminology*, 36, 859–866.

Pollak, R. A. (1976). 'Interdependent Preferences', *American Economic Review*, 66, 309–320.

Ravallion, M. and Lokshin, M. (2005). 'Who cares about relative deprivation?', Policy Research Working Paper No. 3782, Development Research Group, World Bank.

Roth, C. P. (2014). 'Conspicuous Consumption and Peer Effects among the Poor: Evidence From a Field Experiment', Working Paper, Centre for the Study of African Economies, Department of Economics, University of Oxford.

Roychowdhury, P. (2016). 'Visible inequality, status competition and conspicuous consumption: evidence from rural India', *Oxford Economic Papers* (forthcoming), doi: 10.1093/oep/gpw056.

Sacerdote, B. (2001). 'Peer effects with random assignment: results for Dartmouth roommates', *Quarterly Journal of Economics*, 116, 681–704.

Sacerdote, B. (2011). 'Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?', In: E.A. Hanushek, S. Machin and L. Woessmann (Eds.), *Handbook of Economics of Education*, 3, 249–277, Amsterdam: North Holland.

Scott, J. C. (1976). *The Moral Economy of Peasant: Rebellion and Subsistence in South-East Asia*, New Haven: Yale University Press.

Staiger, D. and Stock, J. H. (1997). 'Instrumental variables regression with weak instruments', *Econometrica*, 65, 557–586.

Stark, O. and Taylor, J. E. (1991). 'Migration incentives, migration types: the role of relative deprivation', *Economic Journal*, 101, 1163–1178.

Stock, J., Wright, J. and Yogo, M. (2002). 'A survey of weak instruments and weak identification in generalized method of moments', *Journal of Business and Economic Statistics*, 20, 518–529.

Stock, J. H. and Yogo, M. (2005). 'Testing for weak instruments in linear IV regression', In: D.W. Andrews and J.H. Stock (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, 80–108, New York: Cambridge University Press.

Suri, T. (2004). 'Spillovers in Village Consumption: Testing the Extent of Partial Insurance', Working Paper, Sloan School of Management, Massachusetts Institute of Technology. Available at http://economie.esg.uqam.ca/upload/files/seminaires/suri_spillovers.pdf.

Townsend, R. M. (1994). 'Risk and insurance in village India', *Econometrica*, 62, 539–591.

Table 1. Variable Definitions and Summary Statistics

			Own Characteristics		Peers' Mean Characteristics	
Variable	Definition	N	Mean	SD	Mean	SD
Consumption	Annual total household consumption expenditure (Indian Rupees)	40980	115239	111528	115239	54545
Income	Annual total household income (Indian Rupees)	40980	118130	130239	118130	71653
Household Size	Total number of people in the household	40980	4.85	2.29	4.85	0.91
Age	Age of household head	40980	49.62	13.55	49.62	4.7
Male	= 1 if household head is male; = 0 otherwise	40980	0.86	0.35	0.86	0.11
Literate	= 1 if household head is literate; = 0 otherwise	40980	0.68	0.47	0.68	0.22
Married	= 1 if household head is married; = 0 otherwise	40980	0.81	0.39	0.81	0.12
(Children Proportion)	Number of children/Total number of people in the household	40980	0.24	0.22	0.24	0.09
Teenage Proportion	Number of teens/Total number of people in the household	40980	0.11	0.16	0.11	0.05
Adult Proportion	Number of adults/Total number of people in the household	40980	0.65	0.23	0.65	0.1
(No Married)	= 1 if there are no married people in the household; = 0 otherwise	40980	0.09	0.28	0.09	0.08
One to Five Married	= 1 if there are more than zero and less than five married people in the household; = 0 otherwise	40980	0.87	0.33	0.87	0.09
More than Five Married	= 1 if there are more than five married people in the household; 0 otherwise	40980	0.04	0.19	0.04	0.05
Years in Place	= 1 if the household has been living in the same place for more than 10 years; 0 otherwise	40980	0.97	0.18	0.97	0.08
Urban	= 1 if the household's place of residence is categorized as Urban as per the 2011 Census; 0 otherwise	40980	0.35	0.48	0.35	0.48
Brahmin	= 1 if household head's caste is Brahmin; 0 otherwise	40980	0.05	0.22	0.05	0.11
Forward Caste	= 1 if household head's caste is Non-Brahmin Forward Caste; = 0 otherwise	40980	0.23	0.42	0.23	0.26
OBC	= 1 if household head's caste is Other backward Classes (OBC); = 0 otherwise	40980	0.41	0.49	0.41	0.3
SC	= 1 if household head's caste is Scheduled Caste (SC); = 0 otherwise	40980	0.22	0.41	0.22	0.23
ST	= 1 if household head's caste is Scheduled Tribe (ST); = 0 otherwise	40980	0.09	0.28	0.09	0.21
(Other Caste)	= 1 if household head is a member of some other caste; = 0 otherwise	40980	0.01	0.11	0.01	0.06
Job Loss	= 1 if a household member has lost job in the recent past; = 0 otherwise	40980	0.03	0.16	0.03	0.08
Death	= 1 if there has been a death in the recent past; 0 = otherwise	40980	0.19	0.39	0.19	0.13
Theft	= 1 if there has been an incident of theft in the household in the recent past; = 0 otherwise	40980	0.04	0.19	0.04	0.07
Break-in	= 1 if the household faced an incident of break-in in the recent past; = 0 otherwise	40980	0.01	0.1	0.01	0.03

Notes: The variables in the parentheses are the omitted categories in the following estimations.

Table 2. Peer Effects in Consumption: OLS Estimation

Variables	[1]	[2]	[3]
Endogenous Peer Effect	**0.461*** **	**0.568*** **	**0.207*** **
	(0.013)	**(0.017)**	**(0.026)**
Individual Effects	YES	YES	YES
Idiosyncratic Shocks	YES	YES	YES
Contextual Effects	NO	YES	YES
District Fixed Effects	NO	NO	YES
Observations	40980	40980	40980
Adjusted R-squared	0.324	0.329	0.338

Notes: Dependent variable is Household Annual Total Consumption Expenditure. Endogenous Peer Effect represents the estimated coefficient of Average Peer Household Annual Total Consumption Expenditure. Individual effects are controlled by own demographic characteristics of households that include Income, Household Size, Age, Male, Literate, Married, Teenage Proportion, Adult Proportion, One to Five Married, More than Five Married, Years in Place, Urban, Brahmin, Forward Caste, OBC, SC and ST. Idiosyncratic shocks include Job Loss, Death, Theft and Break-in. Contextual effects are controlled by average peer demographic characteristics. For definition of variables see Table 1. Heteroscedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3. Validity of Instruments: Balancing Test

Dependent Variable	Peer Job Loss	Peer Death	Peer Theft	Peer Break-in	F-statistic	Observations
Income	11716	8246	6630	-15261	0.88	40980
	(13460)	(5877)	(12505)	(23283)	[p=0.474]	
Household Size	0.077	0.156	0.082	-0.589	1.21	40980
	(0.215)	(0.103)	(0.217)	(0.401)	[p=0.305]	
Age	-0.478	1.734***	-1.087	-1.417	2.31	40980
	(1.415)	(0.626)	(1.301)	(2.458)	[p=0.055]	
Male	0.0273	0.023	0.080**	-0.156**	2.99	40980
	(0.035)	(0.017)	(0.033)	(0.064)	[p=0.018]	
Literate	0.011	0.023	0.064	0.010	1.13	40980
	(0.045)	(0.020)	(0.043)	(0.077)	[p=0.340]	
Married	0.047	0.035*	0.093	-0.095	3.12	40980
	(0.039)	(0.018)	(0.037)	(0.071)	[p=0.014]	
Children Proportion	0.013	0.006	-0.002	0.049	0.60	40980
	(0.022)	(0.010)	(0.021)	(0.040)	[p=0.663]	
Teenage Proportion	-0.012	-0.012	0.012	-0.014	0.85	40980
	(0.017)	(0.008)	(0.016)	(0.031)	[p=0.492]	
Adult Proportion	-0.000	0.006	-0.011	-0.035	0.45	40980
	(0.024)	(0.010)	(0.021)	(0.041)	[p=0.772]	
No Married	-0.077**	-0.021	-0.036	0.081	3.39	40980
	(0.030)	(0.013)	(0.026)	(0.054)	[p=0.009]	
One to Five Married	0.077**	-0.002	0.051	-0.120*	2.48	40980
	(0.035)	(0.016)	(0.031)	(0.063)	[p=0.042]	
More than Five Married	0.000	0.023**	-0.015	0.0388	1.92	40980
	(0.019)	(0.009)	(0.018)	(0.034)	[p=0.105]	
Years in Place	-0.008	0.006	-0.016	-0.000	0.43	40980
	(0.019)	(0.007)	(0.018)	(0.037)	[p=0.786]	
Brahmin	-0.014	0.002	-0.010	0.027	0.21	40980
	(0.021)	(0.009)	(0.023)	(0.045)	[p=0.934]	
Forward Caste	0.016	-0.007	-0.027	0.046	0.26	40980
	(0.037)	(0.017)	(0.035)	(0.076)	[p=0.903]	
OBC	-0.023	-0.016	0.033	-0.071	0.49	40980
	(0.042)	(0.019)	(0.041)	(0.080)	[p=0.744]	
SC	0.026	0.022	0.005	-0.023	0.60	40980
	(0.039)	(0.017)	(0.036)	(0.071)	[p=0.663]	
ST	-0.015	-0.000	-0.002	0.025	0.32	40980
	(0.018)	(0.008)	(0.019)	(0.034)	[p=0.868]	
Other Caste	0.009	-0.000	0.000	-0.004	0.13	40980
	(0.015)	(0.0051)	(0.008)	(0.016)	[p=0.973]	

Notes: Each row shows results of regression of a specific household demographic characteristic on different types of average peer shocks. Additional controls in each regression include own idiosyncratic shocks, average peer characteristics and district fixed effects. For definition of variables see Table 1. All regressions are estimated by OLS and include a constant. Heteroscedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4. Peer Effects in Consumption: IV Estimation

Variables	[1]	[2]	[3]
Endogenous Peer Effect	**0.782*** **	**0.848*** **	**0.697*** **
	(0.147)	**(0.128)**	**(0.223)**
Individual Effects	YES	YES	YES
Idiosyncratic Shocks	YES	YES	YES
Contextual Effects	NO	YES	YES
District Fixed Effects	NO	NO	YES
Observations	40980	40980	40980
Adjusted R-squared	0.305	0.321	0.325
Hansen J statistic	6.096	4.355	1.829
	[p=0.107]	[p=0.226]	[p=0.609]
Kleibergen-Paap rk LM statistic	147.5	339.9	154.8
	[p=0.000]	[p=0.000]	[p=0.000]
First stage F-statistic	38.03	89.97	40.01

Notes: Estimation via two-step GMM. Dependent variable is Household Annual Total Consumption Expenditure. Endogenous Peer Effect represents the estimated coefficient of Average Peer Household Annual Total Consumption Expenditure. Individual effects are controlled by own demographic characteristics of households. Contextual effects are controlled by average peer demographic characteristics. For full list of demographic characteristic, idiosyncratic shocks and definition of variables see Table 1 and note below Table 2. Heteroscedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5. Peer Effects in Consumption: Investigating (Group Level) Omitted Variable Bias

Variables	[1]	[2]	[3]
Endogenous Peer Effect	**0.684*** **	**0.681*** **	**0.681*** **
	(0.238)	**(0.236)**	**(0.289)**
Covariate Shocks			
Crop Failure	**3322**		
	(4220)		
Droughts/Floods		**2639**	
		(4449)	
Conflict			**1661**
			(2878)
Observations	40980	40980	40980
Adjusted R-squared	0.326	0.326	0.332
Hansen J statistic	1.926	1.834	1.997
	[p=0.588]	[p=0.608]	[p=0.573]
Kleibergen-Paap rk LM statistic	135.6	137.5	129.5
	[p=0.000]	[p=0.000]	[p=0.000]
First stage F-statistic	34.90	35.45	33.28

Notes: Estimation via two-step GMM. Dependent variable is Household Annual Total Consumption Expenditure. Endogenous Peer Effect represents the estimated coefficient of Average Peer Household Annual Total Consumption Expenditure. All specifications include controls for individual effects, own idiosyncratic shocks, contextual effects and district fixed effects. Individual effects are controlled by own demographic characteristics of households. Contextual effects are controlled by average peer demographic characteristics. For full list of demographic characteristics, idiosyncratic shocks and definition of variables see Table 1 and note below Table 2. Heteroscedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6. Peer Effects in Consumption: Investigating IV Exclusion

Variables	[1]	[2]	[3]	[4]	[5]	[6]
Endogenous Peer Effect	**0.667****	**0.921*****	**0.445**	**0.879*****	**0.551**	**0.769****
	(0.329)	**(0.348)**	**(0.493)**	**(0.287)**	**(0.356)**	**(0.317)**
Peer Idiosyncratic shocks (Instruments)						
Job Loss	**1257**	**-3965**	**7170**			
	(13307)	**(13541)**	**(16503)**			
Death	**643.2**			**-1139**	**1973**	
	(5566)			**(5579)**	**(6033)**	
Theft		**-12946**		**-12609**		**-12230**
		(10010)		**(9876)**		**(9546)**
Break-in			**16226**		**12361**	**9378**
			(26993)		**(23571)**	**(21086)**
Observations	40980	40980	40980	40980	40980	40980
Adjusted R-squared	0.327	0.310	0.335	0.313	0.332	0.321
Hansen J statistic	1.821	0.109	1.493	0.153	1.568	3.25e-05
	[p=0.177]	[p=0.742]	[p=0.222]	[p=0.695]	[p=0.211]	[p=0.995]
Kleibergen-Paap rk LM statistic	69.69	66.46	42.22	101.8	80.03	91.62
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]
First stage F-statistic	36.11	33.92	20.94	52.34	40.67	47.80

Notes: Estimation via two-step GMM. Dependent variable is Household Annual Total Consumption Expenditure. Endogenous Peer Effect represents the estimated coefficient of Average Peer Household Annual Total Consumption Expenditure. All specifications include controls for individual effects, own idiosyncratic shocks, contextual effects and district fixed effects. Individual effects are controlled by own demographic characteristics of households. Contextual effects are controlled by average peer demographic characteristics. For full list of demographic characteristic, idiosyncratic shocks and definition of variables see Table 1 and note below Table 2. Heteroscedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7. Peer Effects in Consumption: Consumption in Logs and Endogenous Income

Variables	[1]	[2]
Endogenous Peer Effect	**0.639*** **	**0.634*** **
	(0.157)	**(0.241)**
Observations	40980	40980
Adjusted R-squared	0.568	0.155
Hansen J statistic	1.454	1.761
	[p=0.693]	[p=0.623]
Kleibergen-Paap rk LM statistic	264.9	134.4
	[p=0.000]	[p=0.000]
First stage F-statistic		
Mean Peer Consumption	71.29	182.97
Household Income		58.18
Mean Peer Income		135.81

Notes: Estimation via two-step GMM. For the specification reported in column [1], the dependent variable is log of Household Annual Total Consumption Expenditure and the Endogenous Peer Effect represents the estimated coefficient of log of Average Peer Household Annual Total Consumption Expenditure. For the specification reported in column [2], the dependent variable is Household Annual Total Consumption Expenditure and the Endogenous Peer Effect represents the estimated coefficient of Average Peer Household Annual Total Consumption Expenditure. Both specifications include controls for individual effects, own idiosyncratic shocks, contextual effects and district fixed effects. Individual effects are controlled by own demographic characteristics of households. Contextual effects are controlled by average peer demographic characteristics. Own and Average Peer Income are instrumented in the specification reported in column (2). For full list of demographic characteristics, idiosyncratic shocks, instruments and definition of variables see Table 1, note below Table 2 and main text. Heteroscedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8. Peer Effects in Consumption: Disaggregate Analysis

	Expenditure Categories			
		Temptation		
Variables	Food	Good	Health	Education
Mean Peer Total Consumption Expenditure	**0.125*** **	**0.482*** **	**0.055**	**0.006**
	(0.0418)	**(0.173)**	**(0.071)**	**(0.046)**
Observations	40980	40980	40980	40980
Adjusted R-squared	0.562	0.186	0.0220	0.143
Hansen J statistic	2.472	1.557	3.677	4.417
	[p=0.480]	[p=0.669]	[p=0.299]	[p=0.220]
Kleibergen-Paap rk LM statistic	154.800	154.800	154.800	154.800
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]
First stage F-statistic	40.01	40.01	40.01	40.01

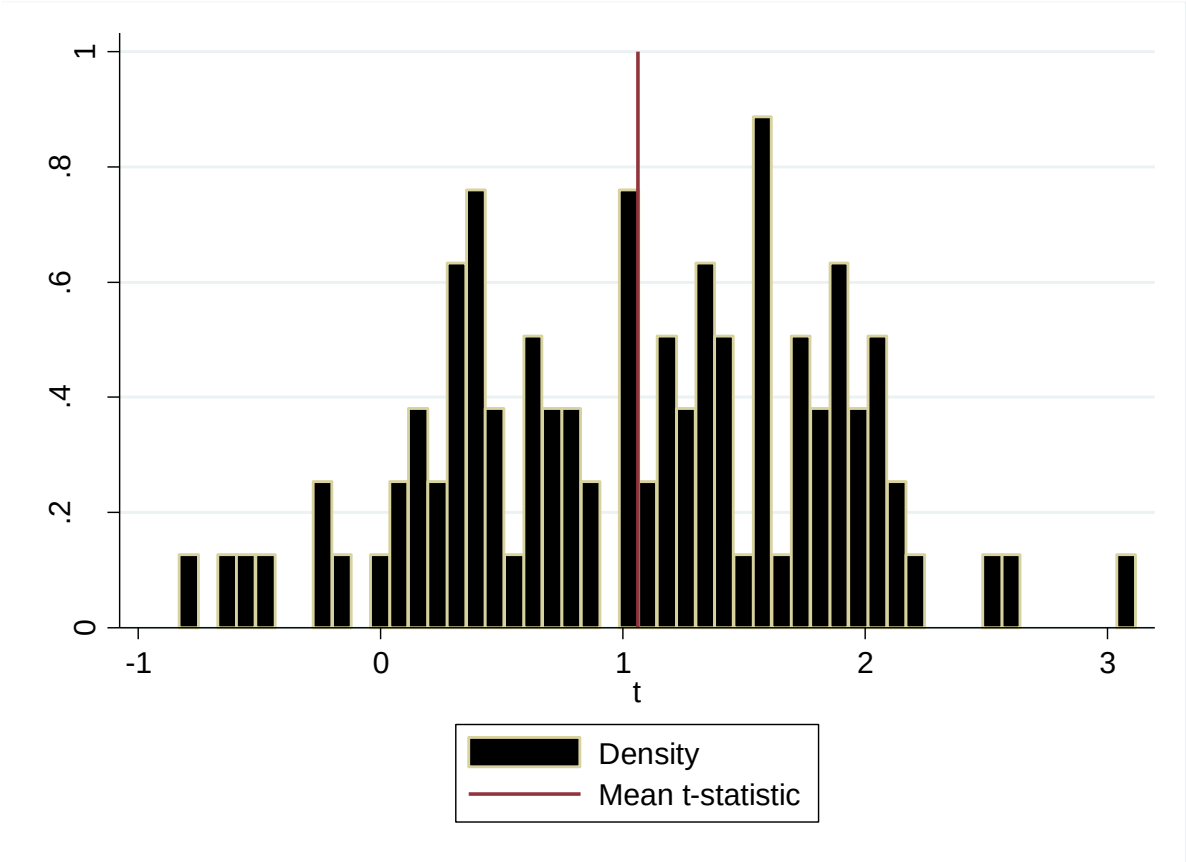
Notes: Estimation via two-step GMM. Dependent variable for regression reported in each column is the respective expenditure category. All specifications include controls for individual effects, own idiosyncratic shocks, contextual effects and district fixed effects. Individual effects are controlled by own demographic characteristics of households. Contextual effects are controlled by average peer demographic characteristics. For full list of demographic characteristic, idiosyncratic shocks and definition of variables see Table 1 and note below Table 2. Heteroscedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 9. Partial Risk Sharing: Subsample Analysis

	Subsamples							
Variables	Urban	Rural	Literate	Not Literate	Upper Caste	Lower Caste	Non-Poor	Poor
Endogenous Peer Effect	**0.873*** **	**0.754*** **	**0.964*** **	**0.195**	**0.887** **	**0.565** **	**0.728*** **	**0.024**
	(0.304)	**(0.227)**	**(0.320)**	**(0.190)**	**(0.360)**	**(0.240)**	**(0.240)**	**(0.079)**
Observations	14171	26809	27747	13233	11465	29515	34168	6812
Adjusted R-squared	0.185	0.195	0.179	0.225	0.190	0.199	0.218	0.735
Hansen J statistic	1.493	0.766	1.000	2.314	0.416	2.815	3.597	7.773
	[p=0.684]	[p=0.858]	[p=0.801]	[p=0.510]	[p=0.937]	[p=0.421]	[p=0.308]	[p=0.0509]
Kleibergen-Paap rk LM statistic	95.99	191.5	83.68	102.7	68.70	121.2	133.6	39.37
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]
First stage F-statistic	25.23	53.03	21.66	25.67	17.26	32.30	34.49	10.61
z-statistic	0.314		2.066		0.744		2.786	

Notes: Estimation via two-step GMM. Dependent variable is Household Annual Total Consumption Expenditure. Endogenous Peer Effect represents the estimated coefficient of Average Peer Household Annual Total Consumption Expenditure. All specifications include controls for individual effects, own idiosyncratic shocks, contextual effects and district fixed effects. Individual effects are controlled by own demographic characteristics of households. Contextual effects are controlled by average peer demographic characteristics. For full list of demographic characteristic, idiosyncratic shocks, definition of variables and definition of subsamples see Table 1, note below Table 2 and main text. Heteroscedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure 1. Falsification Test: Histogram of t-statistics corresponding to Point Estimates of Endogenous Peer Effects



Supplementary Appendix

Peer Effects in Consumption

Punarjit Roychowdhury
Southern Methodist University

FOR ONLINE PUBLICATION ONLY

1 An Alternative Model of Social Interactions in Consumption

In this section, I present a more general theoretical model of peer effects in consumption in which social utility is generated not only due to social conformity but also due to status seeking behavior. As in the model presented in the main text, this model also predicts that a household's consumption is positively related with that of its peers'.

I assume households have common preferences given by the utility function

$$U_{i,r} = u_1(c_{i,r}) + u_2(b_{i,r}) + u_3(\mathbf{\Psi}_{i,r}) + s(c_{i,}\overline{c}_{-i,r}) \tag{1}$$

where $u_1(.)$, $u_2(.)$, and $u_3(.)$ are 'conventional' utility functions which depend on $c_{i,r}$, $b_{i,r}$ and $\mathbf{\Psi}_{i,r}$ respectively (see Section 2 of the main text for definitions of the variables); $s(.)$ captures social utility or utility from social comparisons (i.e., it measures the utility attributable to the deviation of own consumption from average peer consumption). I assume

$$s(.) = \theta\left[-\frac{\phi}{2}(c_{i,r} - \overline{c}_{-i,r})^2\right] + (1-\theta)v(c_{i,r} - \overline{c}_{-i,r}) \tag{2}$$

which is a weighted average of two social subutility functions with $\theta \in (0,1)$. The first social subutility function measures utility due to conformity as in the main model. The social subutility function $v(.)$ on the other hand measures utility due to status seeking behavior or utility from getting ahead of others. I assume $v'(.) > 0$ and $v''(.) < 0$. These assumptions have been used by Corneo (2002) and implies that people enjoy surpassing others' consumption, and that comparison utility is concave (i.e., diminishing marginal comparison utility holds or higher is a household's consumption relative to the mean, lesser is its increase in utility due to increase in status).

The problem of household i is to choose $c_{i,r}$ and $b_{i,r}$ by maximizing (2) subject to a income budget constraint given by:

$$c_{i,r} + b_{i,r} \leq y_{i,r}; c_{i,r}, b_{i,r} \geq 0 \tag{3}$$

For an interior maximum

$$u_1'(c_{i,r}) - u_2'(y_{i,r} - c_{i,r}) - \theta\phi(c_{i,r} - \overline{c}_{-i,r}) + (1-\theta)v'(c_{i,r} - \overline{c}_{-i,r}) = 0 \tag{4}$$

To find how household i responds to a change in others consumption, I differentiate (4) implicitly and simplify get:

$$\frac{\partial c_{i,r}}{\partial \overline{c}_{-i,r}} = \frac{-\theta\phi + (1-\theta)v''(c_{i,r} - \overline{c}_{-i,r})}{u_1''(c_{i,r}) + u_2''(y_{i,r} - c_{i,r}) - \theta\phi + (1-\theta)s_2''(c_{i,r} - \overline{c}_{-i,r})} \tag{5}$$

The numerator on the right-hand side of equation (5) is negative and the denominator is negative by the requirement of the maximization problem to be concave. Thus equation (5) implies that an increase in average consumption causes household i to increase its own consumption. However, now a part of response of $c_{i,r}$ to an increase in $\overline{c}_{-i,r}$ comes from conformity/social norms motive and the other part comes from 'getting ahead of others' motive.

2 Social Multiplier

Endogenous peer effects imply that small exogenous shock at the individual level is magnified through the social interactions process to deliver larger aggregate level social effects. Glaeser and Scheinkman (2001) and Glaeser et al. (2003) define the social multiplier as the ratio of the individual effect from an *exogenous* shock to the aggregate effect from the same shock. In what follows, I derive the social multiplier for my baseline econometric model:

$$c_{i,r} = \alpha + \beta\overline{c}_{-i,r} + \gamma y_{i,r} + \mathbf{x}_{i,r}'\boldsymbol{\varpi} + \overline{\mathbf{x}}_{-i,r}'\boldsymbol{\delta} + \theta\overline{y}_{-i,r} + \mathbf{z}_{i,r}'\boldsymbol{\rho} + \xi_{i,r} \tag{6}$$

Let

$$\Theta_{i,r} = \alpha + \gamma y_{i,r} + \mathbf{x}_{i,r}'\boldsymbol{\varpi} + \overline{\mathbf{x}}_{-i,r}'\boldsymbol{\delta} + \theta\overline{y}_{-i,r} + \mathbf{z}_{i,r}'\boldsymbol{\rho} + \xi_{i,r} \tag{7}$$

Thus, Equation (6) can be rewritten as

$$\begin{aligned} c_{i,r} &= \Theta_{i,r} + \beta\overline{c}_{-i,r} \\ &= \Theta_{i,r} + \frac{\beta}{m_r - 1}\sum_{j\in M_{i,r}} c_{j,r} \end{aligned} \tag{8}$$

where $M_{i,r}$ denotes peer group of household i and $(m_r - 1)$ denotes size of the peer group.

Based on Equation (8), it can show that

$$\frac{1}{m_r}\sum_{j\in M_r} c_{j,r} = \frac{1}{1-\beta}\frac{1}{m_r}\sum_{j\in M_r}\Theta_{j,r} \tag{9}$$

$$\text{or, } \overline{c}_r = \frac{1}{1-\beta}\overline{\Theta}_r \tag{10}$$

where M_r denotes the social group to which household i belongs to. This is the aggregate level equation.

And also we can show that in equilibrium,

$$c_{i,r} = \Theta_{i,r}\left(1 + \frac{\beta^2}{(1-\beta)(m_r - 1 + \beta)}\right) + \frac{\beta}{m_r - 1 + \beta}\sum_{j \in M_{i,r}} \Theta_{j,r} \tag{11}$$

Equation (11) is the individual/household level equation. To simply the analysis, I consider a three-person social group (i.e., $m_r = 3$). Let the individuals be indexed by $v = 1$, 2, 3. Further, I assume $\gamma = \theta = \rho = 0$ and there is a unique exogenous characteristic x that affects household consumption. Thus I have:

$$c_{1,r} = \Theta_{1,r} + \beta\frac{c_{2,r} + c_{3,r}}{2} \tag{12}$$

$$c_{2,r} = \Theta_{2,r} + \beta\frac{c_{1,r} + c_{3,r}}{2} \tag{13}$$

$$c_{3,r} = \Theta_{3,r} + \beta\frac{c_{1,r} + c_{2,r}}{2} \tag{14}$$

$$\Theta_{1,r} = \varpi x_{1,r} + \delta\frac{x_{2,r} + x_{3,r}}{2} + \xi_{1,r} \tag{15}$$

$$\Theta_{2,r} = \varpi x_{2,r} + \delta\frac{x_{1,r} + x_{3,r}}{2} + \xi_{2,r} \tag{16}$$

$$\Theta_{3,r} = \varpi x_{3,r} + \delta\frac{x_{1,r} + x_{2,r}}{2} + \xi_{3,r} \tag{17}$$

In this three-person social group setup, Equation (9) yields

$$\begin{aligned}
\frac{1}{m_r}\sum_{j \in M_r} c_{j,r} &= \frac{1}{1-\beta}\frac{1}{3}\left[\left(\varpi x_{1,r} + \delta\frac{x_{2,r} + x_{3,r}}{2} + \xi_{1,r}\right) + \left(\varpi x_{2,r} + \delta\frac{x_{1,r} + x_{3,r}}{2} + \xi_{2,r}\right)\right. \\
&\quad \left. + \left(\varpi x_{3,r} + \delta\frac{x_{1,r} + x_{2,r}}{2} + \xi_{3,r}\right)\right] \\
&= \frac{1}{1-\beta}\left[\varpi\frac{(x_{1,r} + x_{2,r} + x_{3,r})}{3} + \delta\frac{(x_{1,r} + x_{2,r} + x_{3,r})}{3} + \frac{(\xi_{1,r} + \xi_{2,r} + \xi_{3,r})}{3}\right] \qquad (18) \\
or,\ \overline{c}_r &= \frac{1}{1-\beta}\left[(\varpi + \delta)\overline{x}_r + \overline{\xi}_r\right] \qquad (19)
\end{aligned}$$

and Equation (11) yields

$$
\begin{aligned}
c_{1,r} &= \left(\varpi x_{1,r} + \delta \frac{x_{2,r} + x_{3,r}}{2} + \xi_{1,r}\right)\left(1 + \frac{\beta^2}{(1-\beta)(m_r - 1 + \beta)}\right) \\
&+ \frac{\beta}{m_r - 1 + \beta}\left[\left(\varpi x_{2,r} + \delta \frac{x_{1,r} + x_{3,r}}{2} + \xi_{2,r}\right) + \left(\varpi x_{3,r} + \delta \frac{x_{1,r} + x_{2,r}}{2} + \xi_{3,r}\right)\right] \quad (20)
\end{aligned}
$$

Thus, from the aggregate level Equation (19), a change in $\overline{c}_r$ due to a change in $\overline{\xi}_r$ is given by

$$
\frac{\partial \overline{c}_r}{\partial \overline{\xi}_r} = \frac{1}{1-\beta} \quad (21)
$$

and from the individual level Equation (20), a change in $c_{1,r}$ due to a change in $\xi_{1,r}$ is given by

$$
\frac{\partial c_{1,r}}{\partial \xi_{1,r}} = \left(1 + \frac{\beta^2}{(1-\beta)(m_r - 1 + \beta)}\right) \quad (22)
$$

For any household i, thus, I get:

$$
\frac{\partial c_{i,r}}{\partial \xi_{i,r}} = \left(1 + \frac{\beta^2}{(1-\beta)(m_r - 1 + \beta)}\right) \quad (23)
$$

Hence, the social multiplier Δ is given by:

$$
\Delta = \frac{\frac{\partial \overline{c}_r}{\partial \overline{\xi}_r}}{\frac{\partial c_{i,r}}{\partial \xi_{i,r}}} = \frac{\frac{1}{1-\beta}}{\left(1 + \frac{\beta^2}{(1-\beta)(m_r - 1 + \beta)}\right)} \quad (24)
$$

For sufficiently large social groups ($m_r \longrightarrow \infty$) :

$$
\Delta = \frac{1}{1-\beta} \quad (25)
$$

3 Alternative (Caste-based) Peer Groups

The importance of the caste system as a regulator of social interactions in India has been highlighted in various studies (e.g., Munshi, 2016). This could potentially imply that self-identification is stronger among households of the same caste living in the same region than among households living in the same region but belonging to different castes. To acknowledge this fact, I construct peer groups based on caste affiliation of households to check the sensitivity of my baseline analysis.

The ideal way to construct peer groups for households would be based on caste *and* village/neighborhood of residence. Unfortunately, I am unable to do so purely because of inadequate availability of data (more specifically, for the majority of villages/neighborhoods sampled in the IHDS 2012, there is very little within-village/neighborhood variation in caste of households). As an alternative, I define a household's peer group includes all other households of the household's caste living in its district.[1]

This alternative definition of peer group, however, has a serious limitation. This is due to the fact that the geographical area that districts typically represent are possibly too large for households to form comparator groups based upon. In other words, households may identify more with people of their own caste, but it is impossible for them to be influenced by other households who live in different villages/neighborhoods as it is unlikely that they ever socially interact. Nevertheless, it would be interesting to check the robustness of my main results to this alternative definition of peer group.

Table A4 reports the results. Across all specifications, I find strong evidence of endogenous peer effects in consumption. This is in line with my baseline results (reported in Table 4). My preferred specification, as in Table 4, is reported column (3) which regresses own consumption on average peer consumption controlling for own income, own characteristics, own idiosyncratic shocks, average peer income, average peer characteristics and district fixed effects. I find the estimate of endogenous peer effect to be 0.56 (s.e. = 0.277) which is statistically significant at 1% level of significance. Note that this estimate of endogenous peer effect is only slightly smaller than the one that I had obtained based on my preferred specification of the baseline regression model (where I had defined a household's peer group as all other households in its own village). This indicates that changing the definition of peer groups does note result in any qualitative change of my main results. Overall, thus, I find a consistent picture that is suggestive of strong endogenous peer effects in consumption.

4 Full Risk Sharing

I construct below a simple model of optimal risk allocation and derive the empirical specification implied by this model. Following Bardhan and Udry (1999), I examine an economy in which the village is the relevant risk sharing group. I also assume there is just one time period but multiple possible states of nature. Say there are m_r households in village r, indexed by i. Let M_r denotes the set of households in village r. Further, let there be W states of nature, indexed by w_τ, each with a probability of occurrence of $\pi(w_\tau)$ such that $\sum_{\tau=1}^{W} \pi(w_\tau) = 1$. Without loss of generality, by definition of state, the probabilities of occur-

[1] Under this new definition of peer groups, a household's peer group, on average, consists of 28 households.

rence of each state of nature do not vary by household. Income is exogenously given in each state of nature for household i, by $y_{i,r}(w_\tau)$. The households' utility in each state of nature still depends on consumption ($c_{i,r}(w_\tau)$) and bequests ($b_{i,r}(w_\tau)$) along with observable taste shifter or preference shock ($x_{i,r}(w_\tau)$) (say, demographic characteristics other than income) . The expected utility for household i is given by

$$U_{i,r} = \sum_{\tau=1}^{W} \pi(w_\tau) \left\{ u\left[c_{i,r}(w_\tau), x_{i,r}(w_\tau)\right] + v[b_{i,r}(w_\tau), x_{i,r}(w_\tau)] \right\} \tag{26}$$

Here the u and v depends positively on $c_{i,r}$ and $b_{i,r}$ respectively. Households are risk averse and share the same coefficient of constant absolute risk aversion. As in Mace (1991), preferences are homothetic and are state separable. The instantaneous sub-utility functions are given by

$$u\left[c_{i,r}(w_\tau), x_{i,r}(w_\tau)\right] = -\frac{1}{\sigma_c} \exp[-\sigma(c_{i,r}(w_\tau) - x_{i,r}(w_\tau))],\ \sigma_c > 0 \tag{27}$$

$$v\left[b_{i,r}(w_\tau), x_{i,r}(w_\tau)\right] = -\frac{1}{\sigma_b} \exp[-\sigma(b_{i,r}(w_\tau) - x_{i,r}(w_\tau))],\ \sigma_b > 0 \tag{28}$$

where σ_c and σ_b denotes the coefficient of risk aversion with respect to consumption and bequests respectively.

To consider the implication of risk-sharing, we can imagine a social planner who allocates a Pareto (programming) weight $\kappa_{i,r}$ to each household within village r and maximizes the weighted sum of expected utilities of the m_r households in the village or social group of individual i. To achieve *full* risk sharing with the village, the social planner must choose a risk allocation by maximizing the weighted sum of household utilities subject to constraints of non-negative consumption and total consumption equal to total income in each state. That is,

$$\begin{gathered} Max \sum_i \kappa_{i,r} U_{i,r} \\ s.t. \sum_i c_{i,r}(w_\tau) + \sum_i b_{i,r}(w_\tau) = \sum_i y_{i,r}(w_\tau) \forall w_\tau \\ and\ c_{i,r}(w_\tau), b_{i,r}(w_\tau) > 0\ \forall x_\tau \end{gathered}$$

where $0 < \kappa_{i,r} < 1$ and $\sum_{i=1}^{m_r} \kappa_{i,r} = 1$.

The corresponding Lagrangian can be written as

$$L = \sum_{i=1}^{m_r} \kappa_{i,r} \sum_{\tau=1}^{W} \pi(w_\tau) \left\{ u\left[c_{i,r}(w_\tau), x_{i,r}(w_\tau)\right] + v[b_{i,r}(w_\tau), x_{i,r}(w_\tau)] \right\} \tag{29}$$

$$+\Lambda \left[\sum_{i=1}^{m_r} y_{i,r}(w_\tau) - \sum_{i=1}^{m_r} c_{i,r}(w_\tau) - \sum_{i=1}^{m_r} b_{i,r}(w_\tau) \right] \tag{30}$$

From the first order conditions for this problem for two households i and j, I get

$$\frac{u'\left[c_{i,r}(w_\tau), x_{i,r}(w_\tau)\right]}{u'\left[c_{j,r}(w_\tau), x_{j,r}(w_\tau)\right]} = \frac{\kappa_{j,r}}{\kappa_{i,r}} \ \forall i, j, x_\tau, i \neq j \tag{31}$$

$$\frac{v'\left[b_{i,r}(w_\tau), x_{i,r}(w_\tau)\right]}{v'\left[b_{j,r}(w_\tau), x_{j,r}(w_\tau)\right]} = \frac{\kappa_{j,r}}{\kappa_{i,r}} \ \forall i, j, x_\tau, i \neq j \tag{32}$$

Substituting for the functional form of u and taking logs of the Equation (31), I get

$$\ln \left\{ \frac{\exp[-\sigma(c_{i,r}(w_\tau) - x_{i,r}(w_\tau))]}{\exp[-\sigma(c_{j,r}(w_\tau) - x_{j,r}(w_\tau))]} \right\} = \ln \left(\frac{\kappa_{j,r}}{\kappa_{i,r}} \right) \tag{33}$$

Solving for $c_{i,r}$,

$$c_{i,r} = c_{j,r} + x_{i,r} - x_{j,r} + \frac{1}{\sigma} \left(\ln \kappa_{i,r} - \ln \kappa_{j,r} \right) \tag{34}$$

Notice that a similar expression can be obtained with respect to any of the $J-1 = m_r - 1$ agents in the social group of person i:

$$\begin{aligned}
c_{i,r} &= c_{j+1,r} + x_{i,r} - x_{j+1,r} + \frac{1}{\sigma} \left(\ln \kappa_{i,r} - \ln \kappa_{j+1,r} \right) \\
c_{i,r} &= c_{j+2,r} + x_{i,r} - x_{j+2,r} + \frac{1}{\sigma} \left(\ln \kappa_{i,r} - \ln \kappa_{j+2,r} \right) \\
&\ldots \\
c_{i,r} &= c_{J-1,r} + x_{i,r} - x_{J-1,r} + \frac{1}{\sigma} \left(\ln \kappa_{i,r} - \ln \kappa_{J-1,r} \right)
\end{aligned}$$

Summing across the $J-1$ pairwise combinations of agent and dividing by $m_r - 1$, I get

$$c_{i,r} = \frac{\sum_{j \in M_{i,r}} c_{j,r}}{m_r - 1} + x_{i,r} - \frac{\sum_{j \in M_{i,r}} x_{j,r}}{m_r - 1} + \varsigma_{i,r} \tag{35}$$

where $\varsigma_{i,r} = \frac{1}{\sigma} \left(\ln \kappa_{i,r} - \frac{\sum_{j \in M_{i,r}} \ln \kappa_{j,r}}{m_r - 1} \right)$ is the household specific effect. Equation (35) implies household consumption perfectly co-moves with peer consumption. Moreover it also implies

that under full insurance household level income and idiosyncratic shocks should play no role in determining household consumption, once average village consumption has been controlled for. Thus as noted by Suri (2004) and Chiappori et al. (2014), a test for full risk sharing would be to estimate (35) with using household income as an additional regressor and testing the joint null hypothesis that the coefficient of endogenous peer effect is equal to one and that the coefficient of household income and idiosyncratic shocks are equal to zero. If the the null hypothesis is rejected, full risk sharing does not explain peer influences in consumption.

References

Bardhan, P. and Udry, C. (1999). *Development Microeconomics*, New York: Oxford University Press.

Chiappori, P., Samphantharak, K., Schulhofer-Wohl, S. and Townsend, R. (2014). 'Heterogeneity and risk sharing in village economies', *Quantitative Economics*, 5, 1–27.

Corneo, G. (2002). 'The efficient side of progressive income taxation', *European Economic Review*. 46, 1359–1368.

Glaeser, E. L. and Scheinkman, J. (2001). 'Measuring social interactions', In: S. Durlauf and H. P. Young (Eds.), *Social Dynamics*, 83–131, Cambridge: MIT Press.

Glaeser, E. L., Scheinkman, J. A. and Sacerdote, B. I. (2003). 'The Social Multiplier', *Journal of the European Economic Association*, 1, 345–353.

Mace, B. (1991). 'Full insurance in the presence of aggregate uncertainty', *Journal of Political Economy*, 99, 928–956. [

Munshi, K. (2016). 'Caste Networks in the Modern Indian Economy', In: S. Mahendra Dev and P. G. Babu (Eds.), *Development in India: Micro and Macro Perspectives*, 13-37, New Delhi: Springer.

Suri, T. (2004). 'Spillovers in Village Consumption: Testing the Extent of Partial Insurance', Working Paper, Sloan School of Management, Massachusetts Institute of Technology. Available at http://economie.esg.uqam.ca/upload/files/seminaires/suri_spillovers.pdf

Table A1. Categories of Consumption: IHDS 2012

Food	Vegetables	Soap/detergent	Therapeutic appliances
Rice	Salts/spices	Conveyance	Personal care
Wheat	Tea/Coffee	Diesel/Petrol/CNG	Other personal expenditure
Sugar	Processed food	House/other rent/ loans	Repair/maintenance
Kerosene	Paan/Tobacco	Consumer tax/fees	Insurance premiums
Other Cereal	Fruits/Nuts	Services/servants	Vacations
Pulses	**Temptation Good**	Clothing/bedding	Social Functions
Meat	Eating out	Footwear	**Health**
Sweeteners	Fuel	Furnitures/fixtures	Medical out-patient
Edible Oil	Light	Crockery/utensils	Medical in-patient
Eggs	Entertainment	Household appliances	**Education**
Milk	Telephone	Recreation goods	School/college fees
Milk products	Cosmetics/Toiletries	Jewelry	Private tuition
Cereal Products	Household items	Transport equipment	School books

Table A2. Peer Effects in Consumption: OLS Estimation (Full Results)

Variables	[1]	[2]	[3]
Endogenous Peer Effect	**0.461*** **	**0.568*** **	**0.208*** **
	(0.0132)	**(0.0173)**	**(0.0256)**
Individual Effects			
Income	0.283***	0.291***	0.291***
	(0.00838)	(0.00892)	(0.00893)
Household Size	8327***	8564***	8414***
	(352.4)	(357.5)	(355.7)
Age	60.11	83.41**	104.9**
	(40.62)	(40.94)	(41.33)
Male	-4531**	-4776**	-5317**
	(2156)	(2142)	(2147)
Literate	15047***	16211***	17011***
	(901.6)	(929.8)	(946.0)
Married	-1380	-1704	-2080
	(2329)	(2335)	(2341)
Teenage Proportion	53137***	52975***	52022***
	(3274)	(3264)	(3250)
Adult Proportion	34547***	33183***	31926***
	(3169)	(3161)	(3185)
Zero to Five Married	20652***	20505***	21472***
	(2127)	(2109)	(2124)
More than Five Married	34309***	33236***	33475***
	(4677)	(4657)	(4566)
Years in Place	-6196**	-3461	-3497
	(2767)	(2908)	(2928)
Urban	-48.03	1997	5144***
	(1191)	(1230)	(1713)
Brahmin	-11827*	-9402	-10226*
	(6206)	(6207)	(6133)
Forward Caste	-13286**	-6932	-7412
	(5823)	(5786)	(5697)
OBC	-17472***	-15945***	-16533***
	(5751)	(5728)	(5656)
SC	-29420***	-29953***	-30424***
	(5748)	(5682)	(5604)
ST	-28145***	-27050***	-27395***
	(5827)	(5925)	(5870)
Individual Idiosyncratic Shocks			
Job Loss	1035	1292	4282
	(2538)	(2552)	(2829)
Death	4767***	4579***	4015***
	(1174)	(1168)	(1171)
Theft	14865***	14721***	14807***
	(2741)	(2748)	(2789)

Table A2. Peer Effects in Consumption, OLS Full Results *(Continued)*

Variables	[1]	[2]	[3]
Break-in	130.0	549.5	923.7
	(5869)	(5837)	(5835)
Contextual Effects			
Income		-0.0953***	0.00343
		(0.0131)	(0.0154)
Household Size		-3805***	-3163***
		(978.5)	(1,184)
Age		-254.3*	119.4
		(149.9)	(162.1)
Male		-159.4	-10,961
		(8420)	(9229)
Literate		-12699***	7032*
		(2734)	(3704)
Married		864.4	-3,242
		(8066)	(9088)
Teenage Proportion		-5659	-3478
		(11852)	(12677)
Adult Proportion		7392	-1414
		(10709)	(12168)
Zero to Five Married		-9865	13214
		(8496)	(9302)
More than Five Married		-23304*	-6461
		(12862)	(14629)
Years in Place		-23773***	-25290***
		(7872)	(9472)
Brahmin		2890	-11266
		(11421)	(14217)
Forward Caste		-10770	-19472
		(10609)	(12849)
OBC		-1717	-15194
		(10398)	(12771)
SC		8984	-7720
		(10404)	(12553)
ST		628.2	-12260
		(10460)	(12583)
Constant	-43477***	15876	8212
	(7009)	(15756)	(19702)
District Fixed Effects	NO	NO	YES
Observations	40980	40980	40980
Adjusted R-squared	0.324	0.329	0.338

Notes: Dependent variable is Household Annual Total Consumption Expenditure. Endogenous Peer Effect represents the estimated coefficient of Average Peer Household Annual Total Consumption Expenditure. For definition of variables see Table 1. Heteroscedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A3. Peer Effects in Consumption: IV Estimation (First Stage and Second Stage Full Results)

Variables	[1]	[2]	[3]	[4]	[5]	[6]
Endogenous Peer Effect		**0.782*** **		**0.848*** **		**0.697*** **
		(0.147)		**(0.128)**		**(0.223)**
Individual Effects						
Income	0.102***	0.250***	0.0109***	0.289***	0.00547***	0.289***
	(0.00261)	(0.0166)	(0.00207)	(0.00885)	(0.00161)	(0.00887)
Household Size	-125.1	8,422***	130.2	8,528***	-112.1	8,453***
	(154.4)	(353.6)	(116.8)	(356.0)	(88.40)	(359.5)
Age	38.57*	49.91	-29.21*	92.31**	11.05	101.2**
	(21.63)	(42.45)	(16.41)	(40.79)	(12.08)	(41.52)
Male	-4000***	-3285	-474.3	-4802**	-952.1	-4860**
	(1167)	(2230)	(885.6)	(2152)	(679.7)	(2158)
Literate	5346***	13357***	-550.6	16381***	743.8**	16628***
	(514.3)	(1220)	(402.3)	(934.2)	(298.8)	(956.5)
Married	1209	-1530	409.8	-1696	-148.4	-1985
	(1148)	(2349)	(851.2)	(2335)	(650.9)	(2341)
Teenage Proportion	6320***	50861***	3235***	51501***	608.1	51336***
	(1607)	(3405)	(1212)	(3273)	(902.3)	(3201)
Adult Proportion	6138***	32663***	3631***	31987***	413.5	31609***
	(1585)	(3241)	(1215)	(3164)	(905.6)	(3190)
Zero to Five Married	-206.3	20181***	449.5	20124***	1465**	20742***
	(1140)	(2155)	(878.3)	(2116)	(669.4)	(2157)
More than Five Married	-5906***	36090***	-576.2	33230***	-15.54	33315***
	(1960)	(4744)	(1500)	(4666)	(1122)	(4586)
Years in Place	-18049***	-1310	-4382***	-3122	-2360**	-2901
	(1652)	(3906)	(1245)	(2957)	(1004)	(2939)
Urban	34910***	-11428**	4710***	633.3	6347***	2105
	(5639)	(5197)	(492.0)	(1334)	(529.2)	(2175)
Brahmin	4327*	-13135**	-500.9	-9021	-1165	-9415
	(2564)	(6389)	(2204)	(6255)	(1645)	(6196)
Forward Caste	-54.74	-13166**	-2462	-6233	-1824	-6561
	(2288)	(5919)	(2000)	(5815)	(1480)	(5766)
OBC	-9982***	-14373**	-1799	-15603***	-1591	-15789***
	(2246)	(5951)	(1969)	(5771)	(1468)	(5722)
SC	-8592***	-26630***	-1439	-29471***	-1256	-29738***
	(2271)	(5914)	(1986)	(5724)	(1475)	(5670)
ST	-22457***	-20986***	-2375	-26743***	-1589	-26562***
	(2356)	(6596)	(2093)	(5979)	(1540)	(5914)
Individual Idiosyncratic Shocks						
Job Loss	-81.09	-83.47	-1,172	1,532	1,224	3,254
	(1608)	(2546)	(1196)	(2567)	(890.1)	(2905)
Death	1875***	4036***	1605***	3993***	441.5	3713***
	(634.7)	(1253)	(480.3)	(1218)	(358.0)	(1187)

Table A3. IV Full Results *(Continued)*

Variables	[1]	[2]	[3]	[4]	[5]	[6]
Theft	-1181	15037***	60.00	14329***	-342.7	14914***
	(1288)	(2758)	(931.4)	(2762)	(680.8)	(2801)
Break-in	5926**	-2761	3307*	-1328	2580	-1006
	(2711)	(6053)	(1997)	(5898)	(1573)	(5877)
Contextual Effects						
Income			0.445***	-0.221***	0.369***	-0.179**
			(0.00500)	(0.0576)	(0.00539)	(0.0832)
Household Size			10608***	-6887***	6699***	-6436***
			(410.7)	(1637)	(410.1)	(1829)
Age			-357.2***	-167.3	250.4***	-8.880
			(60.46)	(157.0)	(51.40)	(176.3)
Male			-14157***	3879	-21046***	-179.0
			(3387)	(8621)	(3121)	(10325)
Literate			9868***	-15997***	30500***	-8574
			(1099)	(3034)	(1091)	(7942)
Married			7200**	1137	-1857	-866.5
			(3447)	(8066)	(3116)	(9133)
Teenage Proportion			100107***	-30960*	58047***	-30183*
			(4506)	(17768)	(3878)	(17659)
Adult Proportion			85786***	-14571	35349***	-17559
			(4322)	(15082)	(4014)	(13588)
Zero to Five Married			27029***	-17062*	42488***	-7815
			(3402)	(9074)	(3070)	(13101)
More than Five Married			24316***	-29606**	32488***	-23375
			(5930)	(13234)	(5126)	(16598)
Years in Place			-60641***	-7443	-33491***	-9751
			(3207)	(11236)	(3481)	(12605)
Brahmin			-15528***	6391	-24836***	1302
			(4986)	(11525)	(4331)	(14823)
Forward Caste			-39230***	-79.35	-30292***	-3488
			(4647)	(11448)	(3949)	(13967)
OBC			-39395***	8925	-36399***	3759
			(4592)	(11248)	(3872)	(14405)
SC			-47856***	21773*	-45272***	15216
			(4613)	(11668)	(3878)	(15240)
ST			-58802***	17160	-46567***	11444
			(4664)	(12244)	(3921)	(15611)
Peer Idiosyncratic Shocks						
Job Loss	3080		-10012***		24826***	
	(2974)		(2169)		(3059)	
Death	20410***		26768***		8794***	
	(2180)		(1680)		(1553)	

Table A3. IV Full Results *(Continued)*

Variables	[1]	[2]	[3]	[4]	[5]	[6]
Theft	4125		14160***		8228***	
	(3942)		(2898)		(2821)	
Break-in	70519***		51949***		41407***	
	(8989)		(6379)		(6093)	
Constant	104204***	-77597***	32269***	5243	3982	2296
	(3,185)	(17188)	(6497)	(16488)	(6219)	(20185)
District Fixed Effects	NO	NO	NO	NO	YES	YES
Observations	40980	40980	40980	40980	40980	40980
Adjusted R-squared	0.259	0.305	0.576	0.321	0.766	0.325
Hansen J statistic		6.096		4.355		1.829
		[p=0.107]		[p=0.226]		[p=0.609]
Kleibergen-Paap rk LM statistic		147.5		339.9		154.8
		[p=0.000]		[p=0.000]		[p=0.000]
First stage F-statistic		38.03		89.97		40.01

Notes: Estimation via two-step GMM. Dependent variable is Household Annual Total Consumption Expenditure. Endogenous Peer Effect represents the estimated coefficient of Average Peer Household Annual Total Consumption Expenditure. For full list of demographic characteristic, idiosyncratic shocks and definition of variables see Table 1 and note below Table 2. Heteroscedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A4. Peer Effects in Consumption: Caste-based Peer Groups

Variables	[1]	[2]	[3]
Endogenous Peer Effect	**0.757*** **	**0.844*** **	**0.559*** **
	(0.137)	**(0.088)**	**(0.277)**
Individual Effects	YES	YES	YES
Idiosyncratic Shocks	YES	YES	YES
Contextual Effects	NO	YES	YES
District Fixed Effects	NO	NO	YES
Observations	40629	40629	40629
Adjusted R-squared	0.309	0.320	0.332
Hansen J statistic	17.349	4.832	1.749
	[p=0.000]	[p=0.185]	[p=0.626]
Kleibergen-Paap rk LM statistic	185.9	758.9	96.075
	[p=0.000]	[p=0.000]	[p=0.000]
First stage F-statistic	44.84	179.87	28.93

Notes: Estimation via two-step GMM. Dependent variable is Household Annual Total Consumption Expenditure. Endogenous Peer Effect represents the estimated coefficient of Average Peer Household Annual Total Consumption Expenditure. Individual effects are controlled by own demographic characteristics of households. Contextual effects are controlled by average peer demographic characteristics. For full list of demographic characteristic, idiosyncratic shocks and definition of variables see Table 1 and note below Table 2. Heteroscedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A5. Partial Risk Sharing: Subsample Analysis based on (Relative) Group Vulnerability

	Subsamples					
Variables	More Vulnerable: Literacy	Less Vulnerable: Literacy	More Vulnerable: Caste	Less Vulnerable: Caste	More Vulnerable: Poverty	Less Vulnerable: Poverty
Endogenous Peer Effect	**0.806*** **	**0.897*** **	**0.624*** **	**0.832** **	**0.629** **	**0.767*** **
	(0.193)	**(0.275)**	**(0.224)**	**(0.323)**	**(0.261)**	**(0.267)**
Observations	20403	20577	20584	20396	20235	20745
Adjusted R-squared	0.178	0.190	0.170	0.217	0.236	0.177
Hansen J statistic	0.376	1.712	0.923	1.371	0.876	0.883
	[p=0.945]	[p=0.634]	[p=0.820]	[p=0.712]	[p=0.831]	[p=0.830]
Kleibergen-Paap rk LM statistic	308.1	116.6	153.5	77.80	195.7	113.1
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]
First stage F-statistic	94.37	30.34	41.26	19.79	50.22	30.25

Notes: Estimation via two-step GMM. Dependent variable is Household Annual Total Consumption Expenditure. Endogenous Peer Effect represents the estimated coefficient of Average Peer Household Annual Total Consumption Expenditure. Subsample construction: *More Vulnerable: Literacy* includes those groups (villages/neighborhoods) with proportion of illiterate households above the sample group-median level (where the group median level is calculated as the median of the proportion of illiterate households over all groups in the sample). *Less Vulnerable: Literacy* includes those groups with proportion of illiterate households below the sample group-median level. *More Vulnerable: Caste* includes those groups with proportion of low caste households above the sample group-median level. *Less Vulnerable: Caste* includes those groups with proportion of low caste households below the sample group-median level. *More Vulnerable: Poverty* includes those groups with proportion of below poverty line households above the sample group-median level. *Less Vulnerable: Poverty* includes those groups with proportion of below poverty households below the sample group-median level. All specifications include controls for individual effects, own idiosyncratic shocks, contextual effects and district fixed effects. Individual effects are controlled by own demographic characteristics of households. Contextual effects are controlled by average peer demographic characteristics. For full list of demographic characteristic, idiosyncratic shocks, definition of variables and definition of each subsample see Table 1, note below Table 2 and main text. Heteroscedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A6. Partial Risk Sharing: Subsample Analysis based on (Relative) Group Heterogeneity

	Subsamples					
Variables	More Heterogeneous: Literacy	Less Heterogeneous: Literacy	More Heterogeneous: Caste	Less Heterogeneous: Caste	More Heterogeneous: Poverty	Less Heterogeneous: Poverty
Endogenous Peer Effect	**0.950*** **	**0.742*** **	**1.140*** **	**0.644*** **	**0.731* **	**0.737*** **
	(0.280)	**(0.228)**	**(0.415)**	**(0.223)**	**(0.415)**	**(0.235)**
Observations	16938	24042	13586	27394	9516	31464
Adjusted R-squared	0.180	0.187	0.155	0.205	0.204	0.206
Hansen J statistic	1.973	0.159	0.618	2.938	1.620	1.781
	[p=0.578]	[p=0.984]	[p=0.892]	[p=0.401]	[p=0.655]	[p=0.619]
Kleibergen-Paap rk LM statistic	94.92	197.5	50.56	138.5	108.1	136.8
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]
First stage F-statistic	25.58	51.44	13.54	36.56	30.06	35.07

Notes: Estimation via two-step GMM. Dependent variable is Household Annual Total Consumption Expenditure. Endogenous Peer Effect represents the estimated coefficient of Average Peer Household Annual Total Consumption Expenditure. Subsample construction: *More Heterogeneous: Literacy* includes those groups (villages/neighborhoods) with proportion of illiterate households between 0.25 and 0.75. *Less Heterogeneous: Literacy* includes those groups with proportion of illiterate households below 0.25 or above 0.75. *More Heterogeneous: Caste* includes those groups with proportion of low caste households between 0.25 and 0.75. *Less Heterogeneous: Caste* includes those groups with proportion of low caste households below 0.25 or above 0.75. *More Heterogeneous: Poverty* includes those groups with proportion of below poverty line households between 0.25 and 0.75. *Less Heterogeneous: Poverty* includes those groups with proportion of below poverty households below 0.25 or above 0.75. All specifications include controls for individual effects, own idiosyncratic shocks, contextual effects and district fixed effects. Individual effects are controlled by own demographic characteristics of households. Contextual effects are controlled by average peer demographic characteristics. For full list of demographic characteristic, idiosyncratic shocks, definition of variables and definition of each subsample see Table 1, note below Table 2 and main text. Heteroscedasticity robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.